THE EXAMINATION OF THE PROTEASE INHIBITOR ECOTIN AND N-LINKED GLYCOSYLATION IN *CAMPYLOBACTER* SPECIES, AN INSIGHT INTO CELL PROTECTION IN THE PROTEASE RICH ENVIRONMENT OF THE HUMAN HOST

by

CODY LOWELL CHARLES THOMAS

(Under the Direction of Christine Szymanski)

ABSTRACT

Protection from proteolysis is important for bacteria living in a competitive niche under attack by host defense mechanisms. *Campylobacter* species possess an N-linked protein glycosylation system, where an oligosaccharide is attached to asparagine residues or released as free oligosaccharide. The glycosylation of these proteins can stabilize proteins such as the multidrug efflux pump of *C. jejuni* that helps efflux antibiotics and toxic compounds.

Glycosylation can also mask cleavage sites protecting against proteolytic activity. Oral *Campylobacter* species have also developed an additional method for protection against proteases. Ecotin homologues have been identified in these non-thermophilic *Campylobacter* species. Ecotin is a generic serine-protease inhibitor that can bind and inhibit a wide range of proteases released by the host. The combination of N-glycosylation and ecotin helps provide *Campylobacter* species inhabiting the protease-rich oral cavity with additional mechanisms for bacterial fitness and survival in the human host.

INDEX WORDS: *Campylobacter jejuni*, *Campylobacter rectus*, *Campylobacter showae*, ecotin, N-glycosylation, free oligosaccharide, protease inhibition, neutrophil elastase

THE EXAMINATION OF THE PROTEASE INHIBITOR ECOTIN AND N-LINKED GLYCOSYLATION IN *CAMPYLOBACTER* SPECIES, AN INSIGHT INTO CELL PROTECTION IN THE PROTEASE RICH ENVIRONMENT OF THE HUMAN HOST

By

CODY LOWELL CHARLES THOMAS

B.S., University of Alberta, 2015

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2018

The examination of the protease inhibitor ecotin and N-linked glycosylation in *Campylobacter* species, an insight into cell protection in the protease rich environment of the human host

By

Cody Lowell Charles Thomas

Major Professor: Christine Szymanski

Committee: Robert Maier

Kelly Moremen

Electronic Version Approved:

Suzanne Barbour

Dean of the graduate School

The University of Georgia

August 2018

Dedication

This dissertation is dedicated first and foremost to myself. I never would have expected myself at this juncture of my life if I did not push myself. I also dedicate this to my parents: Lowell Thomas and Verene Thomas, for all their love, patience and wisdom growing up in their care.

Acknowledgements

I would like to express my sincere gratitude to Dr. Christine M. Szymanski for allowing me to conduct my thesis research in her laboratory. Allowing me to transfer with her to the University of Georgia to complete my research at the Complex Carbohydrate Research Center. This work would not be completed if not for her advice and guidance. She is one of the top in her field and I consider myself extremely privileged to have had worked with her and learn from her. I had a wonderful experience throughout my graduate degree.

I would like to thank Dr. Harald Nothaft for being a mentor to me throughout my degree. Harald is a well of knowledge and experience that has taught me almost everything I know in the laboratory. Harald was always there to answer questions and help develop methods to answering questions. I have learnt so much from my mentor Dr. Nothaft.

I would like to thank my co-supervisors, Dr. Robert Maier and Dr. Kelley Moremen for providing feedback and useful suggestions throughout the duration of my degree.

I would like to thank other members of the Szymanski laboratory: Robert Patry, Clay Crippen, Corey Wenzel, Justin Duma, Silke Andresen, Jolene Garber, Jessica Sacher as well as other past and present members. I am thankful to have worked by all your sides in the laboratory.

I would like to thank Dr. Balaz Rada and his laboratory. Collaboration with the Rada lab was a major part of my research. I enjoy every day I would have to go work with neutrophils in the Rada lab, everyone was always so friendly and helpful.

Thank you to everyone for making my graduate studies at the University of Alberta and the University of Georgia so memorable and enjoyable.

Table of contents

List of tables

List of figures

List of abbreviations

Amp	Ampicillin
AtpA	Adenosine triphosphate synthase subunit alpha
BHI	Brain heart infusion
Cm	Chloramphenicol
diNAcBac	di-N-acetylbacillosamine
Dol-P	Dolichol phosphate
Ec	Ecotin
ENGase	Endo-β-N-acetylglucosaminidase
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum associated degradation
fOS	Free oligosaccharides
Fuc	Fucose
FRET	Fluorescence resonance energy transfer
HPAEC-PAD	High performance anion exchange chromatography with pulsed amperometric detection
Hex	Hexose
HexA	Hexuronic aid
HexNAc	N-acetylated hexose
HexNAcA	N-acetyl hexuronic acid
Kan	Kanamycin

Gal	Galactose
GalNAc	N-acetylgalactosamine
GlcNAc	N-acetylglucosamine
Glc	Glucose
kDa	Kilodalton
LC	Liquid chromatography
LLO	Lipid-linked oligosaccharide
LB	Luria-Bertani
LOS	Lipooligosaccharide
MALDI	Matrix assisted laser desorption ionization
MEM	Minimal essential medium
MH	Mueller Hinton
mM	Millimolar
MS	Mass spectrometry
N	Asparagine
NaCl	Sodium Chloride
OTase	Oligosaccharyltransferase
P	Phosphate
PBS	Phosphate buffered saline
Pgl	Protein glycosylation
PNGase	Peptide: N-glycanase
R_f	Distance traveled by compound/ distance traveled by solvent

R_t	Retention time
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TFA	Trifluoracetic acid
TLC	Thin layer chromatography
TOF	Time of flight
Und-PP	Undecaprenyl pyrophosphate
Und-P	Undecaprenyl phosphate
WCP	Wet cell pellet
WT	Wildtype

Preface

Chapter 1 is my original research into the background and introduction of my thesis.

Chapter 2 is my original work and is in the process of being prepared for submission to *PLoS Pathogens* in July 2018. I was responsible for designing the experiments and developing the methods. In addition, I performed the data collection, and analysis of experiments involving the cloning, purification and characterization of the ecotin homologues. I designed experiments to investigate the broad inhibitory range of the ecotins for serine-proteases, involving CmeA protection assays (for trypsin activity), Fluorescence resonance energy transfer (FRET) assays (for factor Xa activity) developed by Dr. Abofu Alemka and enzyme linked immunosorbent assays (ELISA) (for neutrophil elastase activity). Dr. Harald Nothaft created the ecotin-deficient *E. coli* strain that I subsequently used for complementation and live cells assays. To complete live cell neutrophil killing assays, Dr. Ruchi Yadav provided the technical training. I was responsible for manuscript preparation and all co-authors provided their input in the editing process.

Chapter 3 is my original work. I was responsible for method development, involving thin layer chromatography (TLC), porous graphite carbon chromatography, sample preparation for mass spectrometry and high anion exchange with pulsed amperometric detection (HPAEC-PAD). Technical training of HPAEC-PAD equipment was provided by Dr. Harald Nothaft. All samples for mass spectrometry were prepared by me. Dr. Artur Muszynski assisted with all mass spectrometry data generation and analysis, as well as mass spectrometry figure preparation.

Chapter 4 is my original work, I was responsible for developing methods and designing experiments to further on my research for future discovery.

Appendix 1 has been published and reproduced with permission from *Glycobiology,* 2017, vol 27, no. 10, 978-989 as “A conserved DGGK motif is essential for the function of the PglB oligosaccharyltransferase from *Campylobacter jejuni*”. This chapter was the original work of Yasmin Barre and Dr. Harald Nothaft. I was responsible for the cloning of the PglB homologues from *C. curvus* and *C. gracilis*. I assisted in altering the amino acid sequences of the PglB homologues to test the DGGK motif of these *Campylobacter* species. I was also responsible for editing of the manuscript.

Chapter I

General Introduction

1.1 Overview of protein glycosylation

Modification of proteins occurs across all domains of life. These modifications include lipidation, methylation, acetylation, phosphorylation, ubiquitination, proteolysis and glycosylation [1-3]. Glycosylation is one of the most common post-translational modifications of proteins. Glycans have been shown to be important for structural and functional roles for proteins, including: proper folding, protein stability and localization, protection from proteolysis, and evasion from antigenic recognition [4,5]. Oligosaccharides are important for the development, growth and survival of an organism[6]. Protein glycosylation involves the attachment of a carbohydrate unit covalently to the side chain of a specific amino acid residue. There have been various glycosylation types identified in several organisms including: N-glycosylation, O-glycosylation, C-mannosylation, rhamnosylation and glypiation [2].

It has been predicted that more than half of all proteins found in nature are glycosylated and that over 70% of the eukaryotic proteome is glycosylated. Compared to other major biological molecules like proteins, nucleic acids and lipids, there is still much to learn about protein glycosylation in bacteria. The first evidence of protein glycosylation was found in 1938 in egg albumin [7]. It was then discovered that it was an N-acetylglucosamine (GlcNAc) linked to an asparagine (N) in ovalbumin, and since then several carbohydrate modifications on various functional groups on proteins have been described across all domains of life [8].

N- and O- linked protein glycosylation are the most abundant types of post-translational modifications, and are present in all domains of life [9]. O-glycosylation occurs when a carbohydrate is linked through the hydroxyl group of an amino acid [2,9]. O-glycosidic linkages have been reported for every amino acid containing a hydroxyl group, such as serine (S), threonine (T), lysine (K), proline (P), tryptophan (W) and tyrosine (Y), although the first two amino acids are the most commonly modified [2,10-13]. In eukaryotes, proteins can be glycosylated with GlcNAc, N-acetylgalactosamine (GalNAc), mannose (Man), fucose (Fuc) and xylose (Xyl) [10]. In eukaryotic cells, O-glycosylation starts in the Golgi apparatus where the sugars are added sequentially to the hydroxyl groups of exposed serines and threonines [11]. The best studied O-linked glycoproteins are surface-layer proteins [14]. It has been reported that O-linked glycosylation occurs in bacteria with the glycosylation of pilin subunits or flagellar proteins [10,15,16]. In bacteria, this process occurs in the cytoplasm or inner membrane, where nucleotide-activated sugars are attached to S or T residues [17].

N-linked glycosylation involves the attachment of a carbohydrate to the amide nitrogen of asparagine (N). There have been multiple proteins reported to be N-glycosylated in various organisms across all domains of life. N-glycosylation has been studied most in eukaryotes, in particular *Saccharomyces cerevisiae* [18,19]. It was then observed to exist outside of the eukaryotic domain in the archaeon *Halobacterium salinarium* [20]. It was then discovered by Szymanski *et al.* (1999) that the Gram-negative bacterium *Campylobacter jejuni* was capable of N-glycosylation [21]. The two major glycosylation pathways in bacteria are discussed here. The O-linked and N-linked glycosylation pathways and their biological importance in bacterial species is discussed below in more detail.

1.2 Overview of O-linked glycosylation

1.2.1 Eukaryotic O-linked glycosylation

Eukaryotic O-linked glycosylation is commonly initiated in the Golgi apparatus, but it can also occur in the cytoplasm. The most common modification is the transfer of GalNAc onto the hydroxyl group of serine and threonine residues. The glycosylation is completed through a step-wise manner (Fig 1.1.B) adding nucleotide-activated sugars to yield several core structures [22,23].

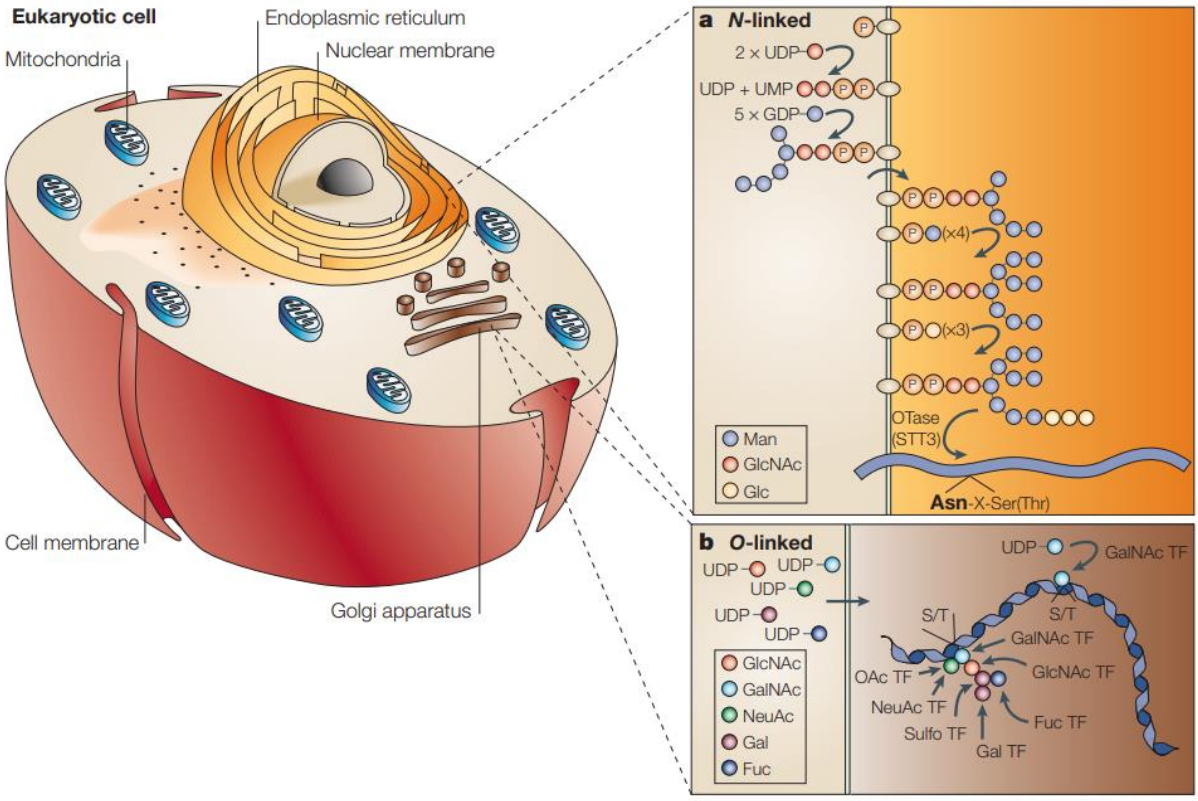


Fig. 1.1. The eukaryotic N- and O-linked glycosylation system

A) N-linked glycosylation begins with the synthesis of dolichol-linked oligosaccharides in the cytosol of the luminal membrane of the endoplasmic reticulum (ER) The heptasaccharide is then

flipped into the lumen of the ER. After translocation into the lumen, the heptasaccharide is further processed to a complete tetradecasaccharide, where it then gets transferred to an asparagine residue within the N-X-S/T consensus sequence by the oligosaccharyltransferase complex (OTase) complex. N-linked glycosylated proteins and free oligosaccharide (fOS) are released into the ER lumen. **B)** O-linked glycosylation happens in the Golgi apparatus. Nucleotide activated sugars are added in a step-wise manner on the hydroxyl residues of serines and threonines. Figure reproduced from Szymanski and Wren (2005) with permission [24] .

1.2.2 Bacterial O-linked glycosylation

1.2.2.1 The classical pathway (sequential transfer)

Since the discovery of bacterial O-glycosylation, it has become the most studied type of protein glycosylation in bacteria [25]. The common mechanism of O-linked glycan addition to proteins is completed in the classical pathway (Fig 1.2.A)[26]. The classical pathway is similar to the method observed in eukaryotes. In this method, the glycosyltransferases directly add nucleotide-activated sugars sequentially onto the hydroxyl group of serine and threonine residues of proteins. This is typically completed in the cytoplasmic space [27,28]. There are bacterial pathogens, including *Campylobacter* and *Helicobacter* species, that modify their flagellins with the addition of sugars to flagellar proteins by the classical pathway of O-glycosylation [29-31].

1.2.2.2 The non-classical pathway (bloc transfer)

The other method of O-linked glycosylation, known as the non-classical pathway (bloc transfer), is used to modify pilin proteins of some bacteria. Examples of organisms that O-glycosylate pilin by this mechanism include *Pseudomonas aeruginosa*, *Neisseria meningitidis* and *Neisseria gonorrhoeae.* The non-classical pathway is slightly different from the classical pathway. In the non-classical pathway, oligosaccharides are assembled onto a lipid carrier undecaprenyl-phosphate by pilin glycosylation enzymes in the cytoplasm [32,33]. The lipid-linked oligosaccharides are flipped into the periplasmic space, where the preassembled glycans are transferred *en bloc* onto the pilin subunits (Fig 1.2.B)**.** The non-classical pathway of glycan attachment occurs only at surface-exposed serine and theorine residues.

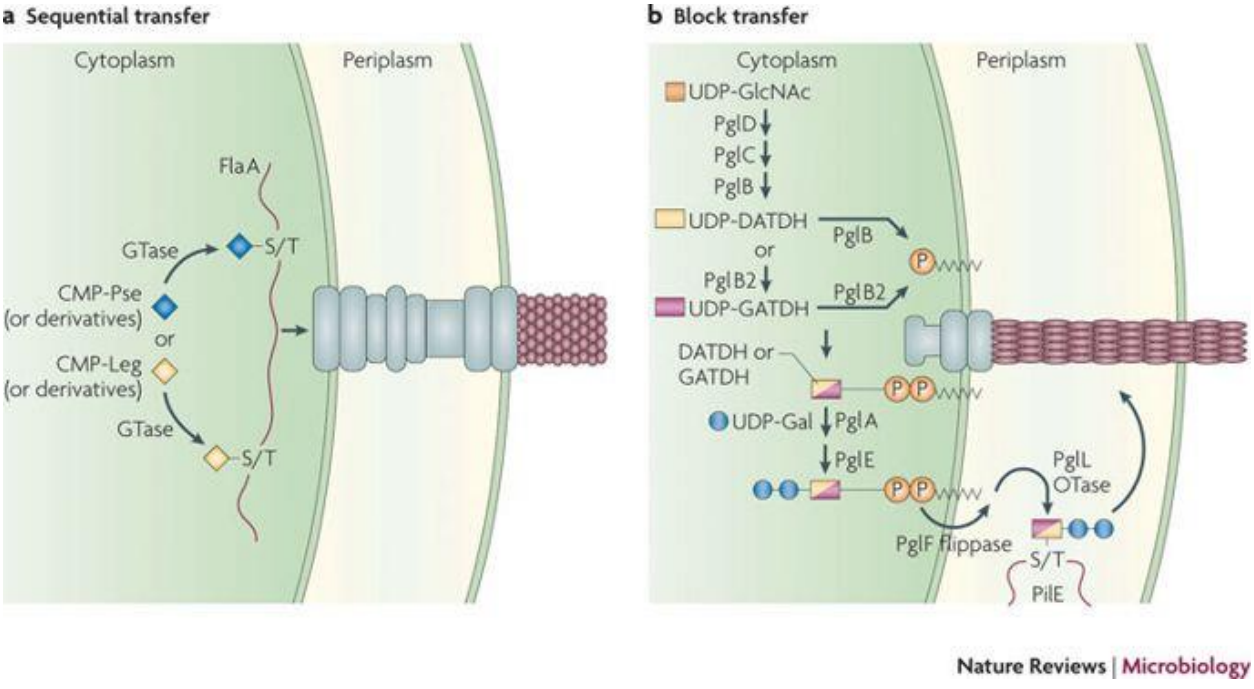


Fig 1.2. Bacterial O-linked glycosylation. A) In the classical pathway, nucleotide-activated sugars are sequentially transferred onto proteins. B) In the non-classical pathway, an oligosaccharide is built on a lipid carrier (undecaprenyl phosphate). The lipid-linked sugar is then flipped into the periplasmic space by a flippase (PglF, in *Neisseria*), and then the oligosaccharide is transferred to surface-exposed serine and theorine residues by an oligosaccharyltransferase (PglL, in *Neisseria*) of pilin subunits. Figure reproduced from Nothaft and Szymanski (2010) with permission[34] .

1.3 Overview of N-linked protein glycosylation

N-linked protein glycosylation has been shown to be present in all three domains of life, and involves the attachment of an oligosaccharide to the amide nitrogen of asparagine within a specific protein sequence. In eukaryotes, the oligosaccharide is assembled on a lipid carrier anchored to the cytosolic side of the endoplasmic reticulum (ER) and then flipped into the ER lumen (Fig. 1.2.A). In bacteria, the oligosaccharide is assembled on a lipid carrier anchored to the cytoplasmic side of the inner membrane by the addition of nucleotide-activated monosaccharides, and is then flipped into the periplasmic space (Fig. 1.3). In eukaryotes, the oligosaccharide is further modified and then transferred *en bloc* and covalently attached to the proteins by an oligosaccharyltransferase (OTase) complex that contains the conserved catalytic Stt3 subunit. In bacteria a homologue to the Stt3 performs this function [35]. Additionally, the OTase has a hydrolase activity and can release the sugar as a free oligosaccharide (fOS) into the ER of eukaryotes or the periplasmic space in bacteria. It has also been shown that in eukaryotes, fOS can be created by degradation of glycoproteins in the cytosol, and while fOS has been observed in bacteria [26], it is unknown whether glycoproteins are recycled and fOS released by the

same mechanism in bacteria . Across all domains of life, OTases contain a WWDYG motif, which is essential for the generation of N-linked glycoproteins.

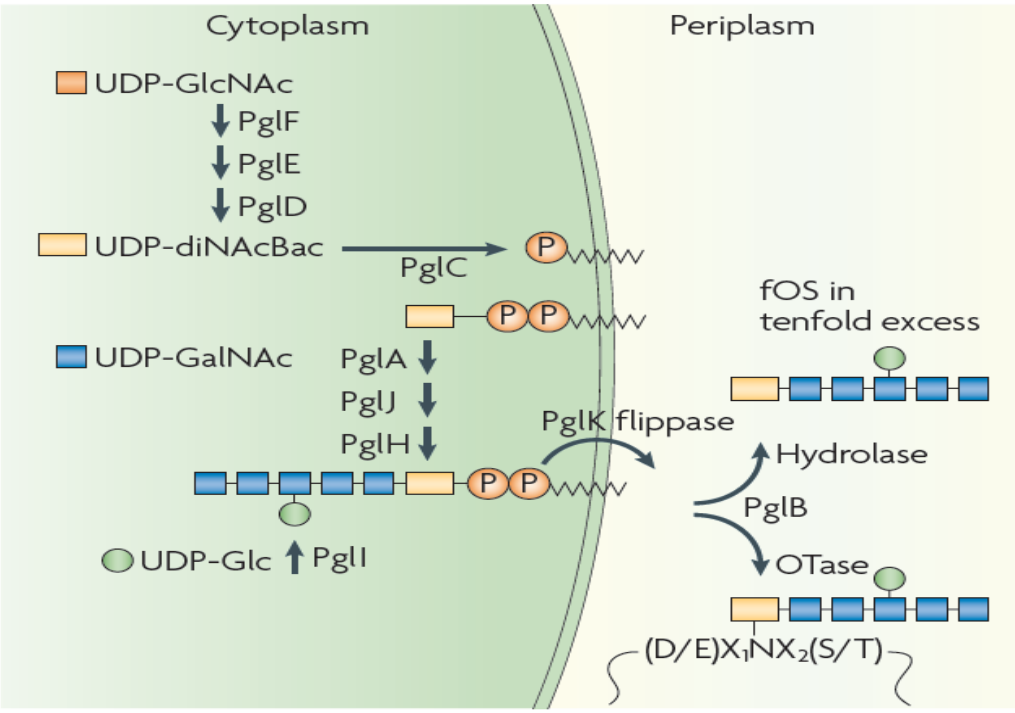


Fig. 1.3. The N-linked glycosylation pathway in *Campylobacter jejuni.* In *C. jejuni*, N-glycosylation occurs in the periplasm, where the heptasaccharide is assembled on a lipid carrier (undecaprenyl phosphate) attached to the cytoplasmic side of the inner membrane. Where by sequential action of Pgl glycosyltransferases (PglF, E, D, A, J, H, I) builds the heptasaccharide. The lipid-linked oligosaccharide (LLO) is then flipped into the periplasmic space by PglK, a flippase. The heptasaccharide is then transferred onto the amide residue of asparagine of proteins with the sequon D/E-X-N-X-S/T (where X is any amino acid except proline) by the OTase, PglB. PglB can also hydrolyze the LLOs and generates free oligosaccharide (fOS), which is

released into the periplasm in a free: protein-linked ratio of approximately 10:1 [26,36]. diNAcBac is shown as a yellow rectangle. Figure reproduced from Nothaft and Szymanski (2010) with permission[34] .

1.3.1 Bacterial N-linked protein glycosylation

N-linked protein glycosylation was first discovered in *C. jejuni* [37]. The N-linked protein glycosylation system of *C. jejuni* is modification that occurs in the periplasmic space. The *pgl* gene locus (Fig 1.3) encodes three sugar-modifying enzymes and five glycosyltransferases. The pathway begins on the cytosolic side of the inner membrane, where a UDP-GlcNAc is converted to UDP-diNAcBac by PglF, a dehydratase; PglE, an aminotransferase; and PglD, an acetyltransferase [38]. PglC then transfers the diNAcBac-P from the UDP-diNAcBac and attaches it to the lipid carrier Und-P, which produces Und-PP-diNAcBac. The oligosaccharide is then built up on the diNAcBac by PglA, PglH and PglJ, which add a total of five GalNAc residues to produce a hexasaccharide. Finally, PglI transfers a glucose branch to complete the heptasaccharide. A transporter protein, PglK then flips the heptasaccharide into the periplasm [39]. The OTase PglB then transfers the heptasaccharide to the N of the consensus sequence D/E-X-N-X-S/T. PglB also contains a hydrolase activity, and releases the heptasaccharide as fOS at a 10:1 ratio of free to protein-linked glycans. It has been proposed that more than 150 *C. jejuni* proteins are modified by the *pgl* pathway, a prediction which is based on the presence of the N-glycosylation consensus sequence as well as the signal peptide required to target the protein to the periplasm, where the Pgl proteins are located [34].

1.3.2 Discovery of the N-glycosylation system in *C. jejuni*

Bacterial O-linked protein glycosylation has been studied in different bacterial species heavily, but the presence of N-linked protein glycosylation was not discovered until 1999 [21]. This discovery was made in the Gram-negative bacterium, *C. jejuni*. *C. jejuni* is a human gastrointestinal pathogen and is the leading cause of foodborne bacterial gastroenteritis in the world. In some cases, campylobacteriosis can lead to the auto-immune disease Guillain-Barré syndrome [40]. Other diseases, such as reactive arthritis and irritable bowel syndrome have also been linked to *Campylobacter* infections [41]. *C. jejuni* is a bacterium of interest as it is known to both O- and N-glycosylate its proteins [42-44]. The glycosylation systems of *C. jejuni* have been shown to play major roles in infection and pathogenesis of this bacterium [21,45]. The N-linked pathway is responsible for the modification of more than 60 proteins [46], and mutations in the pathway result in multiple phenotypes, from reduced colonization of the intestinal tract of chicken and mouse models to reduced adherence and invasion of human epithelial cells [5,47-53]. This makes *C. jejuni* an excellent model organism for bacterial carbohydrate studies.

1.3.3 Diversity of N-glycosylation between *Campylobacter* species

Several epsilonproteobacteria (and all *Campylobacter* species) have been reported to contain N-glycosylation systems [21,54]. The entire *Campylobacter* genus was analyzed and shown to N-glycosylate proteins with an oligosaccharide containing a conserved diNAcBac on the reducing end and a varied non-reducing end (Fig. 1.4) [54]. Like *C. jejuni*, all 29 *Campylobacter* species synthesize their oligosaccharide on undecaprenyl phosphate by the enzymes encoded by

the protein glycosylation (*pgl*) genes, which are clustered together (Fig. 1.4) [34,54,55]. The oligosaccharide of *C. jejuni* has been shown to protect glycoproteins against proteases found in the chicken gut by blocking the cut sites of these enzymes [5]. The non-reducing end of the *C. jejuni* heptasaccharide is a GalNAc, the terminal sugar may also play a role in reducing neutrophil elastase release. GalNAc has been shown to reduce the amount of neutrophil elastase released from neutrophils *in vitro* [56]. Two other species of interest are *Campylobacter rectus* and *Campylobacter showae*, two oral pathogenic *Campylobacter* species. These oral *Campylobacters* contain a hexose as a carbohydrate on the non-reducing end of their N-linked glycoproteins, in which hexose itself has been shown not to affect neutrophil elastase release [56]. The human oral cavity is a reservoir for several Group II non-thermophilic *Campylobacter* species. The most commonly associated *Campylobacter* species within the oral cavity are: *C. rectus, C. showae, C. concisus, C. gracilis, C. curvus, C. hominis* and *C. ureolyticus* [57,58]. Interestingly, these oral *Campylobacter* species also contain an ecotin homologue just upstream of the *pgl* cluster, which the Group I thermophilic *Campylobacters* do not contain. The importance of ecotin will be discussed further in subsequent sections.

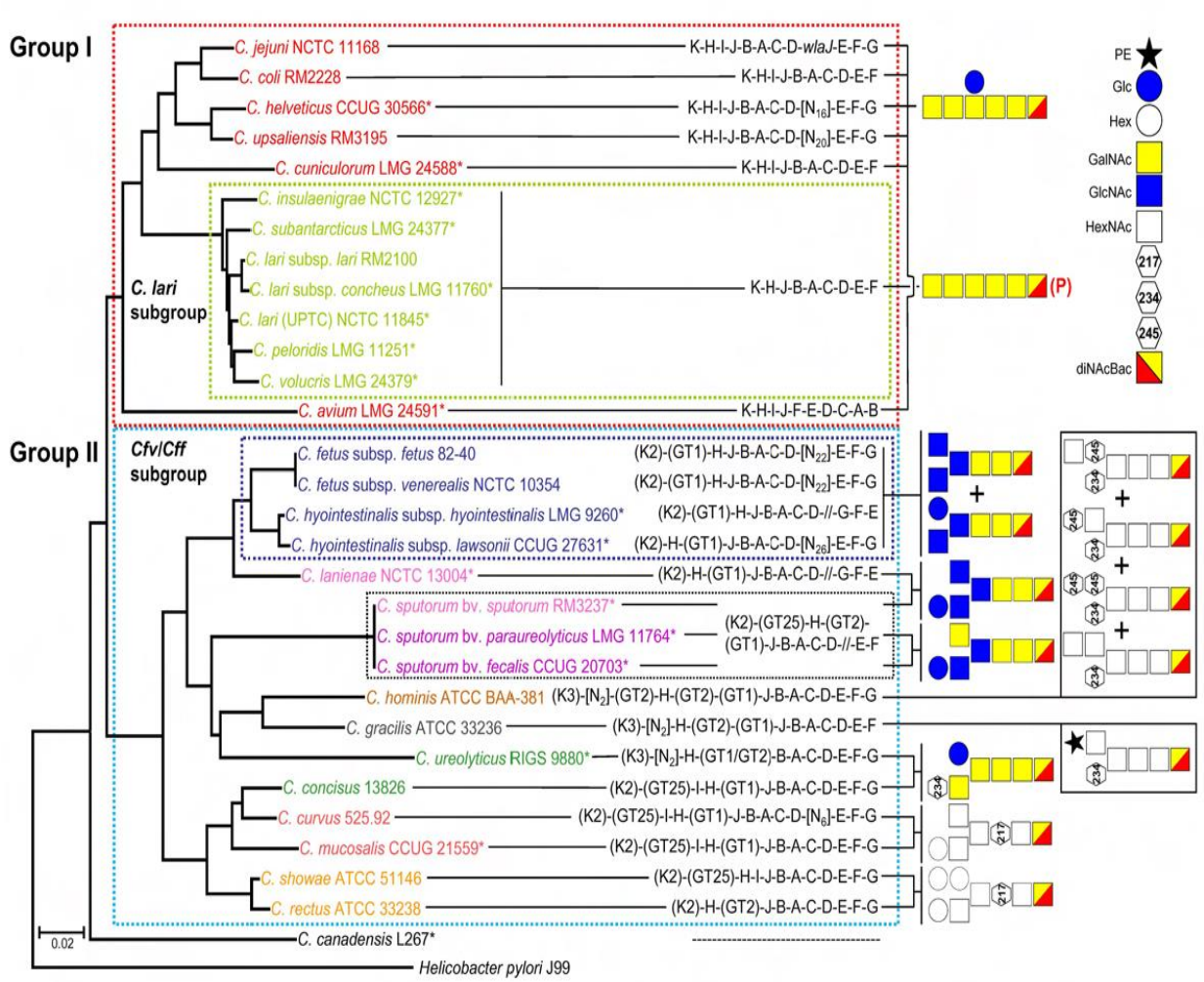


Fig. 1.4. Structure diversity of N-glycans and fOS of *Campylobacter* species. All *Campylobacter* species contain at least one functional *pgl* locus for N-glycosylation. There are some species that contain a second *pglB* homologue of unknown function [59]. *Campylobacter* species can make a diversity of oligosaccharides for N-glycosylation and fOS. A dendrogram of *Campylobacter* species based on their AtpA sequences grouped them into two major subgroups. Group I species are thermophilic species (grow optimally at 42°C), and Group II species are non-thermophilic (grow optimally at 37°C). The reducing end diNAcBac is conserved among all

strains, while the rest of the oligosaccharide is variable between species. Phosphoethanolamine (PE), glucose (Glc), hexose (Hex), N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), di-N-acetylbacillosamine (diNAcBac). Identities of molecules with molecular weights of 217, 234 and 245 Da are unknown. Letters representing each *pgl* gene show the order in each glycosylation cluster by species name. Figure reproduced from Nothaft *et al.* (2012) with permission [60].

1.3.4 Protease protection by glycosylation

1.3.4.1 Glycosylated proteins are protected from proteolytic degradation

O-linked protein glycosylation has been linked with a reduction in protease susceptibility. A study on fibronectin glycosylation showed that glycosylated domains of fibronectin were more resistant to proteolytic degradation than non-glycosylated domains [61]. A previous study on the glycosylation of the transcription factor Sp1 in *Saccharomyces cerevisiae* showed that reducing the amount of O-glycans on the protein lead to faster degradation of the protein by the proteasome [62]. Additionally, a non-glycosylated yeast invertase was shown to be more susceptible to intracellular degradation and trypsin digestion than the native, glycosylated version [63,64]. Another study showed that a glycosylated ribonuclease in *Vigna radiata* or was much more resistant to proteases than a non-glycosylated ribonuclease [65].

1.3.4.2 N-glycan protection of *C. jejuni* from chicken cecum proteases

The roles of *C. jejuni* N-glycosylation are not completely understood, but protein modification does play a role in protease protection. It is believed that N-glycans protect proteins from proteases by limiting access to proteolytic cleavage sites. This was determined following the observation that N-glycosylation of proteins in *C. jejuni* protect proteins and promotes bacterial fitness in chicken cecal contents [5]. In previous studies, a *C. jejuni pglB* mutant unable to N-glycosylate proteins was attenuated in virulence and displayed reduced survival in the chicken gut. To further understand how N-glycans were promoting bacterial fitness, chicken cecum samples were mixed with select *C. jejuni* strains. A protease inhibitor cocktail was able to restore viability of the *C. jejuni pglB* mutant back to *C. jejuni* WT levels, indicating that the N-glycans of *C. jejuni* inhibit proteases found in the chicken host. The engineering of the N-glycosylation system to add glycans to proteins combined with a fluorescence resonance energy transfer (FRET) assay can be used for further study of N-glycan-mediated protease protection. The cleavage of a FRET peptide by proteases results in fluorescence, while the addition of an N-glycan leads to reduced fluorescence. This assay is discussed further in Chapter II.

1.4 Oral *Campylobacter* species

Periodontitis is an inflammatory disease cause by a consortium of microbes and host-defense mechanisms. Bacteria killing mechanisms by neutrophils come as a double-edged sword; reactive oxygen species and bactericidal proteins such as elastase kill bacteria but lso damage host tissues and increase the severity and progression of periodontal diseases [66]. Oral

Campylobacter species have been linked with different stages of periodontitis progression. *C. rectus* is a known pathogen found at elevated levels in diseased human subgingival sites as compared to healthy non-diseased subgingival sites. [67-70]. *C. gracilis* shows closer proximity to 'green complex' oral microbiota while *C. rectus* is typically considered an 'orange complex' organism but can have strong associations with 'red complex' bacteria like *Porphyromonas gingivalis* [71,72]. There are other *Campylobacter* species that have been isolated from oral sites which include *Campylobacter showae, Campylobacter curvus, Campylobacter concisus, Campylobacter sputorum* and *Campylobacter hominis* [67-69,73]. In Chapter II of this thesis, I describe the investigation of two oral *Campylobacter* species, *C. rectus,* the oral *Campylobacter* species known to cause the most oral damage, and *C. showae,* a newly discovered *Campylobacter* species with its precise role in oral disease unknown.

1.5 Serine proteases

Proteases are important and involved in many critical physiological processes such as; digestion, hemostasis, apoptosis, and the immune response[74-77]. Almost one-third of all proteases can be classified as serine proteases, which is named after the nucleophilic serine residue at the active site[78] . Common serine proteases are trypsin which is involved in digestion, factor Xa involved in blood coagulation and elastase involved in pathogen clearance[79-81]. The mechanistic class contains a catalytic triad distinguished by the presence of Asp-His-Ser[78]. The catalytic triad is part of an extensive hydrogen bonding network involved in the cleavage of peptides. Nucleophilic attack happens by the serine hydroxyl group binding to the carbonyl carbon atom of the substrate catalyzed by a histidine imidazole group[82]. This leads to the formation of a tetrahedral intermediate and an imidazolium ion[83]. The intermediate is then broken down by

general acid catalysis to acyl-enzyme and an imidazole base[84]. Then the acylation step occurs where the imidazole group then transfers a proton of the serine hydroxyl group to the amine leaving group. The acyl-enzyme is then deacylated through the reverse reaction pathway of acylation reacting with a water molecule generating another tetrahedral intermediate[85]. The intermediate collapses by acid catalysis from the histidine producing a acid and regenerating the hydroxyl group on serine. The general mechanism of catalysis by serine proteases can be seen in Fig. 1.5.

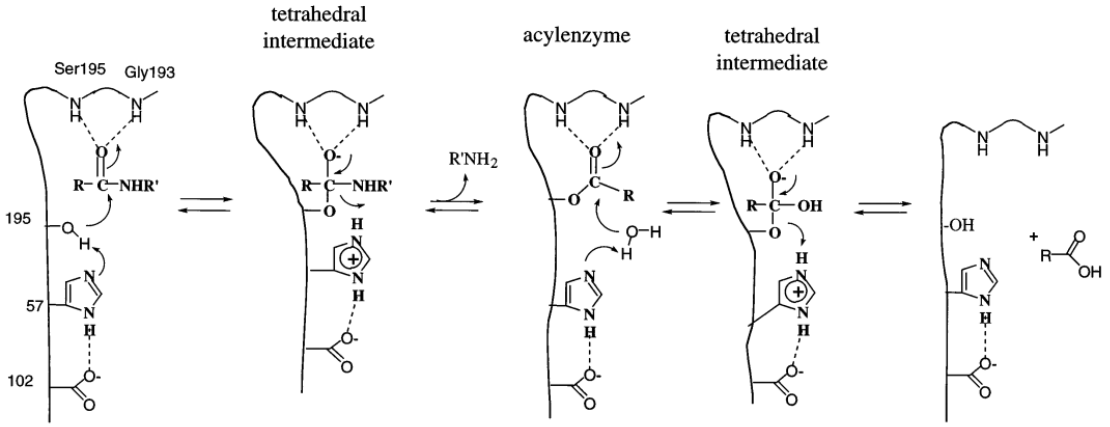


Figure 1.5. The generally accepted mechanism for serine proteases. Figure reproduced from Kraut (1977) with permission[82].

1.6 Ecotin: a periplasmic serine protease inhibitor

First discovered in the *E. coli* periplasmic space, ecotin is a small protein able to inhibit pancreatic serine proteases [86]. The protein was determined to form a homodimer capable of binding two protease molecules. The active site of ecotin was determined [87] and found to bind a wide variety of serine proteases, such as trypsin, chymotrypsin, elastase and factor Xa [81,88,89]. A

majority of protease inhibitors are generally specific to one protease, meaning that the sequence along with the active site correlate with the substrate specificity of the protease. Therefore, the fit between an inhibitor and a protease is very specific. A common strategy for an organism is to contain a family of inhibitors to allow for the inhibition of a variety of proteases [90]. Ecotin seems to be a unique protease inhibitor in the fact that it can bind to variety of proteases [86,89,91].

Ecotin is different from all other serine protease inhibitors in that it has a generic secondary binding site for proteases, resulting from the dimerization of ecotin molecules [87]. The double binding of each ecotin to the protease forms a network that allows it to bind tightly to the proteases (Fig. 1.5)[92]. Ecotins have been described from several species and pathogens, such as *Pseudomonas aeruginosa, Burkholderia pseudomallei* and *Yersinia pestis*. Ecotins have also been associated with the pathogenicity of these species [91,93,94]. There have been 198 ecotin homologues found in nature, 80% are found in bacteria and 78% belong to Proteobacteria [95]. However, the ecotin homologues of species other than *E. coli* have not been fully characterized.

1.6.1 Amino acid composition of ecotin

E. coli ecotin is 142 amino acids in length, and homologues from other species vary by only a few amino acids. It contains a signal peptide that directs it to the periplasm [88]. The primary reactive site of ecotin is the 84th amino acid, known as the P1 residue, which was found to contain a methionine. Methionine in the active site interacts with the substrate-binding pocket of serine proteases such as trypsin, chymotrypsin and elastase, which provides a mechanism for inhibition of these proteases [79,90,96]. The bulky side chain of methionine would be expected to be a problem for the shallow elastase substrate-binding pocket, but by increasing the contact area

with the protease through the secondary binding site, it is able to compensate for the side chain length [97-99]. Ecotin homologues contain either a methionine or leucine at P1 in position 84. Leucine replacement may simply provide a shorter side chain. *C. showae* ecotin contains a methionine at the P1 active site and *C. rectus* ecotin contains a leucine at the P1 active site. The reactive site of the inhibitors reacts with the active site of the enzyme in a substrate manner. The carbonyl group of the leucine or methionine forms van der Waals contact with the oxygen of the serine in the protease. Tetrahedral distortion is then completed by the inhibitor, resulting in the attraction between the carbonyl oxygen of the inhibitor and oxyanion hole of the protease [82]. This distortion stops the reactive site of the protease from being regenerated and results in permanent binding of the inhibitor to the protease [90,100].

The secondary site in ecotin allows it to be different from other serine protease inhibitors, as it contains a generic binding site for proteases. The secondary binding site results from a dimerization of the ecotin proteins (Fig 1.6). The secondary site contains loops similar to those found in antibody binding sites [101], and uses two beta-sheets to bind proteins [102,103]. Therefore, the primary and secondary binding sites allow ecotin to bind its targets specifically (via the primary binding site) and tightly (via the secondary binding site).

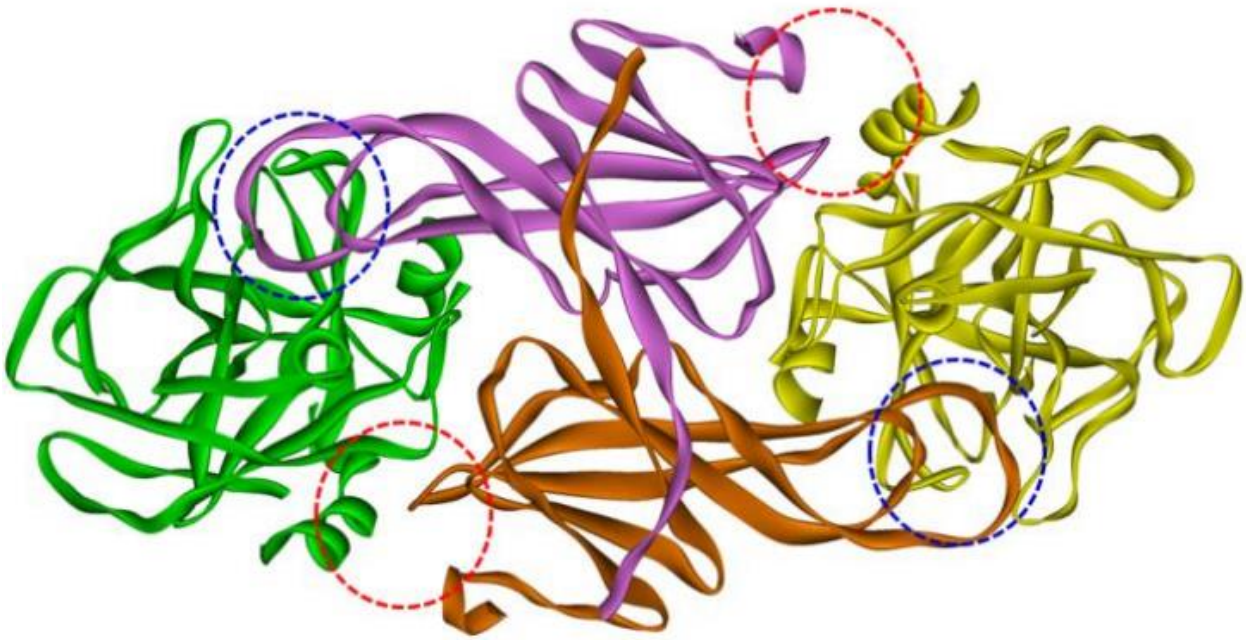


Fig. 1.6. Structure of ecotin homodimer binding to two trypsin molecules. The ecotin homodimer (purple and brown) is shown interacting and binding two trypsin molecules (green and yellow). Ecotins contains primary (red lines) and secondary (blue lines) binding sites, which confer specificity and strength of the interaction, respectively. Figure reproduced from Akermi *et al.* (2017) with permission [92,104].

1.6.2 Ecotin in *Pseudomonas aeruginosa*

Ecotin from *P. aeruginosa* has been shown to be an active inhibitor in the ability to inhibit trypsin, chymotrypsin and NE. A novel feature of ecotin has recently been discovered in *P. aeruginosa*. Ecotin was shown to bind to biofilm matrix exopolysaccharide PsI [105] and inhibit neutrophil elastase. Biofilms help protect communities against antimicrobial agents, and the

ability of *P. aeruginosa* ecotin to inhibit serine proteases was shown to help the biofilm community exhibit a higher tolerance to these proteases [106]. Importantly, this example shows that ecotin is capable of exiting the periplasmic space, either through nonclassical protein secretion methods [107] or cell lysis [108]. This represents a novel mechanism of protection in biofilms for increased tolerance against innate immune response.

1.7 Overview of neutrophil killing

Neutrophils are the first line of host defense in mammals against invading pathogens. Neutrophils are phagocytes that help to kill pathogens by engulfing and degrading cells using oxidative and non-oxidative methods [109]. The traditional method of neutrophil killing involves engulfing the bacteria by phagocytosis and trapping the microorganisms in a phagolysosome (Fig. 1.7). This method activates the NADPH oxidase system, and generates large amounts of reactive oxygen species (ROS), which get released into the phagolysosome and are thought to cause the direct killing of microorganisms [110].

The non-oxidative method of neutrophil-mediated killing involves antimicrobial peptides and proteases, which can be released into the phagolysosome or extracellularly (Fig. 1.7). Among these proteases is a family of structurally related serine proteases: neutrophil elastase, cathepsin G and proteinase 3. The importance of neutrophil elastase was shown to be required for the clearing of Gram-negative bacteria in mice deficient for these proteases [111,112]. This was the first study to show that mice deficient in neutrophil elastase are more susceptible to infection with Gram-negative bacteria than wild-type mice. The proteases of the nucleophile, superfamily

A have a high level of homology, with each containing a serine, histidine and aspartic acid in their catalytic triad [109].

Neutrophil elastase is believed to exert its antibacterial activity by binding to the bacterial outer membrane and causing depolarization and disruption of the membrane. It has been shown that elastase is able to degrade outer membrane protein A (OmpA) of *E. coli*, and when incubated with intact *E. coli* cells, leads to loss of bacterial integrity and reduced viability [113,114]. Using mice deficient in neutrophil elastase, it was shown that there was an increase in *E. coli* as compared to wild-type mice, indicating that neutrophil elastase is essential for clearing and killing Gram-negative bacteria [112,115,116]. Without neutrophil elastase, these bacteria can escape phagolysosomes and survive.

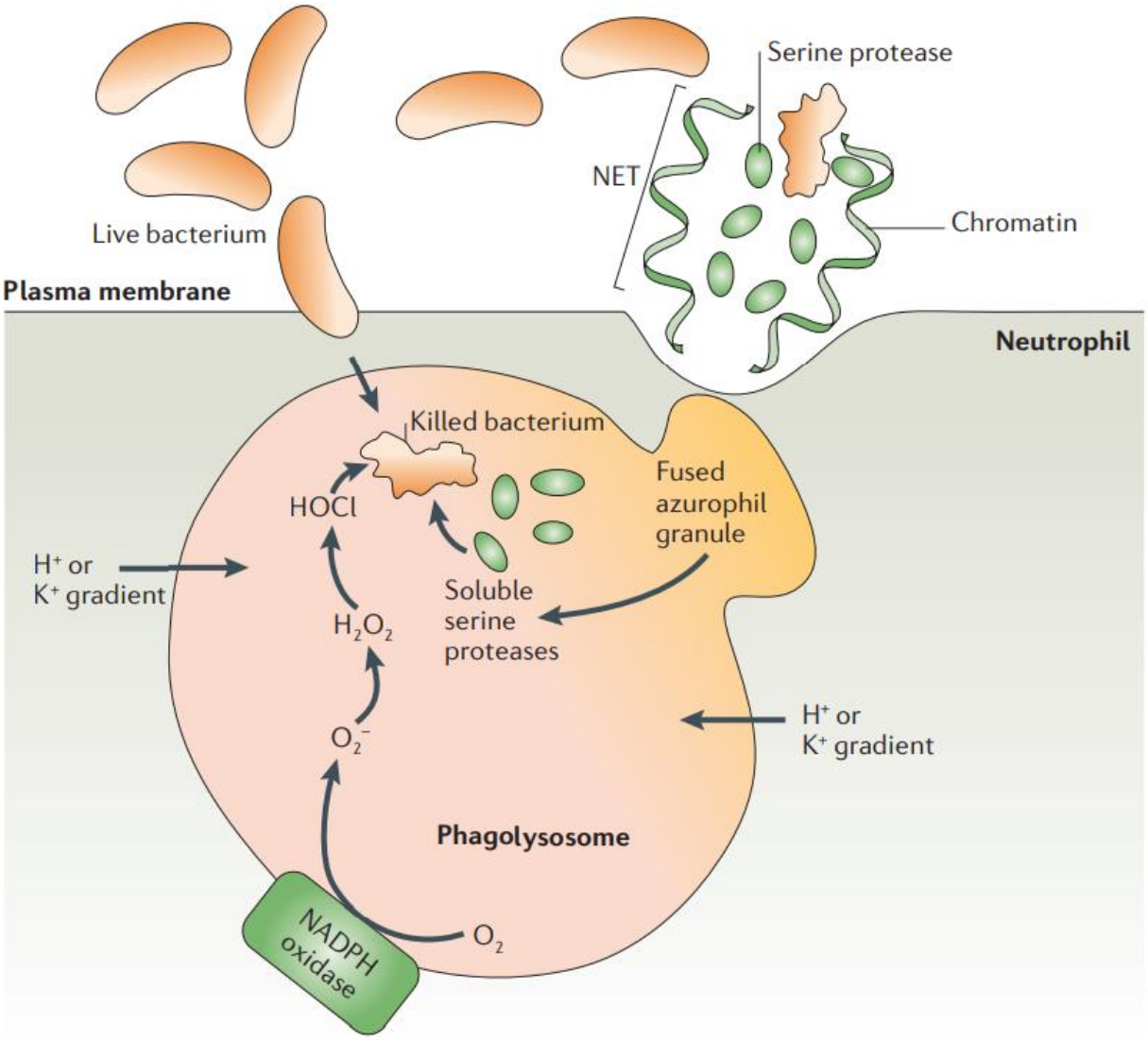


Fig. 1.7. Neutrophil cell killing. Neutrophils can kill bacteria intracellularly or extracellularly. Intracellular neutrophil killing occurs by either the oxidative or non-oxidative method. The oxidative method involves bacterial engulfment into a phagolysosome. In the phagolysosome, reactive oxygen species (ROS) created by the NADPH oxidase system help kill bacteria. Positively charged ions are pumped into the phagolysosome to offset the accumulation of negative ions. In the non-oxidative method, fusion of azurophil granules releases serine

proteases, such as neutrophil elastases, to help kill bacteria. Neutrophils can also degranulate and release neutrophil extracellular traps (NETs) composed of serine proteases and DNA to trap and kill extracellular bacteria. Figure reproduced from Pham *et al.* (2006) with permission [109].

1.8 Thesis objectives

The aims of this thesis work were to better understand the roles of free oligosaccharides (fOS), N-linked glycoproteins and the serine protease inhibitor ecotin in the survival and protection of *Campylobacter* species within mammalian and avian host environments. As described earlier, ecotin is a serine protease inhibitor found in non-thermophilic oral *Campylobacter* species but not found in thermophilic species.

One aim of this thesis was to characterize ecotin from *C. rectus* and *C. showae* and to compare it to the well-understood *E. coli* ecotin. I first cloned and purified these ecotins to test their activity *in vitro*. Once purified, I developed methods to test the ecotin homologues to confirm that they are generic serine protease inhibitors. I showed that they could inhibit the serine protease trypsin using a reporter protein CmeA with SDS-PAGE. I then developed a FRET assay to study the inhibition of factor Xa. I then proceeded to test inhibition of neutrophil elastase *in vitro*, one of the most common proteases encountered by oral bacteria. An ecotin mutant was then created in *E. coli* to study the effects of ecotin *in vivo*. The complementation of ecotin homologues into the *E. coli* ecotin mutant background was done to see if the ecotin homologues play a role in the survival of the cell against whole cell neutrophils and against neutrophil extracellular traps (NETs).

The second aim of my thesis was to learn more about the degradation and recycling of N-glycans and fOS in bacteria, by developing methods to study and follow glycans in bacteria through labelling. I also characterized the glycan accumulated in a *pglB* mutant to help understand how the bacteria may remove glycans from lipid-carriers on the inner membrane.

The final aim of my thesis was to understand how the OTase (PglB) binds to lipid-carriers in *Campylobacter* species. This involved studying the role of the DGGK motif through amino acid substitutions and limited proteolysis in the presence of the LLO substrate to demonstrate its essential OTase function. I also tested whether addition of the DGGK motif that is lacking in certain *Campylobacter* species containing a second inactive homologue of PglB, could restore OTase function. This study helped to understand the mechanism used for LLO binding in N-glycosylation. Furthermore, the study raises the question as to why there exist two copies of the OTase in some *Campylobacter* species.

In summary, this thesis will focus on:

1) Understanding and characterizing the ecotin homologues from oral *Campylobacter* species such as *C. rectus* and *C. showae*.

2) Developing methods to studying N-glycans and fOS recycling in *C. jejuni*.

3) Studying the PglB machinery for exploitation for biotechnology.

1.9 References

1. Cain JA, Solis N, Cordwell SJ. Beyond gene expression: The impact of protein post-translational modifications in bacteria. *Journal of proteomics*. 2014;97:265-286.

2. Spiro RG. Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*. 2002;12(4):56R.

3. Wang Z, Bachvarova M, Morin C, et al. Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: Possible implications in abnormal mucin O-glycosylation. *Oncotarget*. 2014;5(2):544.

4. Hebert DN, Garman SC, Molinari M. The glycan code of the endoplasmic reticulum: Asparagine-linked carbohydrates as protein maturation and quality-control tags. *Trends Cell Biol*. 2005;15(7):364-370.

5. Alemka A, Nothaft H, Zheng J, Szymanski CM. N-glycosylation of *campylobacter jejuni* surface proteins promotes bacterial fitness. *Infect Immun*. 2013;81(5):1674-1682.

6. Varki A. Biological roles of oligosaccharides: All of the theories are correct. *Glycobiology*. 1993;3(2):97-130.

7. Neuberger A. Carbohydrates in protein: The carbohydrate component of crystalline egg albumin. *Biochem J*. 1938;32(9):1435.

8. Johansen PG, Marshall RD, Neuberger A. Carbohydrates in protein. 3. the preparation and some of the properties of a glycopeptide from hen's-egg albumin. *Biochem J*. 1961;78(3):518.

9. Haynes PA. Phosphoglycosylation: A new structural class of glycosylation? *Glycobiology*. 1998;8(1):1-5.

10. Dell A, Galadari A, Sastre F, Hitchen P. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. *International journal of microbiology*. 2011;2010.

11. Steen PVd, Rudd PM, Dwek RA, Opdenakker G. Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol*. 1998;33(3):151-208.

12. Wilson IB, Gavel Y, Von Heijne G. Amino acid distributions around O-linked glycosylation sites. *Biochem J*. 1991;275(2):529-534.

13. Chou C, Omary MB. Mitotic arrest-associated enhancement of O-linked glycosylation and phosphorylation of human keratins 8 and 18. *J Biol Chem*. 1993;268(6):4465-4472.

14. Messner P, Sleytr UB. Bacterial surface layer glycoproteins. *Glycobiology*. 1991;1(6):545-551.

15. Iwashkiw JA, Seper A, Weber BS, et al. Identification of a general O-linked protein glycosylation system in *acinetobacter baumannii* and its role in virulence and biofilm formation. *PLoS pathogens*. 2012;8(6):e1002758.

16. Borud B, Aas FE, Vik o, Winther-Larsen HC, Egge-Jacobsen W, Koomey M. Genetic, structural, and antigenic analyses of glycan diversity in the O-linked protein glycosylation systems of human neisseria species. *J Bacteriol*. 2010;192(11):2816-2829.

17. Szymanski CM, Wren BW. Protein glycosylation in bacterial mucosal pathogens. *Nature Reviews Microbiology*. 2005;3(3):225.

18. Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 1999;1426(2):239-257.

19. Kelleher DJ, Gilmore R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*. 2005;16(4):62R.

20. Mescher MF, Strominger JL. Purification and characterization of a prokaryotic glucoprotein from the cell envelope of halobacterium salinarium. *J Biol Chem*. 1976;251(7):2005-2014.

21. Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P. Evidence for a system of general protein glycosylation in *campylobacter jejuni*. *Mol Microbiol*. 1999;32(5):1022-1030.

22. Yamashita Y, Chung YS, Horie R, Kannagi R, Sowa M. Alterations in gastric mucin with malignant transformation: Novel pathway for mucin synthesis. *JNCI: Journal of the National Cancer Institute*. 1995;87(6):441-446.

23. Hounsell EF, Davies MJ, Renouf DV. O-linked protein glycosylation structure and function. *Glycoconj J*. 1996;13(1):19-26.

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

25. Schaffer C, Graninger M, Messner P. Prokaryotic glycosylation. *Proteomics*. 2001;1(2):248-261.

26. Nothaft H, Szymanski CM. Protein glycosylation in bacteria: Sweeter than ever. *Nature Reviews Microbiology*. 2010;8(11):765.

27. Feldman MF. Industrial exploitation by genetic engineering of bacterial glycosylation systems. In: *Microbial glycobiology*. Elsevier; 2009:903-914.

28. Nachamkin I. Chronic effects of *campylobacter* infection. *Microbes Infect*. 2002;4(4):399-403.

29. Josenhans C, Vossebein L, Friedrich S, Suerbaum S. The *neuA/flmD* gene cluster of *helicobacter pylor*i is involved in flagellar biosynthesis and flagellin glycosylation. *FEMS Microbiol Lett*. 2002;210(2):165-172.

30. Josenhans C, Ferrero RL, Labigne A, Suerbaum S. Cloning and allelic exchange mutagenesis of two flagellin genes of *helicobacter felis*. *Mol Microbiol*. 1999;33(2):350-362.

31. Thibault P, Logan SM, Kelly JF, et al. Identification of the carbohydrate moieties and glycosylation motifs in *campylobacter jejuni* flagellin. *J Biol Chem*. 2001;276(37):34862-34870.

32. Power PM, Roddam LF, Rutter K, Fitzpatrick SZ, Srikhanta YN, Jennings MP. Genetic characterization of pilin glycosylation and phase variation in *neisseria meningitidis*. *Mol Microbiol*. 2003;49(3):833-847.

33. Power PM, Roddam LF, Dieckelmann M, et al. Genetic characterization of pilin glycosylation in *neisseria meningitidis*. *Microbiology*. 2000;146 (Pt 4)(Pt 4):967-979.

34. Nothaft H, Szymanski CM. Protein glycosylation in bacteria: Sweeter than ever. *Nature Reviews Microbiology*. 2010;8(11):765.

35. Schwarz F, Aebi M. Mechanisms and principles of N-linked protein glycosylation. *Curr Opin Struct Biol*. 2011;21(5):576-582.

36. Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proceedings of the National Academy of Sciences*. 2009;106(35):15019-15024.

37. Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P. Evidence for a system of general protein glycosylation in *campylobacter jejuni*. *Mol Microbiol*. 1999;32(5):1022-1030.

38. Linton D, Dorrell N, Hitchen PG, et al. Functional analysis of the *campylobacter jejuni* N-linked protein glycosylation pathway. *Mol Microbiol*. 2005;55(6):1695-1703.

39. Alaimo C, Catrein I, Morf L, et al. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J*. 2006;25(5):967-976.

40. Acheson D, Allos BM. *Campylobacter jejuni* infections: Update on emerging issues and trends. *Clinical infectious diseases*. 2001;32(8):1201-1206.

41. Chamovitz BN, Hartstein AI, Alexander SR, Terry AB, Short P, Katon R. *Campylobacter jejuni*-associated hemolytic-uremic syndrome in a mother and daughter. *Pediatrics*. 1983;71(2):253-256.

42. Wagenaar JA, French NP, Havelaar AH. Preventing *campylobacter* at the source: Why is it so difficult? *Clinical infectious diseases*. 2013;57(11):1600-1606.

43. Keithlin J, Sargeant J, Thomas MK, Fazil A. Systematic review and meta-analysis of the proportion of *campylobacter* cases that develop chronic sequelae. *BMC Public Health*. 2014;14(1):1203.

44. Heimesaat MM, Lugert R, Fischer A, et al. Impact of *campylobacter jejuni* cj0268c knockout mutation on intestinal colonization, translocation, and induction of immunopathology in gnotobiotic IL-10 deficient mice. *PloS one*. 2014;9(2):e90148.

45. Wacker M, Linton D, Hitchen PG, et al. N-linked glycosylation in *campylobacter jejuni* and its functional transfer into *E. coli*. *Science*. 2002;298(5599):1790-1793.

46. Scott NE, Parker BL, Connolly AM, et al. 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*. *Molecular & Cellular Proteomics*. 2011;10(2):MCP201.

47. Larson CL, Christensen JE, Pacheco SA, Minnich SA, Konkel ME. *Campylobacter jejuni* secretes proteins via the flagellar type III secretion system that contribute to host cell invasion and gastroenteritis. In: *Campylobacter, third edition.* American Society of Microbiology; 2008:315-332.

48. Szymanski CM, Burr DH, Guerry P. *Campylobacter* protein glycosylation affects host cell interactions. *Infect Immun*. 2002;70(4):2242.

49. Karlyshev AV, Everest P, Linton D, Cawthraw S, Newell DG, Wren BW. The *campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology*. 2004;150(6):1957-1964.

50. Hendrixson DR. A phase-variable mechanism controlling the *campylobacter jejuni* FlgR response regulator influences commensalism. *Mol Microbiol*. 2006;61(6):1646-1659.

51. Van Sorge NM, Bleumink N, Van Vliet SJ, et al. N-glycosylated proteins and distinct lipooligosaccharide glycoforms of *campylobacter jejuni* target the human c-type lectin receptor MGL. *Cell Microbiol*. 2009;11(12):1768-1781.

52. Scott NE, Nothaft H, Edwards AVG, et al. Modification of the *campylobacter jejuni* N-linked glycan by EptC protein-mediated addition of phosphoethanolamine. *J Biol Chem*. 2012;287(35):29384-29396.

53. Kakuda T, Koide Y, Sakamoto A, Takai S. Characterization of two putative mechanosensitive channel proteins of *campylobacter jejuni* involved in protection against osmotic downshock. *Vet Microbiol*. 2012;160(1-2):53-60.

54. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *campylobacter* genus. *Molecular & Cellular Proteomics*. 2012;11(11):1203-1219.

55. Nothaft H, Szymanski CM. Bacterial protein N-glycosylation: New perspectives and applications. *J Biol Chem*. 2013;288(10):6912-6920.

56. Kamel M, Hanafi M, Bassiouni M. Inhibition of elastase enzyme release from human polymorphonuclear leukocytes by N-acetyl-galactosamine and N-acetyl-glucosamine. *Clin Exp Rheumatol*. 1991;9(1):17-21.

57. Zhang L, Man SM, Day AS, et al. Detection and isolation of *campylobacter* species other than *C. jejuni* from children with crohn's disease. *J Clin Microbiol*. 2009;47(2):453-455.

58. Zhang L. Oral *campylobacter* species: Initiators of a subgroup of inflammatory bowel disease? *World Journal of Gastroenterology: WJG*. 2015;21(31):9239.

59. Barre Y, Nothaft H, Thomas C, et al. A conserved DGGK motif is essential for the function of the PglB oligosaccharyltransferase from *campylobacter jejuni*. *Glycobiology*. 2017;27(10):978-989.

60. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *campylobacter* genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219.

61. Bernard BA, Yamada KM, Olden K. Carbohydrates selectively protect a specific domain of fibronectin against proteases. *J Biol Chem*. 1982;257(14):8549-8554.

62. Han I, Kudlow JE. Reduced O glycosylation of Sp1 is associated with increased proteasome susceptibility. *Mol Cell Biol*. 1997;17(5):2550-2558.

63. Chu FK, Maley F. The effect of glucose on the synthesis and glycosylation of the polypeptide moiety of yeast external invertase. *J Biol Chem*. 1980;255(13):6392-6397.

64. Brown JA, Segal HL, Maley F, Trimble RB, Chu F. Effect of deglycosylation of yeast invertase on its uptake and digestion in rat yolk sacs. *J Biol Chem*. 1979;254(10):3689-3691.

65. Birkeland AJ, Christensen TB. Resistance of glycoprotiens to proteolysis, ribonuclease: Ribonuclease-A and ribonuclease-B compared. *JOURNAL OF CARBOHYDRATES-NUCLEOSIDES-NUCLEOTIDES*. 1975;2(1):83-90.

66. Sorsa T, TjÃ¤derhane L, Konttinen YT, et al. Matrix metalloproteinases: Contribution to pathogenesis, diagnosis and treatment of periodontal inflammation. *Ann Med*. 2006;38(5):306-321.

67. Tanner ACR, Haffer C, Bratthall GT, Visconti RA, Socransky SS. A study of the bacteria associated with advancing periodontitis in man. *J Clin Periodontol*. 1979;6(5):278-307.

68. Tanner A, Socransky SS, Goodson JM. Microbiota of periodontal pockets losing crestal alveolar bone. *J Periodont Res*. 1984;19(3):279-291.

69. Moore W, Moore LV. The bacteria of periodontal diseases. *Periodontol 2000*. 1994;5(1):66-77.

70. Dzink JL, Socransky SS, Haffajee AD. The predominant cultivable microbiota of active and inactive lesions of destructive periodontal diseases. *J Clin Periodontol*. 1988;15(5):316-323.

71. Rawlinson A, Eley A, Bennett KW, Goodwin L. Bacteroides gracilis in periodontal health and disease. *Microb Ecol Health Dis*. 1994;7(4):201-205.

72. Henne K, Fuchs F, Kruth S, Horz H, Conrads G. Shifts in *campylobacter* species abundance may reflect general microbial community shifts in periodontitis progression. *Journal of oral microbiology*. 2014;6(1):25874.

73. Etoh Y, Dewhirst FE, Paster BJ, Yamamoto A, Goto N. *Campylobacter showae* sp. nov., isolated from the human oral cavity. *Int J Syst Bacteriol*. 1993;43(4):631-639.

74. Johnson DE. Noncaspase proteases in apoptosis. *Leukemia*. 2000;14(9):1695.

75. Joseph K, Ghebrehiwet B, Kaplan AP. Activation of the kinin-forming cascade on the surface of endothelial cells. *Biol Chem*. 2001;382(1):71-75.

76. Coughlin SR. Thrombin signalling and protease-activated receptors. *Nature*. 2000;407(6801):258.

77. Barros C, Crosby JA, Moreno RD. Early steps of sperm–egg interactions during mammalian fertilization. *Cell Biol Int*. 1996;20(1):33-39.

78. Blow DM. The tortuous story of asp… his… ser: Structural analysis of α-chymotrypsin. *Trends Biochem Sci*. 1997;22(10):405-408.

79. Hochstrasser K, Fritz H. The amino acid sequence of the double-headed protein proteinase inhibitor from dog submandibular glands, I. structural homology to the pancreatic secretory trypsin inhibitors *Hoppe-Seyler's Z Physiol Chem*. 1975;356(10):1659-1662.

80. azzaq Belaaouaj A, Kim KS, Shapiro SD. Degradation of outer membrane protein A in *escherichia coli* killing by neutrophil elastase. *Science*. 2000;289(5482):1185-1187.

81. Lauwereys M, Stanssens P, Lambeir AM, Messens J, Dempsey E, Vlasuk GP. Ecotin as a potent factor xa- inhibitor. . 1993;69(6):783.

82. Kraut J. Serine proteases: Structure and mechanism of catalysis. *Annu Rev Biochem*. 1977;46(1):331-358.

83. Johnson CH, Knowles JR. The binding of inhibitors to α-chymotrypsin. *Biochem J*. 1966;101(1):56.

84. Henderson R. Structure of crystalline α-chymotrypsin: IV. the structure of indoleacryloyl-α-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme. *J Mol Biol*. 1970;54(2):341-354.

85. Hartley BS, Kilby BA. The reaction of p-nitrophenyl esters with chymotrypsin and insulin. *Biochem J*. 1954;56(2):288.

86. Chung CH, Ives HE, Almeda S, Goldberg AL. Purification from *escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. *J Biol Chem*. 1983;258(18):11032-11038.

87. Mcgrath ME, Gillmor SA, Fletterick RJ. Ecotin: Lessons on survival in a protease-filled world. *Protein Science*. 1995;4(2):141-148.

88. McGrath ME, Erpel T, Browner MF, Fletterick RJ. Expression of the protease inhibitor ecotin and its co-crystallization with trypsin. *J Mol Biol*. 1991;222(2):139-142.

89. Seymour JL, Lindquist RN, Dennis MS, et al. Ecotin is a potent anticoagulant and reversible tight-binding inhibitor of factor xa. *Biochemistry (N Y)*. 1994;33(13):3949-3958.

90. Laskowski Jr M, Kato I. Protein inhibitors of proteinases. *Annu Rev Biochem*. 1980;49(1):593-626.

91. Ireland PM, Marshall L, Norville I, Sarkar-Tyson M. The serine protease inhibitor ecotin is required for full virulence of *burkholderia pseudomallei*. *Microb Pathog*. 2014;67:55-58.

92. Akermi N, Mkaouar H, Jeblaoui A, et al. Relevant patented biotechnological applications of ecotin: An update. *Recent patents on biotechnology*. 2017.

93. Eggers CT, Murray IA, Delmar VA, Day AG, Craik CS. The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. *Biochem J*. 2004;379(Pt 1):107.

94. Clark EA, Walker N, Ford DC, Cooper IA, Oyston PC, Acharya KR. Molecular recognition of chymotrypsin by the serine protease inhibitor ecotin from *yersinia pestis*. *J Biol Chem*. 2011;286(27):24015-24022.

95. Lee HR, Seol JH, Kim OM, et al. Molecular cloning of the ecotin gene in *escherichia coli*. *FEBS Lett*. 1991;287(1-2):53-56.

96. Papamokos E, Weber E, Bode W, et al. Crystallographic refinement of japanese quail ovomucoid, a kazal-type inhibitor, and model building studies of complexes with serine proteases. *J Mol Biol*. 1982;158(3):515-537.

97. Rydel TJ, Ravichandran KG, Tulinsky A, et al. The structure of a complex of recombinant hirudin and human alpha-thrombin. *Science*. 1990;249(4966):277-280.

98. Bode W, Engh R, Musil D, et al. The 2.0 AXray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J*. 1988;7(8):2593-2599.

99. Stubbs MT, Laber B, Bode W, et al. The refined 2.4 xray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: A novel type of proteinase inhibitor interaction. *EMBO J*. 1990;9(6):1939-1947.

100. Neves DE, Markley JL, Welch ME, Laskowski M. NMR-studies of the reactive site arginine of soybean trypsin-inhibitor (STI). . 1979;38(3):474.

101. Janin J, Chothia C. The structure of protein-protein recognition sites. *J Biol Chem*. 1990;265(27).

102. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol*. 1987;196(4):901-917.

103. Branden C, Tooze J. Prediction, engineering, and design of protein structures. *Chapter*. 1991;16:247.

104. Gillmor SA, Takeuchi T, Yang SQ, Craik CS, Fletterick RJ. Compromise and accommodation in ecotin, a dimeric macromolecular inhibitor of serine proteases1. *J Mol Biol*. 2000;299(4):993-1003.

105. Tseng BS, Reichhardt C, Merrihew GE, et al. A biofilm matrix-associated protease inhibitor protects *pseudomonas aeruginosa* from proteolytic attack. *mBio*. 2018;9(2):543.

106. Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: A common cause of persistent infections. *Science*. 1999;284(5418):1318-1322.

107. Wang G, Chen H, Xia Y, et al. How are the non-classically secreted bacterial proteins released into the extracellular milieu? *Curr Microbiol*. 2013;67(6):688-695.

108. Webb JS, Thompson LS, James S, et al. Cell death in *pseudomonas aeruginosa* biofilm development. *J Bacteriol*. 2003;185(15):4585-4592.

109. Pham CT. Neutrophil serine proteases: Specific regulators of inflammation. *Nature Reviews Immunology*. 2006;6(7):541.

110. Dinauer MC, Lekstrom-Himes JA, Dale DC. Inherited neutrophil disorders molecular basis and new therapies. *ASH Education Program Book*. 2000;2000(1):303-318.

111. Belaaouaj A, McCarthy R, Baumann M, et al. Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis. *Nat Med*. 1998;4(5):615.

112. Reeves EP, Lu H, Jacobs HL, et al. Killing activity of neutrophils is mediated through activation of proteases by K flux. *Nature*. 2002;416(6878):291.

113. Zasloff M. Antimicrobial peptides of multicellular organisms. *Nature*. 2002;415(6870):389.

114. azzaq Belaaouaj A, Kim KS, Shapiro SD. Degradation of outer membrane protein A in *escherichia coli* killing by neutrophil elastase. *Science*. 2000;289(5482):1185-1187.

115. Weinrauch Y, Drujan D, Shapiro SD, Weiss J, Zychlinsky A. Neutrophil elastase targets virulence factors of enterobacteria. *Nature*. 2002;417(6884):91.

116. Belaaouaj A, McCarthy R, Baumann M, et al. Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis. *Nat Med*. 1998;4(5):615.

Chapter II

Characterization of the ecotin homologs in *Campylobacter rectus* and *Campylobacter showae* [1]

[1] Thomas C, Nothaft H, Yadav R, Alemka A, Rada B, Szymanski CM. To be submitted to *PLoS Pathogens*.

Abstract

Ecotin is a potent periplasmic serine protease inhibitor that was first described in *Escherichia coli*. Ecotin can inhibit a broad range of serine proteases including those typically released by the innate immune system such as neutrophil elastase. We have identified ecotin orthologues in various *Campylobacter* species, including *Campylobacter rectus* and *Campylobacter showae* that reside in the oral cavity and are implicated in the development and progression of periodontal disease in humans. To investigate the function of the Campylobacter ecotins *in vitro,* the orthologues from *C. rectus* and *C. showae* were recombinantly expressed and purified from *E. coli*. Using CmeA degradation/protection assays, fluorescence resonance energy transfer and neutrophil elastase activity assays we found that ecotins from *C. rectus* and *C. showae* inhibit elastase, factor Xa and trypsin, but not the *Campylobacter jejuni* serine protease HtrA or its orthologue in *E. coli*, DegP. To further evaluate their *in vivo* function, an *E. coli* ecotin mutant was complemented with both *C. rectus* and *C. showae* ecotins. Using a neutrophil killing assay we demonstrate that the low survival rate of the *E. coli* ecotin mutant can be rescued upon expression of ecotins from *C. rectus* and *C. showae* thus implicating a similar role of these proteins in the native host in order to ensure survival under the harsh protease-rich environment of the oral cavity.

2.1 Introduction

Bacteria have developed various strategies to survive in their ecological niche. In the human host, proteases are vital players of the immune system for protecting against and clearing pathogens[1]. One of the particular areas is the oral cavity with the protection the innate immune system's neutrophil cells. Polymorphonuclear neutrophils (PMN) are the first line of defense when it comes to the immune system and they are the most prominent circulating leukocytes[2]. Neutrophils play a pivotal role in the defense against invading microorganisms. They contain a variety of neutrophil serine proteases (NSPs) the predominant one being neutrophil elastase (NE). Neutrophils can engulf invading bacteria via endocytosis and form a phagolysosome that kills the microbes by the action of these proteases and reactive oxygen species. On the other hand, neutrophils can release these proteases at inflammatory sites during activation. Aside from assisting in pathogen destructions, NSPs are involved in inflammatory human health conditions like chronic lung diseases[3-6]. Here, neutrophils are overstimulated, resulting in excessive accumulation of NSPs and in the destruction of host cells.

Bacteria have developed a variety of mechanism to evade neutrophil-mediated killing, including launching a general survival response, avoiding contact, preventing phagocytosis, surviving inside the neutrophil, inducing cell death and avoiding killing in NETs.

Bacterial countermeasures against the action of proteases is the production of protease inhibitors like the serine protease inhibitor ecotin, another means of protecting against proteases is the glycosylation of proteins to mask proteolytic cleavage sites. Glycosylation is the most common post-translational modification of proteins in nature. Glycans can be covalently attached to the amide nitrogen of asparagine residues in N-linked protein glycosylation. Although known to exist in Eukaryotes and Archaea, the foodborne pathogen *Campylobacter jejuni* was the first bacteria

shown to possess a general N-linked glycosylation (*pgl* locus) system[7]. It has recently been shown that N-glycosylation indeed protects *C. jejuni* surface proteins from the action of gut proteases. When incubated with chicken cecal contents that contained a variety of gut proteases, a mutant with a defect in the N-glycosylation pathway had a significantly lower survival rate when compared to the wild-type strain, a phenotype that could be rescued by the addition of a protease inhibitor cocktail[8]. Whilst the *pgl* locus in *C. jejuni* exclusively harbors genes required for the biosynthesis and of the N-linked carbohydrate, further examination of *pgl* loci from all *Campylobacter* species[9], revealed the presence of an open reading frame for a generic serine protease inhibitor, an ecotin homolog of *E. coli,* in certain non-thermophilic *Campylobacter* species (Fig.1) that are described to inhabit the oral cavity and have been associated with the onset of periodontal disease[10].

In these species, the N-glycan heptasaccharide is believed to function just like the *C. jejuni* heptasaccharide that mask protease cleavage sites and protects glycoproteins from proteolysis. However, in the oral cavity of the human host the N-glycosylation system may not be sufficient to protect from proteolysis and ecotin homologs might provide an additional survival advantage in the periodontal pockets that contain high levels of neutrophil serine proteases as neutrophils degranulate and release large amounts of these proteases to help clear away the bacteria[11]. The serine proteases of neutrophils have recently been shown to be particularly important for clearing infections[12]. The ability to inhibit these serine proteases may allow for the survival of *Campylobacter* species like *C. rectus* and *C. showae*.

Ecotin was first described in *E. coli* as a periplasmic protease inhibitor that exhibits a broad specificity toward exogenous serine proteases like trypsin, chymotrypsin, factor Xa and elastase but not against metallo-, aspartyl and sulfhydryl or its own proteases[13],[14],[15]. General serine protease

inhibitors appear to have risen through convergent evolution resulting in being comprised of very different amino acid sequences but are similar in regards to their structure and function[16] and are deemed beneficial for bacteria that can reside in human host environments.

A novel feature of ecotin was recently discovered in *Pseudomonas aeruginosa*[17]. Here, ecotin was shown to escape the cell and bind to biofilm matrix exopolysaccharides (PsI)[17] that are known to protect bacterial communities against antimicrobial proteins[18]. Although it is unknown whether *P. aeruginosa* ecotin is secreted through nonclassical secretion methods[19] or cell lysis[20], it has been shown that ecotin enhances the survival and persistence of *P. aeruginosa* in biofilms found in cystic fibrosis patients.

In this study we investigated if ecotin homologues from *Campylobacter* species share the broad specificity towards serine-proteases. We tested the *C. rectus* and *C. showae* ecotin homologues against a panel of proteases and investigated their capability of inhibiting the periplasmic bacterial serine proteases, DegP and HtrA from *E. coli* and *C. jejuni*, respectively. We then employed an *in vivo* cell killing assay and demonstrated that a generated ecotin-deficient *E. coli* strain complemented with ecotin homologues from *C. showae* and *C. rectus* showed comparable survival rates when incubated with whole cell neutrophils or purified NETs, indicating that the *Campylobacter* ecotin homologues fulfill a similar function in the native host.

2.2 Materials and methods

2.2.1 Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table.1. *Escherichia. coli* was grown on LB agar or in 2xYT broth at 37^{o}C with shaking at 220 rpm. *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^{o}C. *Campylobacter rectus, Campylobacter showae, Campylobacter gracilis, Campylobacter hominis* and *Campylobacter curvus* were grown on Blood Heart Infusion (BHI) supplemented with 5% horse blood under anaerobic conditions. The antibiotics ampicillin (100 µg/mL), chloramphenicol (25 µg/mL) and trimethoprim (25 µg/mL) were added to the growth medium when needed for selection.

2.2.2 Construction of plasmids

Ecotin genes from *E. coli*, *C. rectus, C. showae, C. hominis, C. gracilis* and *C. curvus* were PCR-amplified from chromosomal DNA with the respective oligonucleotides (Table 2). Obtained PCR products were purified and digested with the respective restriction nucleases and ligated into plasmid pET22b digested with the same enzymes (Table 1). After transformation of *E. coli* DH5α, positive clones were identified by plasmid-restriction analysis and further confirmed by DNA sequencing. On these constructs Ecotins are expressed as C-terminal 6xHis-tagged fusion proteins. In the case of *Campylobacter* ecotins the native signal sequence as determined by SignalP (http://www.cbs.dtu.dk/services/SignalP/) was replaced with the *pelB* leader peptide present on plasmid pET22b. Ecotin from *E. coli* was expressed with its native signal sequence.

2.2.3 Construction of *E. coli* ecotin mutant (*ecotin*::*kan*)

The *E. coli* BL21 ecotin deletion mutant (EC-*ecotin*::kan) was constructed following the protocol of Wanner and Datsenko[21]. Briefly, the kanamycin (*kan*) cassette from plasmid pKD4 was PCR amplified with oligonucleotides pKD4-ecotin-F and pKD4-ecotin-R. The purified PCR product was then electroporated into *E. coli* BL21 carrying plasmid pKD46 grown in 2xYT + 1% arabinose. After out-growth for 1 h at 30°C cells were spread on LB-kan agar and grown at 37°C. Candidate colonies (*ecotin*::kan = KanR, AmpS) that have the ecotin gene replaced with the kan cassette by simultaneously having lost the temperature-sensitive plasmid pKD46 were confirmed by PCR analysis of their chromosomal DNA with oligonucleotides EC-ecotin-F and EC-ecotin-R that hybridize outside of the recombination event. One candidate from which the correct PCR product with a size of 2119 bp was obtained was used for further analyses.

2.2.4 Expression and purification of ecotins proteins in *E. coli*

E. coli BL21 containing pET22b-ecotin expression plasmids were grown in LB broth at 37°C until an OD600 of 0.6 was reached. Ecotin expression was induced by the addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.3 mM. Cells were further grown overnight (18 hrs) and harvested by centrifugation 8,000 x g at 4^{o}C. Cells were re-suspended in PBS and passed through a homogenizer (EmulsiFlex-C5, Avestin at 10000 PSI for 5 minutes). Obtained cell lysates were centrifuged at 16000 g for 30 minutes at 4^{o}C. The resulting supernatant was run through Ni-NTA column washed 3 times with 15 ml of 15 mM imidazole and bound ecotin proteins were eluted with 6 ml 300 mM imidazole-PBS. Aliquots of elution fractions of approx. 1 ml each were analyzed by a 12% SDS-PAGE. Fractions that contain Ecotins (Data

not shown) were dialyzed with 10kD MWCO tubing against 4 L PBS at 4ºC overnight with PBS changed after 12 hours. Purified Ecotins are stored at 4ºC until further use.

2.2.5 Complementing of the *E. coli ecotin(eco)*::*kan* mutant

The *E. coli* BL21 *eco*::*kan* mutant was transformed with the pET22b-ecotin expression plasmids by electroporation. After selection on LB amp plates select colonies were inoculated and grown in ten-milliliters of LB broth at 37ºC in to an OD_{600} of 0.6 and induced with IPTG. IPTG concentrations used to induce ecotin expression were: *EC-ecotin* = 0.1 mM, *CR-ecotin* = 0.4 mM and *CS-ecotin* = 0.4 mM. Cultures were grown for an additional 4 hrs under the same conditions and cells were harvested by centrifugation at 8000 xg for 30 minutes at 4ºC. Cells were re-suspended in ice-cold 1 x PBS and directly used in the bacterial neutrophil killing assay with intact neutrophils (as described below).

2.2.6 Trypsin serine protease protection assay

CmeA from *C. jejuni* was used as protein substrates. The *C. jejuni* CmeA protein was purified from the expression vector pMW2 as previously described [22]. CmeA (250 µg) was incubated with ecotin (250 µg) or PBS (control) and mixed with (250 µg) of trypsin (GIBCO®). Samples were incubated at 37ºC or 42ºC with 15 µl aliquots taken every hour. Aliquots were immediately mixed with protein loading dye and analyzed by 15% SDS-PAGE followed by Coomassie staining.

2.2.7 Self-protease HtrA and DegP assay

Cloning of *htrA* from *C. jejuni*. The *htrA* gene (including its native periplasmic secretion signal peptide) was amplified from chromosomal DNA of *C. jejuni* with oligonucleotides htrA-

NdeI-F and htrA-XhoI-R. Obtained, purified and NdeI-XhoI digested PCR product was ligated into plasmid pET22b digested with the same enzymes. After transformation of *E. coli* DH5α plasmids isolated from selected colonies were analysed by DNA restriction and confirmed by DNA sequencing. One thise construct Cj-HtrA is C-terminally fused to a 6xHis-Tag sequence. One positive plasmid candidate was used to transform *E. coli* BL21. Purification of HtrA and DegP: HtrA-His was expressed and purified from *E. coli* BL21/pET22b-htrA-his6 grown in LB broth at 37^{o}C to an OD_{600} of 0.6. HtrA-His expression was induced with 500 mM IPTG for 5 hours. Cells were harvested, lyzed and protein HtrA-His was purified as similar to Ecotins described above. DegP-His overexpression and purification was performed as previously described[23]. The soluble, His-tagged version of CmeA from *C. jejuni*, encoded on plasmid pWA2 was produced and purified as described[22] and was used as the proteolytic substrate. The assay contained His-tagged CmeA, DegP or HtrA and ecotin from either *E. coli, C. rectus* or *C. showae* (all at 0.250 µg/µl) or an equal volume of PBS as a negative control in a total volume of 150 µl. Samples were initially mixed on ice and then incubated at 45^{o}C. Aliquots of 15 µl were taken after 0, 1, 3, 6 hours of incubation, immediately mixed with protein loading dye, incubated for 5 min at 95°C and stored at -20°C until samples were analyzed by 12.5% SDS-PAGE followed by Coomassie staining.

2.2.8 FRET Assay

The FRET peptide: Terminally labelled fluorophore/quencher (FRET) peptide (Dabcyl-DQNATIDGRKQ-Edans, Edans-fluorophore and Dabyl-quencher) carrying protease cleavage sits (i.e. IDGR) were custom ordered from GenScript, Inc. Peptides were resuspended in 10% isopropanol in deionized water to a final concentration of 573 µM and stored at -20^{o}C until use.

Peptides and reactions containing peptides were protected from light and wrapped in aluminum foil at all times unless stated otherwise.

The FRET peptide reaction (Fig. 4): In an opaque 96 well plate, concentration of ecotin ranging from 0 µg- 1 µg were mixed with 1 µl of FRET peptide and 20 µl of 10 X factor Xa buffer in a total volume of 200 µl. The samples were loaded into a plate reader with an excitation wavelength of 355 nm and emission of 530 nm. The first run was blanked and then 100 ng of factor Xa was added to the corresponding wells. Fluorescence was measured every 5 minutes over a timeframe of 1 hour. A schematic of the FRET assay is depicted in Figure. 2.

2.2.9 Neutrophil elastase titration curve assay

A standard curve for the fluorescence produced by the BioVision Neutrophil Elastase Activity Kit (#K383-100) with varying amounts of purified NE from 0 ng – 100 ng per well was created. NE assay buffer and purified NE was added to a final volume of 50 µl per well. Samples were measure for fluorescence at excitation 380 nm and emission 500 nm. Plate was measured every 5 minutes for an hour. Data was plotted to create a standard curve of the amount of fluorescence produced by the end-point per amount of NE available in solution. For testing of ecotin variants to inhibit the purified NE, an amount of 50 ng elastase was used in the each well with ecotins available. Amounts of variant ecotins added to test (0 ng, 10 ng, 20 ng, 30 ng, 40 ng, 50 ng, 75 ng and 100 ng) against 50 ng elastase. Keeping total volume 50 µl per well. All substrates were added and measured every 5 minutes for one hour. Data was graphed and end-pointed compared back to NE titration curve to show how much elastase remained in solution compared to how much ecotin was present.

2.2.10 Neutrophil elastase inhibition assay

The 96-well plate based Fluorometric Neutrophil Elastase Activity Assay Kit (BioVision #K383-100) was used to determine ecotin protection ability against NE. The kit contains a NE substrate-fluorophore compound that is cleaved in the presence of NE resulting in fluorescence. This kit provides a screening tool to determine the efficiency of ecotin to inhibit NE and determining the activity of NE in solution.

In an opaque 96 well plate, concentration of ecotin ranging from 0 ng to 750 ng were mixed with 2 µl of NE from the BioVision Neutrophil Elastase Activity Kit (#K383-100). Ecotin and NE assay buffer were added to a final volume of 50 µl in each well. Addition 2 µl of NE enzyme (5 ng/µl) was added to each well. Samples were measure for fluorescence at excitation 380 nm and emission 500 nm. Samples were blanked before addition of NE. Then 10 µl of purified NE (5 ng/µl) was added to each well and measure every 5 minutes for 40 minutes for fluorescence.

2.2.11 Isolation of Neutrophils

Whole blood was drawn from healthy adult volunteers at the Health Center of University of Georgia under informed consent according to procedures approved by the Institutional Review Boards at the University of Georgia. Polymorphonuclear leukocytes (PMN) were purified as previously described [24]. Briefly, PMNs were isolated from human blood by dextran sedimentation and Percoll gradient centrifugation.

2.2.12 Neutrophil killing assay

Bacterial killing by human neutrophils was measured as described with some modifications [24]. Isolated neutrophils were washed 2-times with 1 mL 1X HBS + 10% Hepes + 10% serum + 5%

glucose. A 9:1 ratio of bacteria to neutrophil ratio (45 million bacteria: 5 million PMN cells) was used for the assay. Prepared bacteria were washed 2-times with 1 x PBS. 90 µl of 1 x 10^8 cells/ml were mixed with 10 µl serum, and incubated at room temperature for 5 minutes. Subsequently, 50 µl of the bacterial solution were mixed with 450 µl PMN cells and tubes were incubated with shaking (200 rpm), at 37ºC. Aliquots were taken at 0, 10, 20 minutes. For each time point, 30 µl of cells/ neutrophil solution were mix in 96 well plate, with 270 µl 1X HBSS + saponin solution and incubated for 5 min on ice to open neutrophil cells and release live bacteria. 30 µl of cell/saponin solution were then mixed with 270 µl 1X HBSS solution to wash away saponin. At this point samples can be kept on ice until all aliquots are processed. 40 µl of solution were then transferred to 160 µl of LB in the corresponding wells of a new 96 well plate for analysis. To do so 96 well plates were incubated at 37ºC for 10 minutes and read in an EON plate reader. Analysis was done in kinetic mode for 8 hours measuring every 2 minutes.

2.2.13 Neutrophil supernatant killing assay

Isolated NETs were examined using the BioVision Neutrophil Elastase Activity Kit (#K383-100) to determine activity of NE in solution. Prepared bacteria were washed 2-times with 1x PBS solution. Subsequently, 50 µl of 1 x 10^7 cells/ml were mixed with 150 µl of NETs or 150 µl of 1x PBS solution. Samples were incubated at 37ºC for 30 minutes. A 10-fold dilution series was created with 1x PBS. Dilution series had 10 µl plated onto respected agar plates and incubated overnight at 37 ºC**.**

2.2.14. Preparation of NETs.

Neutrophil extracellular traps (NETs) were prepared by stimulating purified human neutrophils seeded on 96-well microplate with 100 nM phorbol-myristate-acetate (PMA) for 4-6 hours. Following stimulation, supernatants of neutrophils were carefully removed and replaced by equivalent volume of sterile Hank's Balanced Salt Solution. NETs attached to the bottom of the wells were subjected to limited DNAse digestion (1 U/ml DNAse I (Sigma Aldrich), 15 minutes) as previously described *(please cite this paper: https://www.ncbi.nlm.nih.gov/pubmed/24670966).* DNAse activity was then stopped by adding 1 mM EGTA. NETs were collected and centrifuged at 1,000 Xg to remove cells and cell debris. NETs were stored at -80 °C until use for experiments.

2.3 Results

2.3.1 Identification and *in silico* analyses of Campylobacter ecotin orthologues

Campylobacter ecotin homologues were identified by NCBI-BLAST searching using either the amino acid sequence of the *E. coli* or *C. rectus* ecotins as a query. The % amino acid identity/similarity compared to the *E. coli* ecotin is depicted in Fig. 1a. Based on the low % of similarity, sequences were also analyzed by the structure prediction program, PHYRE2 (www.sbg.bio.ic.ac.uk/phyre2). Exemplified by the analyses of *C. rectus and C. showae* ecotins that show only 27% and 33% identity on the amino acid level to the *E. coli* ecotin, the structural analysis resulted in 90% confidence comparing protein fold prediction. Since the *E. coli* ecotin is known to be a periplasmic protein, secretion signals in the Campylobacter ecotins were identified

using SignalP (Fig. 1). Based on the observed low probabilities for the native signal sequences, they were replaced with the *pelB* signal on plasmid pET22b to ensure proper periplasmic secretion upon expression in the *E. coli* system. Further analyses of the primary structure revealed that the P3 and P4 sites were well conserved between all ecotins suggesting similar protease targets, however, the P1 active site implicated to target the active site of the protease showed some variation between the species. Interestingly, similar to the *E. coli* ecotin, the *C. showae* contains a methionine at this position (Fig.1)[25], whereas the ecotin from *C. rectus* harbors a leucine in the P1 site (Fig .1), an amino acid with similar properties but a shorter side chain. This variation in the P1 active site could indicate differences in protease targets or no change in protein structure or binding.

2.3.2 Expression and purification of ecotins

To investigate their protease inhibition properties, ecotins from *E. coli, C. rectus* and *C. showae* were expressed as C-terminal 6xHisTag fusions and purified from *E. coli*. To ensure proper secretion into the periplasmic space, the native signal sequence of ecotin from *C. showae* and *C. rectus* was replaced with the *pelB* secretion signal. Analysis of ecotin purified at different time points after induction with IPTG by western blotting (Fig. 5) showed decreasing amounts of a signal peptide containing variant that migrates at a higher molecular weight (~18 kDa) after longer induction times. This indicates that processing of the *pelB*-ecotin sequence and translocation of the mature ecotin protein to the periplasmic space is slower when compared to the ecotin of *E. coli*. Therefore, cells grown for 24 hours after induction were used for purification of the heterologously expressed ecotin proteins.

2.3.3 Trypsin inhibition assay

Purified ecotin homologues from *C. showae* and *C. rectus* were tested for their ability to inhibit the serine protease, trypsin. Ecotin from *E. coli* previously shown to inhibit this protease[13] was used as a control. In the absence of ecotin, complete degradation of the *C. jejuni* multidrug efflux pump protein, CmeA was observed within 1 hour of incubation at 45ºC and within 3 hours of incubation at 37ºC. This indicated that CmeA partly unfolds at the higher temperature of 45ºC potentially exposing trypsin sites that are less accessible at the lower temperature of 37ºC. In the presence of ecotin from either *E. coli, C. rectus,* or *C. showae* no degradation of CmeA could be observed at 37ºC or 45ºC incubation over the time frame of the assay indicating that the *Campylobacter* ecotin homologues are indeed active in inhibiting the serine proteinase trypsin (Fig.6).

2.3.4 Protection against self-proteases HtrA and DegP

Next, we investigated the potential of ecotins to inhibit the self-protease DegP from *E. coli* and/or the orthologue, HtrA, from *C. jejuni*. First, we confirmed that CmeA is stable over the time frame of the assay. No degradation of CmeA could be observed after prolonged incubation at 37°C or 45°C (Fig. 11, lanes 2 to 5). In the presence of HtrA, approximately 50% of CmeA is degraded after 3 hours of incubation, and no intact CmeA could be detected after 6 hours of incubation, indicating that HtrA of *C. jejuni* is active after expression and purification from *E. coli* (Fig. 11, lanes 10 to 13). The presence of ecotin from *E. coli, C. rectus,* or *C. showae* had no effect on the proteolytic activity of HtrA (Fi. 11, lanes 14 to 25) since no difference in HtrA proteolysis of

CmeA could be observed in the presence of ecotin when compared to the absence of ecotin. Similar results were observed when DegP was used instead of HtrA (Fig. 12); DegP was able to completely degrade CmeA after 3 hours of incubation and the addition of the ecotins did not reduce the proteolytic activity of this serine protease.

2.3.5 FRET assay to measure *in vitro* protection efficiency of ecotin against factor Xa

In this study, we have adapted (based on the work of[26]) a high throughput Fluorescence Resonance Energy Transfer (FRET) (96-well) plate assay to investigate the protease inhibition properties of our ecotins from *E. coli*, *C. rectus* and *C. showae* against factor Xa. Our method uses a terminally labelled fluorophore/quencher peptide (Dabcyl-DQNATIDGRKQ-Edans) with Edans-fluorophore and Dabcyl-quencher carrying a factor Xa protease cleavage site (IDGR). In the absence of ecotin, factor Xa cleaves the peptide at the IDGR site, separating the fluorophore from the quencher resulting in increased levels of fluorescence (Fig. 7), whereas in the presence the quencher resulting in increased levels of fluorescence (Fig. 7), whereas in the presence of of ecotin the factor Xa should be inactivated and unable to cleave the peptide resulting in little to no fluorescence.

A time-and concentration dependent assay was performed to determine the optimal amount of ecotin and protease and to optimize the ecotin to protease ratio. To do so, the FRET peptide was incubated with ecotin or PBS (as a control) with the fluorescence analyzed before (t=0) and the relative fluorescence units (RFU) measured every 5 minutes over the time frame of 1 hour. The reactions contained increasing amounts of ecotin in combination with a constant amount of factor

Xa to determine protection efficiency. In parallel, reactions were performed that did not contain factor Xa and maintained a basal fluorescence (noise) of 1 to 10 RFU after incubation for 60 min.

First the *E. coli* ecotin containing reactions were monitored. In the absence of ecotin, 350 RFU were observed after 60 min of incubation with factor Xa indicating cleavage of the peptide (Fig. 7), whereas addition of equimolar amounts of *E. coli* ecotin resulted in 0 RFU after 60 min of incubation, indicating that no degradation has occurred. This clearly demonstrated that the ecotin effectively binds and inhibits factor Xa under these experimental conditions. Next, a titration curve was done using decreasing ratios of ecotin to factor Xa (Fig. 7). Here, an increase in fluorescence with decreasing amounts of ecotin was observed, indicating that factor Xa is not completely inhibited at lower ecotin concentrations.

Next, the *C. rectus* and *C. showae* ecotin containing reactions were monitored. The RFU of the reactions in the absence of ecotins were 350 at 60 min after factor Xa addition (Fig. 7). A 1:1 ratio of *C. rectus* or *C. showae* ecotin to factor Xa resulted in a RFU of ~50 and ~10 after 60 min with factor Xa, respectively. The final RFU were similar to the t=0 RFU of the reaction indicating that the peptide had not been degraded and that the *Campylobacter* ecotins effectively inhibit factor Xa. Titration curves as ratio of ecotins to factor Xa showed that with reduced amounts of ecotin, an increase in the fluorescence could be observed, indicating incomplete inhibition of factor Xa at low ecotin concentrations (Fig. 7).

In summary, it can be concluded that *E. coli* ecotin and ecotin homologs from oral *Campylobacter* species show comparable levels of factor Xa inhibition.

2.3.6 Ecotin inhibition of the serine protease neutrophil elastase

Ecotins were tested for their ability to inhibit NE. The reactions contained varying amounts of ecotin with a controlled amount (100 ng) of NE to produce a titration curve. First, reactions were performed in the absence of ecotin for 40 minutes with RFUs measured in 5 min increments, the PBS control (without elastase) was monitored for background RFU levels at all time points. An increase in RFU due to NE could be observed indicating cleavage of the substrate. All reactions that did not contain elastase maintained a basal fluorescence at 1 to 20 RFU over the timeframe of the assay.

Next, the inhibitory properties of ecotins were evaluated. The titration curve of the efficiency of the *E. coli* ecotin resulted in a decrease in active NE with increasing amounts (10 ng to 100 ng) of *E. coli* ecotin. In the absence of ecotin, $2.0x10^4$ RFU were observed after 40 min of incubation whereas at a 1:1 ratio of *E. coli* ecotin to NE no increase in RFU could be observed (Fig. 8). Similar results were observed when the inhibitory properties of *C. rectus* and *C. showae* ecotin homologues were monitored. In the absence of the *Campylobacter* ecotins, the RFU reached $2.0x10^4$ RFU after 40 min of incubation with NE (Fig. 8). A 1:1 ratio of *Campylobacter* ecotin to NE did not result in an increase after 40 min of incubation. The titration curve of the efficiency of the *C. rectus* and *C. showae* ecotin showed similar results when compared to the *E. coli* ecotin, with only 10 ng of active NE remaining with 100 ng ecotin present. Therefore, it can be concluded that both *Campylobacter* ecotins have similar inhibitory properties for NE when compared to the *E. coli* ecotin.

To confirm the above observation, a titration curve to determine the unbound portion of NE was performed. An increase in fluorescence to $1.1x10^4$ RFU was observed in the absence of

ecotin. As depicted in Figures 9 and 10, the endpoint titration shows that with increasing amounts of ecotins, the amount of free NE decreases. Overall, the titration curves indicate similar binding of NE by the *E. coli* and the *C. rectus* ecotins and potentially a slightly lower NE inhibition efficiency of the *C. showae* ecotin.

2.3.7 Neutrophil killing assay

To evaluate the protease inhibitory properties of ecotins *in vivo*, a whole cell neutrophil killing assay was employed taking advantage of an *E. coli* ecotin mutant complemented *in trans* with ecotins from *E. coli*, *C. rectus* and *C. showae*. First, $1x10^7$ cells of the respective strains, including the *E. coli* wild-type, were incubated with neutrophils for 0, 5, and 10 min. Serial dilutions of aliquots taken at each time point were plated to determine the cell survival rate. At t=10 min all *E. coli* ecotin mutant cells had been killed resulting in an incubation time of 481 minutes (Supplement Tables 1-3). In contrast, *E. coli* wildtype had $8x10^6$ cells remaining (80% survival) (Fig. 14). For the *E. coli* mutant complemented with its own ecotin *in trans*, $1x10^7$ CFU (100% survival) was observed at this time point, while expression of ecotins from *C. rectus* and *C. showae* resulted in a 85% ($8.5x10^6$ CFU) and 70% ($7.0x10^6$ CFU) survival rate, respectively. After 20 min of incubation with neutrophils a 30% survival rate of the *E. coli* wild-type ($3x10^6$ CFU) was observed, neutrophils incubated with the *E. coli* ecotin mutant complemented with ecotins from *E. coli*, *C. rectus* and *C. showae* resulted in 25% cell survival ($2.5x10^6$ CFU) for all three strains.

2.3.8 Neutrophil supernatant killing assay

The ability of ecotins to protect *E. coli* intact cells was further evaluated in killing assays with purified NETs. The concentration of elastase in the NETs was determined to be 3 ng/µl using the BioVision Neutrophil Elastase Activity Kit (#K383-100). In the PBS controls, similar CFU $1.0x10^7$ were present at the start and the endpoint of the assays for all strains (100% survival). The *E. coli* wild-type mixed with NETs showed $1.0x10^6$ cells remaining (10% survival), a significant decrease in survival when compared to the PBS control (Fig. 15a). The *E. coli* ecotin mutant had $1.0x10^4$ cells remaining (0.1% survival) when compared to the PBS control (Fig. 15b). The ecotin homologues (*E. coli, C. rectus* and *C. showae*) complemented back into the *E. coli* ecotin mutant rescued the survival rate to $1.0x10^6$ cells remaining (10% survival) when compared to the PBS control, versus the $1.0x10^4$ cells remaining (0.1% survival) seen in the ecotin mutant (Fig. 15c, d, e).

2.4 Discussion

In the mammalian host, the immune system employs an array of antimicrobial agents to help kill and remove invading or foreign bacteria. Neutrophils play a large part of the innate immune system and are one of the first cells to act on invading bacteria. The use of oxygen radicals is a method of attack, but the use of serine-proteases is equally important in the clearing of bacteria[12]. The importance of NE killing has been shown in mice that lack NE, but still contain the oxidative method of attack[27]. Among the survival strategies that bacteria employ to protect against the host-immune system are the formation of biofilms[28], protein glycosylation[8], a toxic radical

dissipation enzymes, and the expression of protease inhibitors. One prominent example of the latter is the periplasmic general serine-protease inhibitor ecotin.

Ecotin was first discovered in *E. coli* and it has been shown to be important in protecting against NE[14] and against other serine-proteases. Ecotin proteins form homodimers that are able to bind up to two protease molecules through the formation of a hetero-tetramer. The ability to inhibit a wide range of proteases is derived from two active sites. The primary active site contains hydrophobic amino acids like methionine or leucine that can bind the catalytic triad of many serine-proteases. The secondary binding site comes from the second ecotin molecule that binds non-specifically providing additional affinity for the protease [17]. These two points of contact mechanisms result in a strong binding affinity for a broad range of proteases. Ecotin is found in Gram-negative pathogenic bacteria that tend to encounter eukaryotic hosts. Bacteria such as *Pseudomonas, Shigella, Escherichia, Burkholderia* and *Klebsiella* have been identified to contain ecotin homologues[14]. In *P. aeruginosa,* ecotin has been shown to escape out of the periplasm and bind to biofilms helping protect the bacteria in the lungs from host-attack[17]. An ecotin homologue has also been discovered in the plant pathogen *Pantoea citrea*. This is of interest as this pathogen does not see a mammalian host system. Although, little is known about proteases outside the mammalian host it has been shown that the *P. citrea* ecotin is less potent in binding NE as compared to the *E. coli* ecotin suggesting that the plant serine protease inhibitor may have evolved to recognize alternate proteases specific to its host[14]. Here we present, for the first time the characterization of ecotin homologues present in oral *Campylobacter* species. Interestingly, the corresponding open reading frames are located within the protein glycosylation locus. N-glycosylation of proteins has already been shown to protect against chicken gut proteases for *C. jejuni*[8]. However, for the survival in the oral cavity of mammals, bacteria, would have to develop

means to protect themselves against neutrophil killing. So far it is unknown why certain *Campylobacter* species contain ecotin homologues while others i.e. the thermophilic *Campylobacter* species do not. It is possible that for oral *Campylobacter* species, N-glycosylation of surface proteins is not sufficient to ensure protection against proteases i.e NE the most prominent serine proteases found in the oral cavity and that is most likely one of the most important enzymes for oral *Campylobacter* species to inhibit to ensure survival in the mammalian host.

In order to gain insight on the function of the *Campylobacter* ecotin homologues we compared amino acid sequences against the characterized *E. coli* ecotin. The amino acid sequence shows low homology, but the theoretical structure could be modelled with 100% confidence. The P3-P4 positions are highly conserved between the species. Interestingly, the P1 active site showed variation, with *C. rectus* containing a leucine whereas the *E. coli* and the *C. showae* contain a methionine in P1. The methionine at P1 site has been hypothesized to be responsible for the broad specificity of the inhibitor since it fits into many primary binding sites of serine-proteases[29]. The leucine at P1 for *C. rectus* could be an indication for a more specific or different protease targeting ecotin. However, our results show that *C. rectus* and *C. showae* ecotins have nearly identical inhibitor properties toward the tested proteases, however, variations in the specificities towards other, untested proteases cannot be ruled out.

To gain insight into ecotin function and substrate specificity, we cloned, expressed and purified ecotins from *E. coli*, *C. rectus* and *C. showae* since i.e. *C. rectus* is a known periodontitis-causing pathogen[30]. *C. showae* is also considered an oral *Campylobacter* species, however, no clinical symptoms have been attributed to this species so far[31]. *E. coli* ecotin was used as a control and for assay development. Testing the potential to inhibit a wide range of serine-proteases, assays were developed to investigate inhibition of trypsin, factor Xa, NE and the bacterial, periplasmic

serine protease DegP from *E. coli* and HtrA from *C. jejuni*. The ability to inhibit trypsin was tested using a CmeA protection assay with trypsin present. The ecotin homologues of *C. rectus* and *C. showae* behave similar to the *E. coli* ecotin in their ability to inhibit trypsin, resulting in the protection of the CmeA protein. To investigate the inhibition of factor Xa, a FRET assay was developed modified from[32], here a FRET peptide only produces fluorescence when cleaved by active factor Xa , and when the protease is inhibited, fluorescence will be decreased. Indeed, we could demonstrate that the elevated levels of fluorescence in the presence of the protease completely disappears when ecotin and protease are mixed in a 1:1 molar ratio indicting that the hetero-tetramer observed for *E. coli* ecotin is also formed by the *Campylobacter* ecotins - indicating these ecotin variants are not only able to inhibit factor Xa but also that this inhibition most likely occurs via the same mechanism i.e formation of a hetero-tetrameric complex. Similarly, comparable inhibition rates of NE were observed for *C. rectus* and *C. showae* when compared to the *E. coli* ecotin indicating that they have similar equilibrium inhibitory constants K_i for *E. coli* ecotin towards NE = 0.012 nM ± 0.004 nM[14]. Moreover, the oral *Campylobacter* ecotin homologues behave similar to the *E. coli* ecotin *in vitro* in their ability to inhibit a wide range of serine-proteases (e.g. trypsin, factor Xa and NE).

This shows the broad range inhibition of serine-protease by ecotin homologues. The ecotins were not effective in protecting CmeA from degradation by the self-protease HtrA from *C. jejuni* or DegP from *E. coli*. The inability to inhibit the periplasmic serine proteases HtrA/DegP might stem from the ability of the proteins to form barrel shaped proteasomes[33]. This barrel could potentially prevent the ecotins from entering and inhibiting the active site of HtrA. This supports previous observations that ecotins protect against exogenous proteases[14].

Neutrophil elastase is a serine protease produced by activated neutrophils and secreted

during inflammation[34]. It is an important host-defense molecule of the innate immune system for the protection against pathogenic bacteria[35]. Bacteria that interact with the human host may have developed or evolved methods to help survival in human hosts allowing the bacteria to deal with serine proteases. Neutrophil elastase attacks the outer membrane protein A (OmpA) of *E. coli* inhibiting its growth[36]. NE cleaves the OmpA and allows NE access into the periplasm, where it cleaves periplasmic and inner membrane proteins resulting in loss of cell viability and inhibition of growth [14]. Ecotin may represent a way in which the bacterium is able to defend itself against the neutrophil powerhouses of the innate immune system. An ecotin deficient *E. coli* strain was generated to test the ability of ecotin to protect against whole cell neutrophil killing and purified NET killing. The ecotin deficient strain was complemented with the ecotin homologues from *C. rectus* and *C. showae*. The ecotin deficient strain was highly sensitive to neutrophil killing. The complemented ecotin homologues rescued the *E. coli* ecotin mutant back to wild-type levels of survival when mixed with intact neutrophils. Bacterial cells showed reduced fitness in the presence of neutrophils, but when ecotin was present there was prolonged survival of *E. coli*. Comparable results were seen in the assay completed with NETs, since all the ecotin homologues rescued growth back to wild-type levels. This indicated that ecotins help slow down and reduce the killing ability of neutrophils by inhibiting the NE.

In summary, the data presented here demonstrates that ecotins are a key protein expressed by *C. rectus* and *C. showae* for survival in the oral cavity of mammalian hosts. This protein possesses the ability to inhibit NE, trypsin, and factor Xa. Ecotin helps protect the bacteria from attack by NE that can cause damage to surrounding host cells releasing nutrients that can be consumed for growth. Oral *Campylobacter* species likely possess ecotin in addition to their N-

glycosylation systems to provide a competitive advantage in survival in the oral cavity niche containing high levels of neutrophil activity.

Acknowledgements

The authors thank Tracy Raivio at University of Alberta for the DegP expression plasmid and Kelly Moremen and Robert Maier at the University of Georgia for helpful discussions. CMS is an Alberta Innovates Strategic Chair in Bacterial Glycomics. The authors also thank the healthy volunteers for their blood donations and the staff of the UGA Health Center for drawing blood and their continuous support for our research.

Table 2.1: Strain 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-	
BL21		
RK212.2		
C. jejuni		
11168 NCTC	Clinical isolate used for genome sequencing	[37]
C. rectus		[30]
C. showae		[31]
C. hominis		
C. curvus		
C. gracilis		
Plasmids		
pET22B		
pET22B(rectus)		This study
pET22B(showae)		This study
pET22B(hominis)		This study
pET22B(curvus)		This study
pET22B(gracilis)		This study
pKD4		[21]
pKD46		[21]
CmeA		[22]
DegP		[23].
HtrA		This study

Table 2.2: Primers used in this study

Primer	Sequence
pKD4-ecotin-F	ATGAAGACCATTCTACCTGCAGTATTGTTTGCCGCTTTCGTGTAGGCTGGAGCTGCTTCG
pKD4-ecotin-R	TTAGCGAACTACCGCGTTGTCAATTTTCTCTTCCGCCTTCATGGGAATTAGCCATGGTCC
EC-ecotin-F	TAACCTTCAGCGACATCATCGG
EC-ecotin-R	AACCGGCTCGGGCGTTGGATGTC
Ecotin-Cre-NdeI-F	TTAGTGAGCATATGAGAAAAATTTTATTTGCTACGTTGGCTTTAGCGCCGATGC
Ecotin-Cre-XhoI-R	ATATCTCGAGTTTTGGCCTTTCTATTTTTGGTTTTATTGATTTTTCAAACC
Ecotin-Csho-NdeI-F	TTAGTGAACATATGAGGAAAATTTTACTTTTTATCGCGGCTTGCGCGTTGCCG
Ecotin-Csho-SalI-R	ATATGTCGACTTTATTTTTCCTTTTTAATTTTTTTGGTTCTATCG
Ecotin-Cho-NdeI-F	TTAGTTTTCATATGAGATTTTTTTTGATTTTTATTTTGGCGGTAAGTTTCAGTTTCG
Ecotin Cho-XhoI-R	ATATCTCGAGTTTGGCTTTTTTATCCAAAATTTCCGATTTTTCTAGC
Ecotin-Cgr-NdeI-F	ATTAGCAACATATGAGAAAAAGCGTATTTTTCTTTTTGTTGCCGCTG
Ecotin-Cgr-XhoI-R	ATATCTCGAGTTTCATCTTGCGGATTTTAGGCTCTTTGC
Ecotin-Ccu-NdeI-F	AATGTGAACATATGAGGAAAATTTTGTCTTTTTTGGCAGCTGCGACG
Ecotin Ccu-XhoI-R	ATATCTCGAGTTTTTCGACCTTTGCATTTTTTTCTTCAAATTTTTCATAAAGCC
htrA-NdeI-F	ATTAAATCATATGAAAAAGATTTTTTTATCATTAAGTTTAGC
htrA-XhoI-R	TTTCTCGAGTTTAAGCACAAGCAAAGTCGCAAAACC

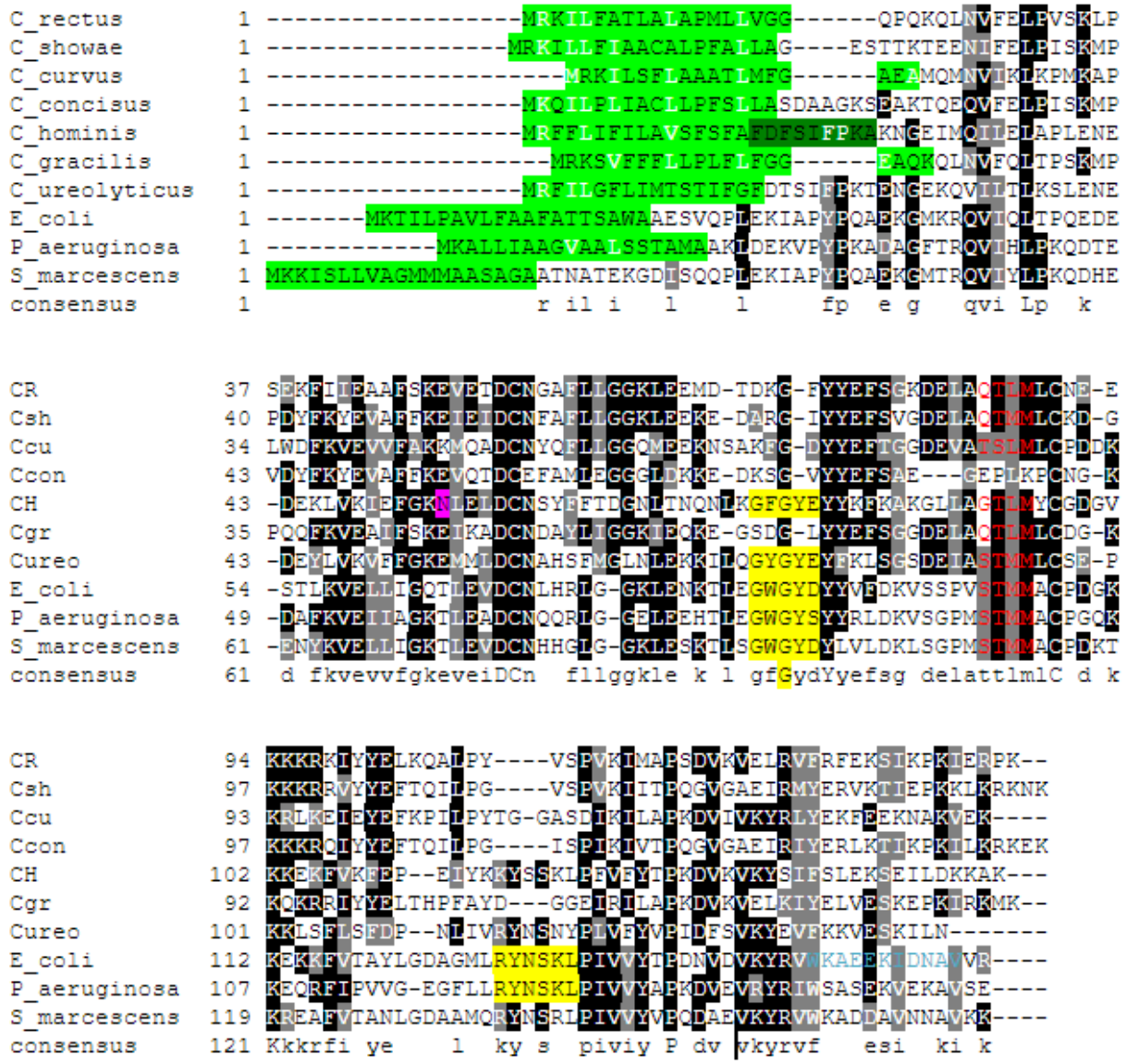


Figure 2.1: Box shade of *Campylobacter* species containing a homologue of ecotin compared to the *E. coli* ecotin and other ecotin sequences. Black shading indicated greater >50% amino acid identity. Grey shading is >50% similarity in amino acid charge. In red is the primary binding site. In yellow is the secondary binding site. In blue is the dimerization interface (polypeptide binding). In green is the signal peptide according to SignalP (cut-off 0.5, except for *C. ureolyticus*, here the signal has been empirically predicted since no signal peptide is predicted even with a cut-off of 0.3).

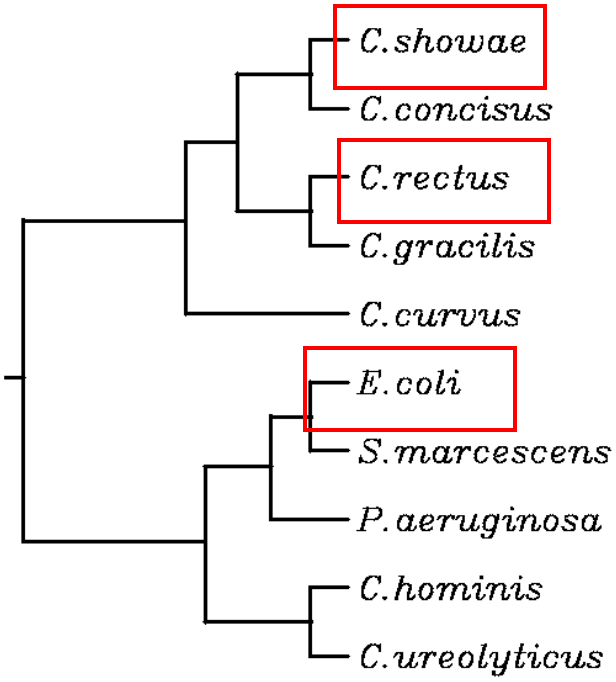


Figure 2.2: Ecotin orthologues separated by differences in protein sequence according to CLUSTALW alignment guide tree. The three ecotins compared in this study are highlighted in a red box.

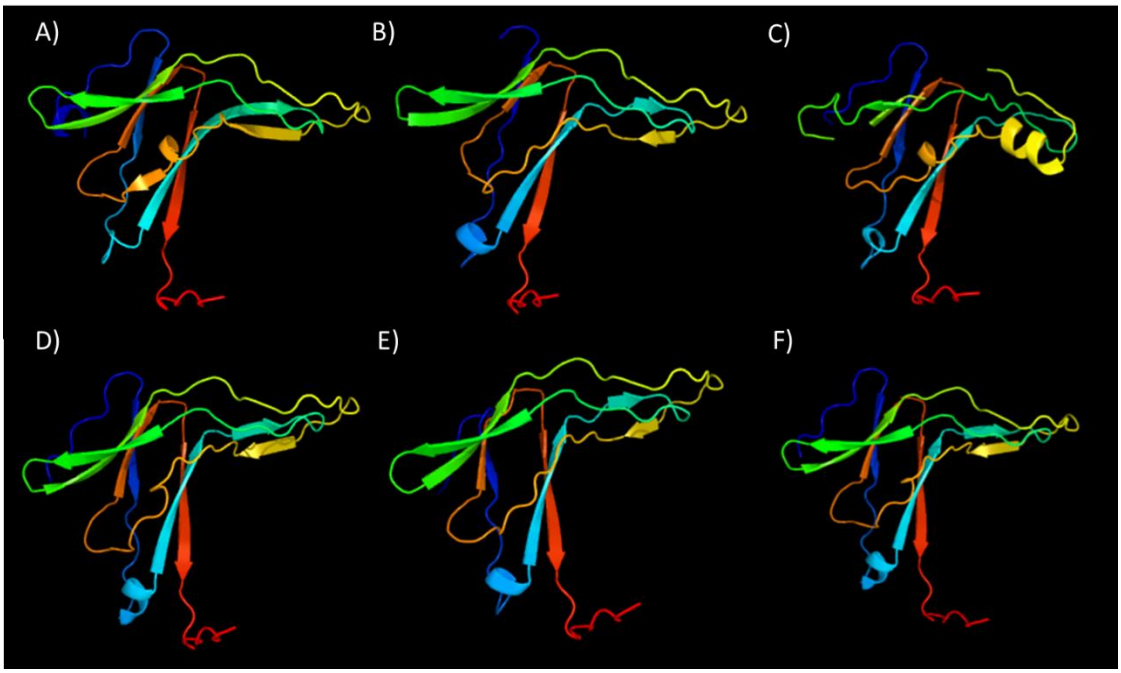


Figure 2.3. Analysis of ecotins using Protein Homology/analogy Recognition Engine V2.0 (PHYRE) by insertion of amino acid sequence to get theoretical 3D image of the proteins. The 3D images show very similar structures to each other. Comparing percent amino acid similarity to *E. coli* ecotin: *C. rectus*=27%, *C. showae*=33%, *C. curvus*= 34%, *C. gracilis*=31%, *C. concisus*=25%. A) *E. coli* ecotin B) *C. rectus* ecotin C) *C. showae* ecotin D) *C. concisus* ecotin E) *C. gracilis* ecotin F) *C. curvus* ecotin

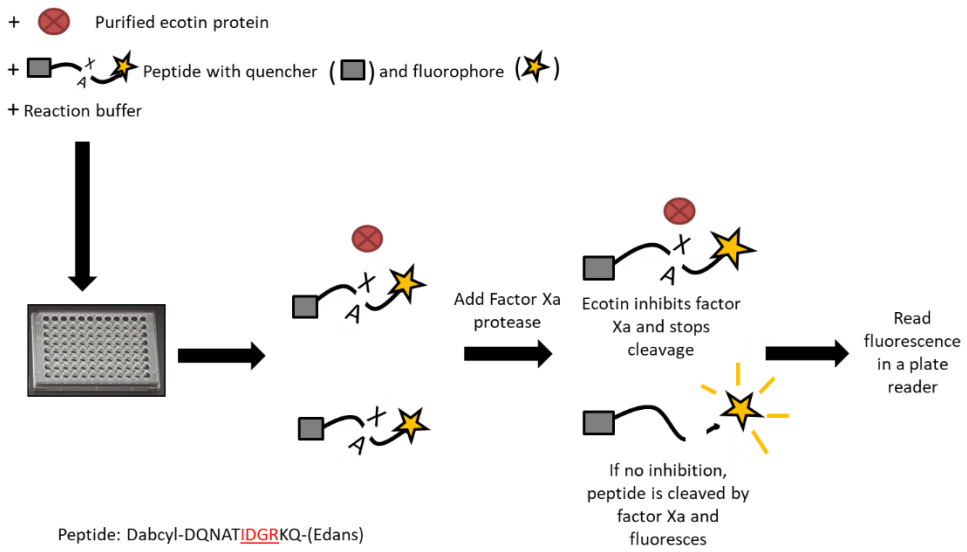


Figure 2.4. A diagrammatic illustration of the FRET assay. Peptides are incubated with purified ecotin homologues and aliquoted into 96-well plates for FRET analysis with or without factor Xa enzyme. If the peptide is cleaved, fluorescence is produced. factor Xa cut-site indicated by red amino acids in the peptide.

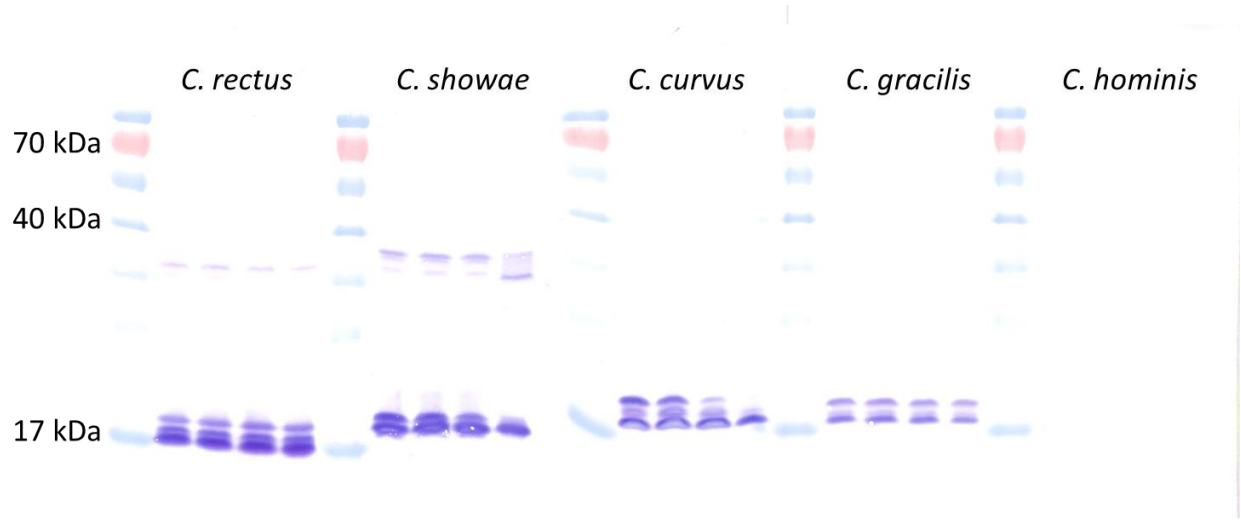
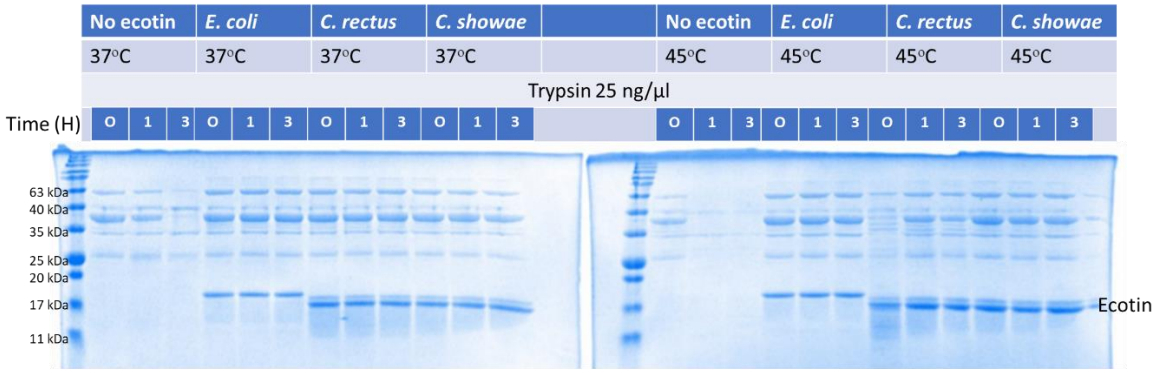


Figure 2.5. A western blot of cell lysates to identify production of *Campylobacter* ecotins expressed in *E. coli* BL21 collected 1, 2, 4 and 24 hours after induction with IPTG bound with anti-His antibodies. The band above 15 kDa, ~18 kDa represents the ecotin from each indicated *Campylobacter* species.



14

Figure 2.6. Coomassie stain of trypsin protease protection assay of ecotins. Assay was carried out at 37ºC and 45ºC to test ecotin inhibition stability. Samples contain 125 ng/µl CmeA (38 kDa), 25 ng/µl trypsin, 2.5 mM $MgCl_2$ and 200 ng/µl ecotin. Proteins not present in certain reactions were replaced by adding the same volume of H_2O. When ecotin is absent, there is degradation of CmeA by trypsin with an increase of degradation at higher temperatures. With ecotin present, there is no degradation of CmeA by trypsin at either temperature

.

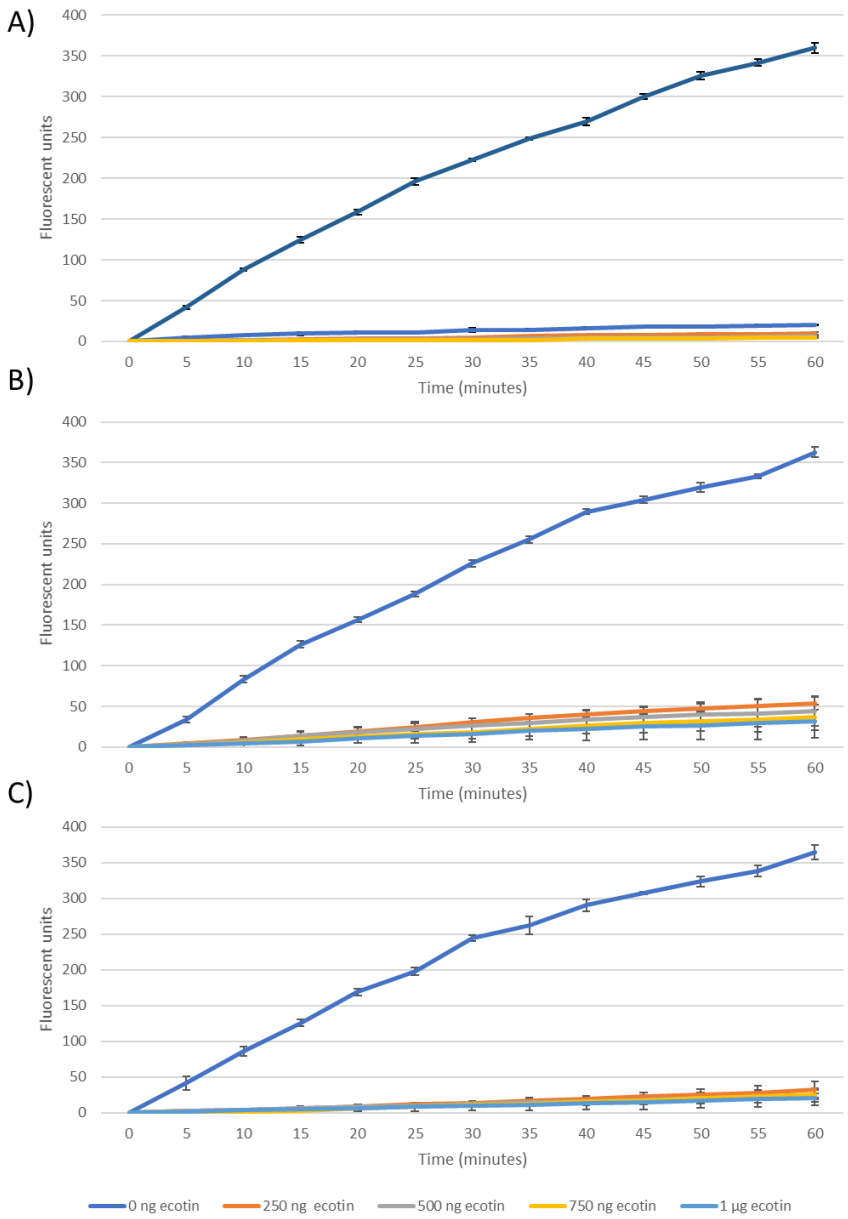


Figure 2.7. Factor Xa protection using the FRET assay (excitation=355nm, emission= 530nm). A) *E. coli*, B) *C. rectus* and C) *C. showae* ecotin protection assay with factor Xa. Only samples without ecotin showed increasing amounts of fluorescence indicating the FRET peptide was being cleaved by factor Xa. Samples containing purified ecotins showed very minimal to no fluorescence indicating that the FRET peptides are not cleaved and therefore maintaining the quencher in close contact with the fluorophore.

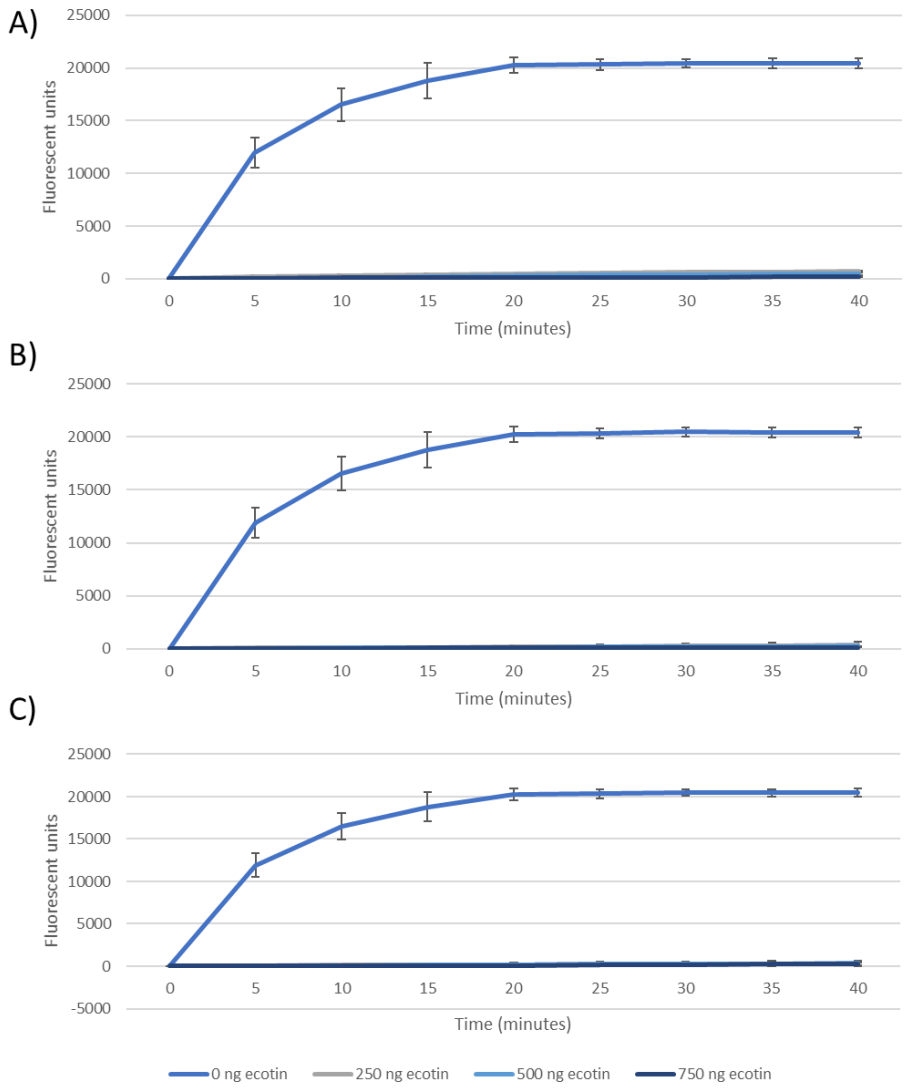


Figure 2.8. *E. coli*, *C. rectus* and *C. showae* ecotin protection assay with purified NE using a NE activity assay (BioVision). The assay detects the activity of the purified NE by its ability to cleave a substrate. When the substrate is cleaved by NE, a fluorophore is released. Fluorescence was measured at excitation= 380 nm and emission= 500 nm. Samples were mixed with 1 µg of NE and the indicated amounts of ecotin.

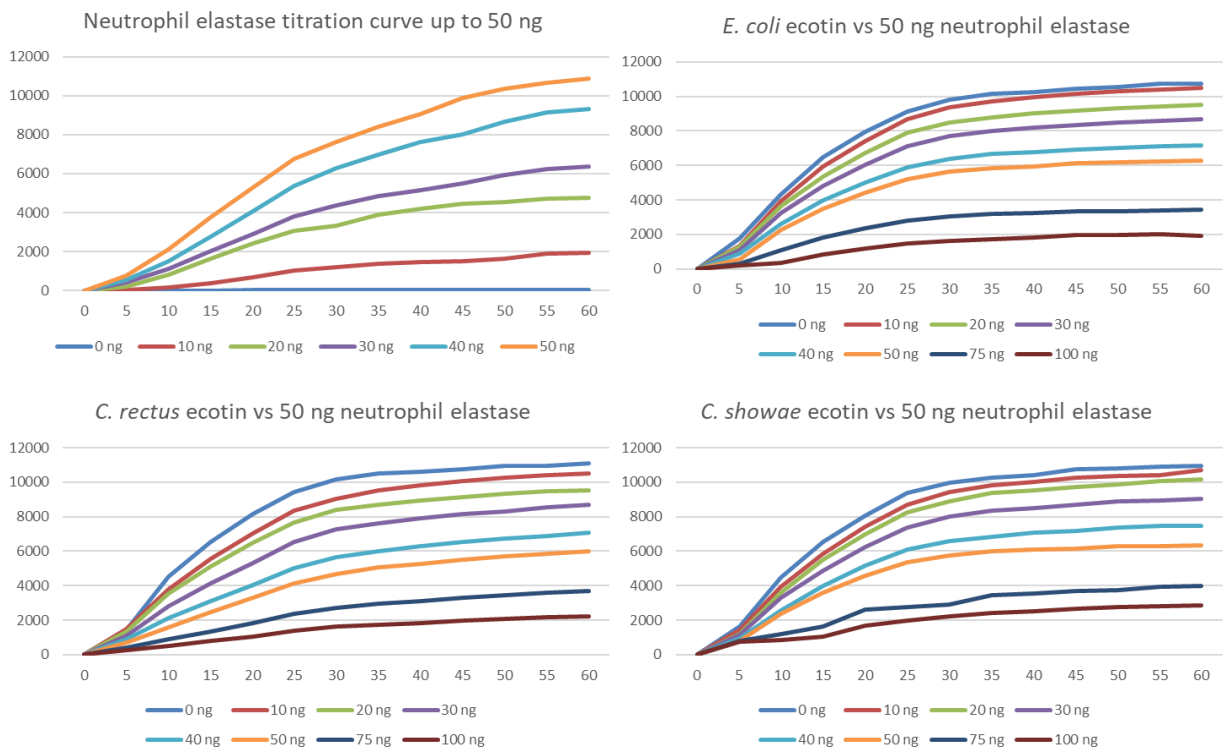


Figure 2.9. Titration assay of ecotin homologues with 50 ng of NE

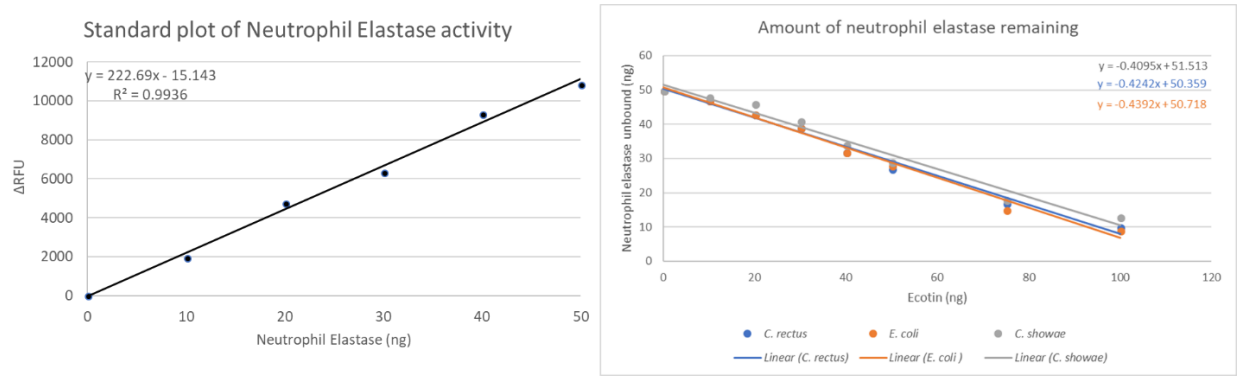


Figure 2.10. Amount of NE remaining linear plot of ecotin homologues binding to free NE.

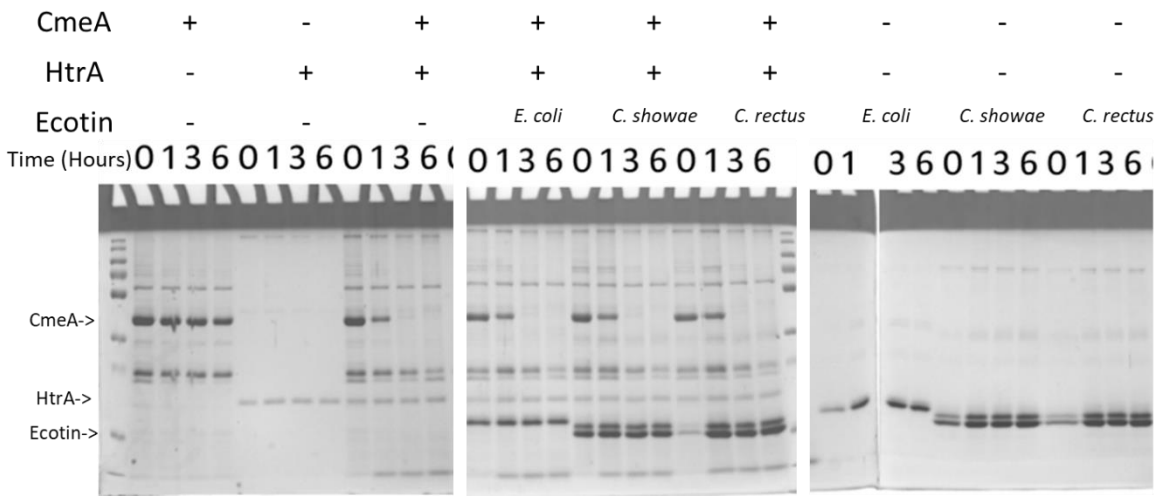


Figure 2.11. Self-protease protection assay with HtrA detecting degradation of *C. jejuni* CmeA protein by coomassie stained SDS-PAGE.

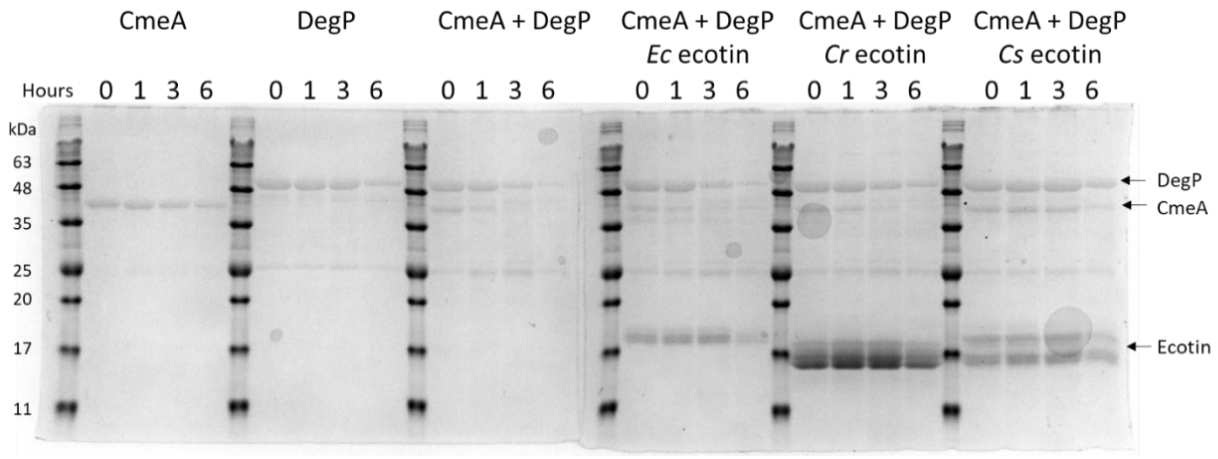


Figure 2.12. Self-protease protection assay with *E. coli* DegP detecting degradation of *C. jejuni* CmeA protein by coomassie stained SDS-PAGE.

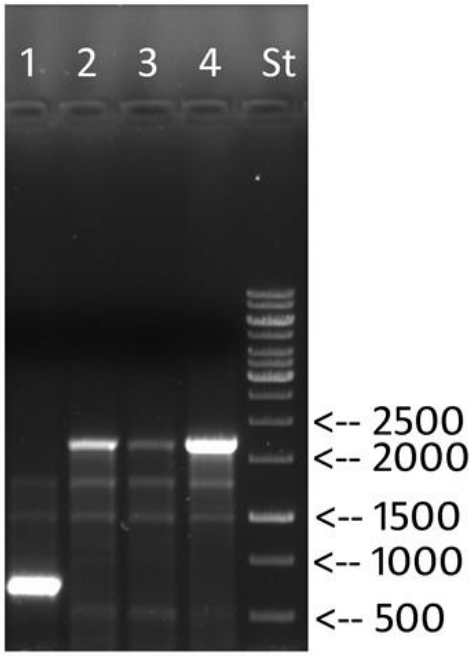


Figure 2.13. Confirmation of the *E. coli* ecotin knock-out mutant/kan cassette insertion by PCR. Lane 1, PCR with wild-type DNA, lanes 2 to 4, PCR with three mutant candidates. St = DNA ladder, relevant markers are indicated on the left in bp.

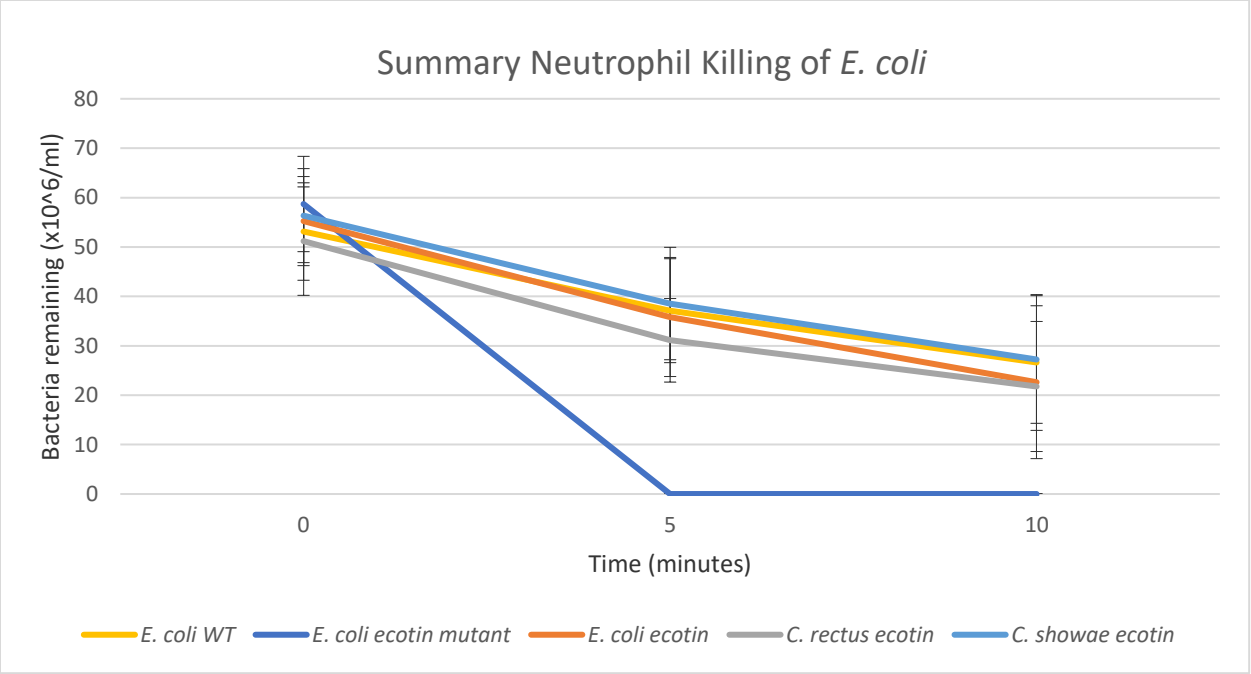


Figure 2.14. Neutrophil killing assay of *E. coli* WT, ecotin mutant, ecotin mutant complemented with either *E. coli*, *C. rectus* or *C. showae* ecotin on pET22B. A 9:1 ratio of bacteria to PMN cells were mixed with washed neutrophil cells and incubated at 37ºC for 0, 5 and 10 minutes. Samples were mixed with saponin designed to break open neutrophil cells and to release alive bacteria. Cells were inoculated in LB in a 96 well plate and run on a kinetic cycle in a plate reader detecting OD_{600}. A standard curve for each strain was created and data plotted to show how many bacteria remained in the sample. Error bars represent the standard deviation of the mean between experiments.

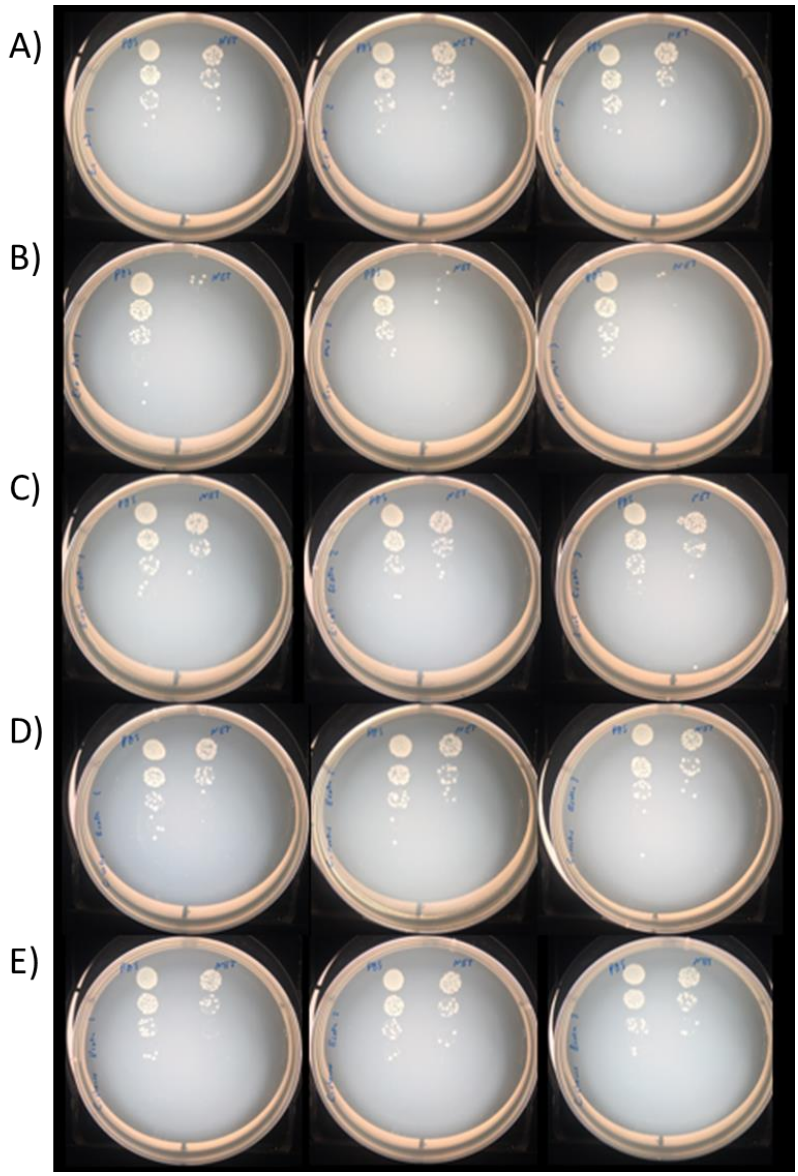


Figure 2.15. A) *E. coli* wild-type mixed with neutrophil NET supernatant solution. B) *E. coli* ecotin mutant mixed with neutrophil NET supernatant solution. C) *E. coli* ecotin mutant complement with *E. coli* ecotin mixed with neutrophil NET supernatant solution. D) *E. coli* ecotin mutant complement with *C. rectus* ecotin mixed with neutrophil NET supernatant solution. E) *E. coli* ecotin mutant complement with *C. showae* ecotin mixed with neutrophil NET supernatant solution.

2.5 References:

1. Heutinck KM, ten Berge IJ, Hack CE, Hamann J, Rowshani AT. Serine proteases of the human immune system in health and disease. *Mol Immunol*. 2010;47(11-12):1943-1955.

2. Korkmaz B, Horwitz MS, Jenne DE, Gauthier F. Neutrophil elastase, proteinase 3, and cathepsin G as therapeutic targets in human diseases. *Pharmacol Rev*. 2010;62(4):726-759.

3. Lee WL, Downey GP. Leukocyte elastase: Physiological functions and role in acute lung injury. *American journal of respiratory and critical care medicine*. 2001;164(5):896-904.

4. Shapiro SD. Proteinases in chronic obstructive pulmonary disease. *Biochemical Society Transactions*. 2002.

5. Moraes TJ, Chow C, Downey GP. Proteases and lung injury. *Crit Care Med*. 2003;31(4):S194.

6. Owen CA. Roles for proteinases in the pathogenesis of chronic obstructive pulmonary disease. *International journal of chronic obstructive pulmonary disease*. 2008;3(2):253.

7. Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P. Evidence for a system of general protein glycosylation in *campylobacter jejuni*. *Mol Microbiol*. 1999;32(5):1022-1030.

8. Alemka A, Nothaft H, Zheng J, Szymanski CM. N-glycosylation of *campylobacter jejuni* surface proteins promotes bacterial fitness. *Infect Immun*. 2013;81(5):1674-1682.

9. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *campylobacter* genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219.

10. Macuch PJ, Tanner A. *Campylobacter* species in health, gingivitis, and periodontitis. *J Dent Res*. 2000;79(2):785-792.

11. Stapels DA, Geisbrecht BV, Rooijakkers SH. Neutrophil serine proteases in antibacterial defense. *Curr Opin Microbiol*. 2015;23:42-48.

12. Belaaouaj A, McCarthy R, Baumann M, et al. Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis. *Nat Med*. 1998;4(5):615.

13. Chung CH, Ives HE, Almeda S, Goldberg AL. Purification from *escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. *J Biol Chem*. 1983;258(18):11032-11038.

14. Eggers CT, Murray IA, Delmar VA, Day AG, Craik CS. The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. *Biochem J*. 2004;379(Pt 1):107.

15. Ireland PM, Marshall L, Norville I, Sarkar-Tyson M. The serine protease inhibitor ecotin is required for full virulence of *burkholderia pseudomallei*. *Microb Pathog*. 2014;67:55-58.

16. Laskowski Jr M, Kato I. Protein inhibitors of proteinases. *Annu Rev Biochem*. 1980;49(1):593-626.

17. Tseng BS, Reichhardt C, Merrihew GE, et al. A biofilm matrix-associated protease inhibitor protects *pseudomonas aeruginosa* from proteolytic attack. *mBio*. 2018;9(2):543.

18. Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: A common cause of persistent infections. *Science*. 1999;284(5418):1318-1322.

19. Wang G, Chen H, Xia Y, et al. How are the non-classically secreted bacterial proteins released into the extracellular milieu? *Curr Microbiol*. 2013;67(6):688-695.

20. Webb JS, Thompson LS, James S, et al. Cell death in *pseudomonas aeruginosa* biofilm development. *J Bacteriol*. 2003;185(15):4585-4592.

21. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*. 2000;97(12):6640-6645.

22. Feldman MF, Wacker M, Hernandez M, et al. Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in *escherichia coli*. *Proc Natl Acad Sci U S A*. 2005;102(8):3016-3021.

23. Raivio TL, Silhavy TJ. Transduction of envelope stress in *escherichia coli* by the cpx two-component system. *J Bacteriol*. 1997;179(24):7724-7733.

24. Rada BK, Geiszt M, Kaldi K, Timar C, Ligeti E. Dual role of phagocytic NADPH oxidase in bacterial killing. *Blood*. 2004;104(9):2947-2953.

25. McGrath ME, Hines WM, Sakanari JA, Fletterick RJ, Craik CS. The sequence and reactive site of ecotin. A general inhibitor of pancreatic serine proteases from *escherichia coli. J Biol Chem*. 1991;266(10):6620-6625.

26. Rodems SM, Hamman BD, Lin C, et al. A FRET-based assay platform for ultra-high density drug screening of protein kinases and phosphatases. *Assay and drug development technologies*. 2002;1(1):9-19.

27. Belaaouaj A, McCarthy R, Baumann M, et al. Mice lacking neutrophil elastase reveal impaired host defense against gram negative bacterial sepsis. *Nat Med*. 1998;4(5):615.

28. Costerton JW, Stewart PS, Greenberg EP. Bacterial biofilms: A common cause of persistent infections. *Science*. 1999;284(5418):1318-1322.

29. McGrath ME, Hines WM, Sakanari JA, Fletterick RJ, Craik CS. The sequence and reactive site of ecotin. A general inhibitor of pancreatic serine proteases from *escherichia coli. J Biol Chem*. 1991;266(10):6620-6625.

30. Tanner ACR, Haffer C, Bratthall GT, Visconti RA, Socransky SS. A study of the bacteria associated with advancing periodontitis in man. *J Clin Periodontol*. 1979;6(5):278-307.

31. Etoh Y, Dewhirst FE, Paster BJ, Yamamoto A, Goto N. *Campylobacter showae* sp. nov., isolated from the human oral cavity. *Int J Syst Bacteriol*. 1993;43(4):631-639.

32. Rodems SM, Hamman BD, Lin C, et al. A FRET-based assay platform for ultra-high density drug screening of protein kinases and phosphatases. *Assay and drug development technologies*. 2002;1(1):9-19.

33. Boehm M, Lind J, Backert S, Tegtmeyer N. *Campylobacter jejuni* serine protease HtrA plays an important role in heat tolerance, oxygen resistance, host cell adhesion, invasion, and transmigration. *European Journal of Microbiology and Immunology*. 2015;5(1):68-80.

34. Pham CT. Neutrophil serine proteases: Specific regulators of inflammation. *Nature Reviews Immunology*. 2006;6(7):541.

35. Reeves EP, Lu H, Jacobs HL, et al. Killing activity of neutrophils is mediated through activation of proteases by K flux. *Nature*. 2002;416(6878):291.

36. azzaq Belaaouaj A, Kim KS, Shapiro SD. Degradation of outer membrane protein A in *escherichia coli* killing by neutrophil elastase. *Science*. 2000;289(5482):1185-1187.

37. Parkhill J, Wren BW, Mungall K, et al. The genome sequence of the food-borne pathogen *campylobacter jejuni* reveals hypervariable sequences. *Nature*. 2000;403(6770):665.

Chapter III

The fate of N-glycans and fOS in *C. jejuni*

3.1 Introduction

Campylobacter jejuni is a Gram-negative, microaerophilic bacterium that belongs to the epsilon class of proteobacteria [1]. *C. jejuni* is a major cause of bacterial gastroenteritis in humans worldwide [2]. *C. jejuni* is commonly found in avian species like chickens, where it develops a commensal relationship with the host [1]. Infections usually happen through the consumption of undercooked chicken, or contaminated water [3]. The gastroenteritis caused by *Campylobacter* is usually self-limiting and is not typically treated with antibiotics [4]. *C. jejuni* can also cause secondary complications including Guillain-Barré syndrome (GBS), where the bacteria mimics gangliosides of nerve cells resulting in an auto-immune attack of these cells [5]. With the secondary complications or infections in immunocompromised patients, antibiotic therapy can become necessary. The usual antibiotic treatment for *C. jejuni* is macrolides and fluoroquinolones, but there has been increased reports describing antibiotic resistance to these antimicrobials which has become a public health concern [4,6]. *C, jejuni* has a few resistance mechanisms for the broad range of antimicrobials [7-9], but the major cause of antimicrobial resistance comes from the drug efflux pump, CmeABC [4]. Glycosylation of CmeABC has been shown to be required for function of the efflux pump [10].

All *Campylobacter* species possess a conserved *pgl* locus that allows for N-linked glycan modification of proteins and production of free oligosaccharides (fOS) [11]. The machinery of the Pgl enzymes are located in the cytoplasm and inner membrane. Nucleotide-activated sugars are assembled on a lipid carrier Und-P. The assembled oligosaccharide is then flipped into the periplasm, where the OTase (PglB) can transfer the oligosaccharide onto proteins or hydrolyze and release the oligosaccharide as fOS.

Disruption of the *pgl* genes results in multiple pleiotropic effects in *C. jejuni*. Currently, there are more than 60 proteins that have been shown to be co- or post-translationally modified with N-glycans in *C. jejuni* with many more expected to be glycosylated [12]. Knockout mutants of the *pgl* pathway disrupt adherence and invasion of bacteria *in vitro* and mouse and chicken colonization *in vivo* [11,13,14]. N-glycosylation influences many aspects of pathogenesis, from stability of proteins to protection against proteases [7,15]. The various roles that N-glycans play is still not fully understood. The production of fOS in *C. jejuni* appears to be an intentional process rather than just a byproduct of inefficient transfer of glycans by PglB since fOS plays a role in osmotic stability [16,17], but it is unknown if it plays any other role in the cell. It has been shown previously that aminosugars (eg. GalNAc and GlcNAc) can inhibit the release of elastase and reactive oxygen species (ROS) from human polymorphonuclear leukocytes [18,19]. It could be possible that N-glycans and fOS containing aminosugars on the non-reducing end could potentially play a role in inhibiting elastase and ROS from host immune cells.

Quality control mechanisms for N-glycosylated proteins in prokaryotes is unknown but in eukaryotes there are a few methods for breakdown and recycling N-glycans and fOS. One of the mechanisms is the endoplasmic reticulum associated degradation (ERAD) pathway. The ERAD system is dedicated to recycling N-linked glycoproteins that are misfolded [20,21]. When

glycoproteins get targeted for degradation, the protein is retro-translocated into the cytosol where a specific set of proteins help remove the N-glycans, so the protein can be degraded by the 26S proteasome [22]. An N-glycanase (PNGase) removes the N-glycans by cleaving the nitrogen bond between the GlcNAc and the asparagine releasing the glycan as fOS in the cytosol [23-26]. The gene encoding PNGase (*Ngly1* in humans and *Png1* in yeast) is widely distributed throughout eukaryotes [27]. In humans, mutations in the *Ngly1* gene have phenotypic consequences for patients resulting in neurological dysfunction, abnormal tear production, and liver disease [28-30]. Interestingly a knockout of another glycanase: endo-β-N-acetylglucosaminidase (ENGase) rescues these phenotypic consequences but the effects are not fully understood currently. ENGase cleaves the N-glycans between the first two GlcNAcs releasing fOS and an N-GlcNAc-modified protein [31,32]. These N-GlcNAc proteins have trouble being degraded by the 26S proteasome resulting in protein aggregation resulting in these *Ngly1* disorders [33,34].

It is well known how *C. jejuni* makes these glycans and assembles them, but nothing is currently known about the quality control and recycling of the glycans in *C. jejuni*. There is no N-glycanase homologue that has been found in the *C. jejuni* genome. Also, the reducing end sugar for *C. jejuni* is a di-N-acetylbacillosamine as compared to a GlcNAc found in eukaryotes. An enzyme to cleave the di-N-acetylbacillosamine bound to asparagine has not been currently found.

Another possible mechanism of recycling is nonspecific degradation by reactive oxygen species. ROS include superoxide radicals, hydrogen peroxide and hydroxyl radicals that are formed from the reduction of oxygen [35]. If cellular iron is present, the ferrous iron can react with hydrogen peroxide through what is known as the Fenton reaction to generate hydroxyl radicals, the strongest oxidant in an aqueous environment [35]. ROS can damage DNA, proteins and lipids

non-specifically [36] and likely also carbohydrates. *C. jejuni* has a set of enzymes to help protect against the damaging effects of oxidative stress. Three of these enzymes are KatA (catalase), SodB (superoxide dismutase) and AhpC (alkyl-hydroxyperoxidase) that protect against oxidative stress [37-39]. As seen in figure 3.1. The creation of mutants in any of these three enzymes can result in the build-up of ROS within *C. jejuni* [40]. The use of non-specific reactive oxygen species could be a potential method of degrading the N-glycosylated proteins and fOS when no enzyme is present to remove the oligosaccharide.

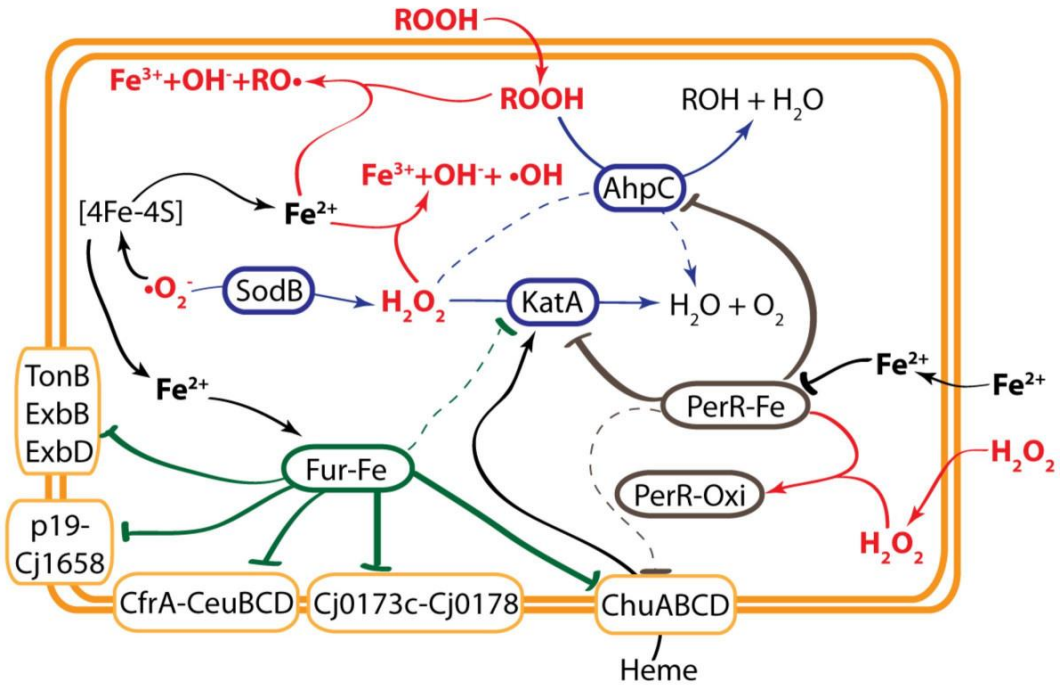


Figure 3.1. Schematic representation of oxidative stress metabolism in *C. jejuni*. Figure reproduced from Palyada *et al.* (2009) with permission[40].

3.2 Materials and methods

3.2.1 Bacterial strains and growth conditions

C. jejuni NCTC 11168 [41] was grown on BHI agar or BHI broth for 18 hours under microaerobic conditions. *C. rectus* RM3267 [42] and *C. showae* ATCC 51146 [43] were grown under anaerobic conditions as previously described [16]. *C. jejuni pglB, sodB, katA* and *ahpC* mutants were grown in the presence of kanamycin at a final concentration of 25 μg/mL.

3.2.2 Ethanol fOS extraction and analysis by TLC

Bacterial cells were harvested from agar or liquid and suspended in BHI. Cultures were centrifuged at 4700 g for 30 minutes at 4^{o}C. The pellets were suspended and washed with 4 mL of deionized water per gram of wet cell pellet (WCP). The cells were centrifuged at 4700 g for 30 minutes and the supernatant removed. The obtained pellets were resuspended in 75% ethanol at 1.5 mL per gram WCP. The resuspension was incubated at 70^{o}C for 30 minutes. The solution was centrifuged at 16000 g at RT and the supernatant was collected and diluted to 20% ethanol with deionized water and frozen at -80^{o}C and lyophilized. The obtained pellet was resuspended in cold methanol, 500 μL per gram of WCP. Sample was vortexed for 2 minutes and centrifuged at 16000 g for 30 minutes at RT. The supernatant was collected and evaporated in a rotatory speed vacuum. The obtained pellet was resuspended in 120 μL deionized water per WCP. Samples had 5 μL spotted onto a TLC plate (aluminum backed, silica coated) and run with a mobile phase of 3:3:2 of acetic acid: n-propanol: water. The carbohydrates were visualized using p-anisaldehyde staining [44].

3.2.3 Analysis of fOS by HPAEC-PAD

The fOS samples extracted from *C. jejuni* as described above were used. Samples were passed through a PGC cartridge (Extract Clean TM SPE Carbp 150 mg/ 4 mL, Grace Davison Discovery Sciences) as described previously [45]. The samples were lyophilized and resuspended in 100 μL of deionized water and 44 μL of the fOS preparation in the presence of 4 M trifluoroacetic acid (TFA) for 2 hours at 100ºC. TFA was removed by a rotary speed vacuum and adjusted to a final volume of 220 μL. GalNAc standards were prepared to a final concentration of 0, 25, 50, 100 and 150 μM and hydrolyzed in TFA as before. The samples were analyzed by HPAEC-PAD on a Dionex ICS3000 system equipped with a CarboPac®PA100 (9 x 250 mm) coupled with a PA100 guard column (3x 50 mm) at a flow rate of 1.0 mL/min by loading 25 μL. The amount of fOS (nmoles) was determined by the GalNAc peak areas of TFA hydrolyzed fOS divided by 5 and values plotted against the GalNAc standard curve.

3.2.4 Extraction of fOS from TLC plates for analysis by MS/MS

One hundred μL of fOS was loaded onto a TLC as described above. The samples were scratched off the plate using a scalpel and resuspended in 1 mL deionized water. The sample was centrifuged 16000 g for 30 minutes and the supernatant was extracted and lyophilized. The obtained pellet was resuspended in 30 μL deionized water and analyzed by MS/MS as follows. Resuspended fOS was spotted onto a Bruker Daltonic MTP Ac800 AnchorchipTM target plate and air dried. A volume of 0.65 μL of 2,5-dihydroxybenzoic acid (DHB, 10 mg/mL in 80% H2O and 20% MeOH containing 0.1% trifluoroacetic acid, TFA) was spotted on top and allowed to dry. Mass spectra were obtained in the positive mode of ionization using a Bruker Daltonics (Bremen, GmbH) UltrafleXtreme MALDI TOF/TOF mass spectrometer. The FlexAnalysis,

BioTools and Sequence Editor software packages provided by the manufacturer were used for analysis of the mass spectra. The MS/MS spectra were obtained manually with CID (collisional induced dissociation) set to off. Elemental composition of analytes was determined using a Bruker Daltonics (Billerica, MA) Apex Qe 9.4T FTICR MS instrument using the MALDI source.

3.2.5 Degradation of fOS by the Fenton-reaction

The fOS purified as described above was used in a Fenton reaction. fOS was mixed with 2M H_2O_2 and 13.6 g/L of $FeSO_4$ at a 1:1 ratio of 5 µL each with 10 µL fOS and incubated at room temperature for 1, 2, 3, 4 hours. The controls were included with fOS and mQH_2O or $FeSO_4$ or H_2O_2. The samples were loaded onto a TLC as described above. Lower concentrations of Fenton reaction components were also tested. The concentrations of hydrogen peroxide and $FeSO_4$ for the experiment can be seen in Table 1. The samples were incubated for 1 hour at room temperature and loaded onto a TLC as described above.

Table 3.1. Concentrations of H_2O_2 and $FeSO_4$ for the Fenton reaction for degradation of fOS.

Tube	fOS	H_2O_2	$FeSO_4$
1	10 µL	2 M	13.6 g/L
2	10 µL	200 mM	1.36 g/L
3	10 µL	20 mM	0.136 g/L
4	10 µL	2 mM	0.0136 g/L
5	10 µL	0.2 mM	0.00136 g/L
6	10 µL	0 mM	0 g/L

3.3 Results

3.3.1 Analysis and quantification of fOS by HPAEC-PAD

Purified fOS was analyzed and quantified by HPAEC-PAD that separates carbohydrates based on charge. The TFA hydrolysis of fOS results in the breakdown of oligosaccharides into

monosaccharides, were they can be quantified using appropriate monosaccharide standards (in this case GalNAc). A GalNac TFA hydrolyzed monosaccharide standard was used to identify and accurately determine the concentrations of fOS per gram of WCP (data not shown). *C. jejuni* TFA hydrolyzed samples were compared back to the GalNAc standard to determine the concentration of fOS per gram of WCP with two sample runs one by agar growth and one by liquid broth growth (Fig 3.2B, 3.3B). The calculated fOS per gram of WCP was plotted in a bar graph (Fig 3.2C, 3.3C). *C. jejuni* wildtype had 0.28 mg of fOS per gram of WCP for agar growth and had 0.29 mg of fOS per gram of WCP for liquid growth in broth. The *C. jejuni pglB* mutant had below <0.05 mg fOS-like material per gram of WCP in both liquid and agar growth. The *C. jejuni katA* mutant had 0.15 mg per gram of WCP for agar growth and 0.13 mg per gram of WCP for liquid growth, containing half the amount of fOS as the *C. jejuni* wildtype. The *C. jejuni ahpC* mutant produced 0.25 mg per WCP for agar growth and 0.15 mg per gram of WCP in liquid growth. The *C. jejuni sodB* mutant had varying results between the two runs with 0.16 mg of fOS per gram of WCP in agar and 0.3 mg per gram of WCP in liquid BHI.

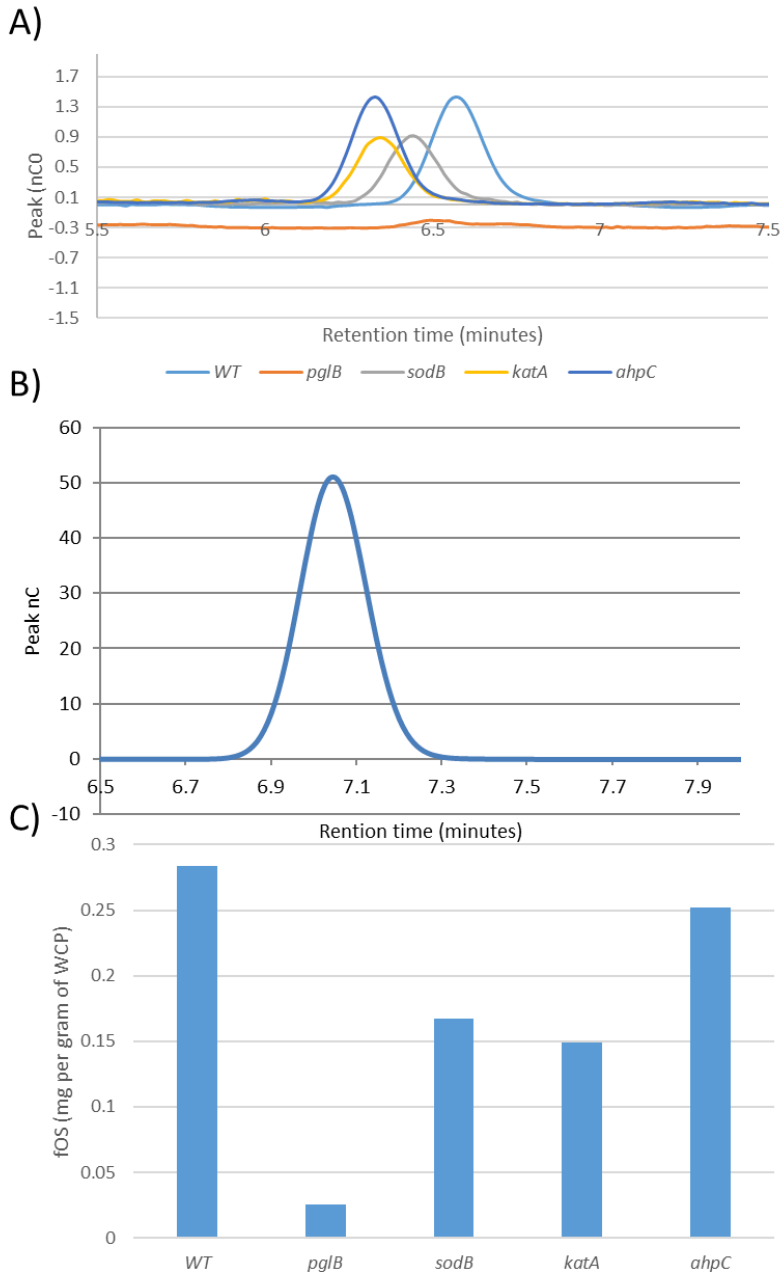


Figure 3.2. Analysis of fOS by HPAEC-PAD from *C. jejuni* 11168 ROS mutants grown on BHI agar. A) fOS from the different *C. jejuni* WT and mutants hydrolyzed by TFA and analyzed by HPAEC-PAD. B) GalNAc monosaccharide standard incubated with TFA for comparison to *C. jejuni* fOS. C) Bar-graph of the amount of fOS in mg per gram of wet cell pellet. Data summarized from HPAEC-PAD data from figures 3.8-3.13.

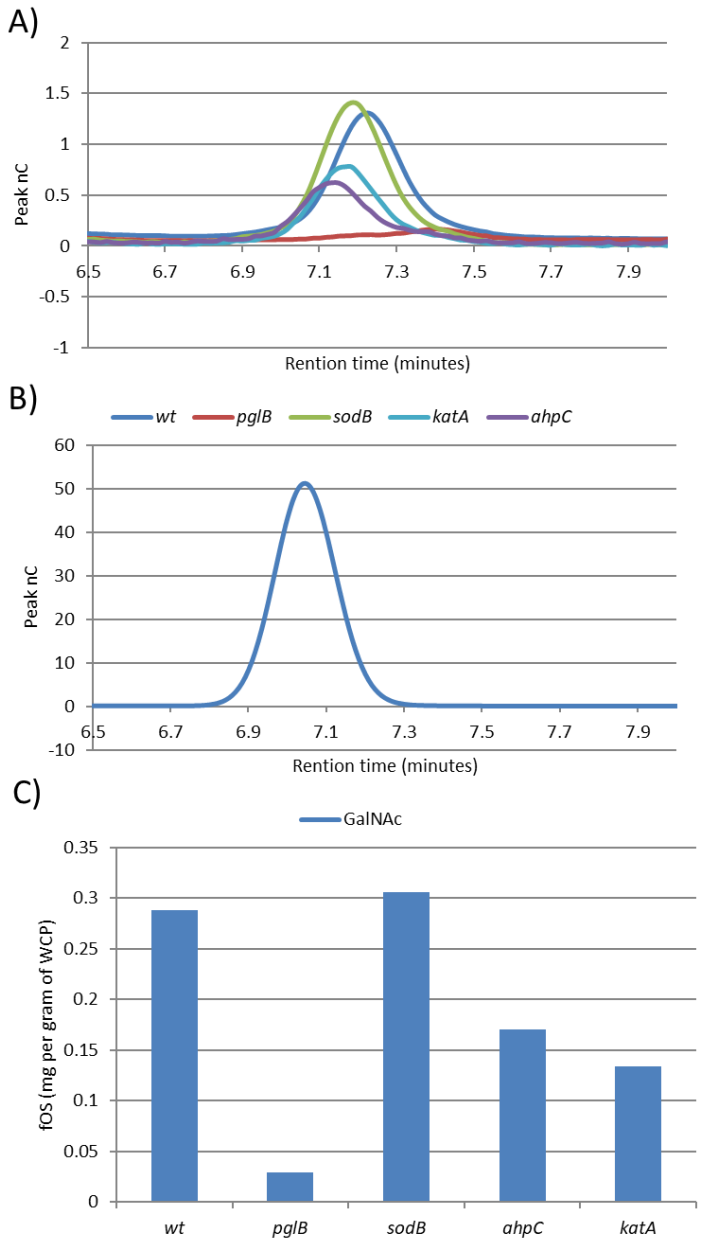


Figure 3.3. Analysis of fOS by HPAEC-PAD from *C. jejuni* 11168 ROS mutants grown in BHI broth. A) Sample of all fOS from the different *C. jejuni* WT and mutants broken apart by TFA and analyzed by HPAEC-PAD. B) GalNAc monosaccharide standard incubated with TFA for comparison of *C. jejuni* fOS. C) Bar-graph of the amount of fOS in mg per gram of wet cell pellet. Data summarized from HPAEC-PAD data from figures 3.14-3.19.

3.3.2 Degradation of fOS by Fenton reaction

We next wanted to determine if fOS could be degraded by hydrogen peroxide through the Fenton reaction. *C. jejuni* fOS has a R_f value of 0.36 when run on the TLC under the conditions described (Fig. 3.4). The fOS control is shown to be stable over the course of the assay of 18 hours. When fOS is mixed with hydrogen peroxide or iron sulfate ($FeSO_4$) alone it is stable and not broken down over the course of the assay. The Fenton reaction at 2 M hydrogen peroxide breaks down fOS in under 1 hour. We lowered the concentration of the Fenton reaction (Fig. 3.5A) to observe if breakdown of fOS occurs at the lower concentrations in 1 hour. There was some degradation seen with a lighter spot at the concentration of 200 mM hydrogen peroxide but no degradation of fOS was observed at 20 mM and lower concentrations.

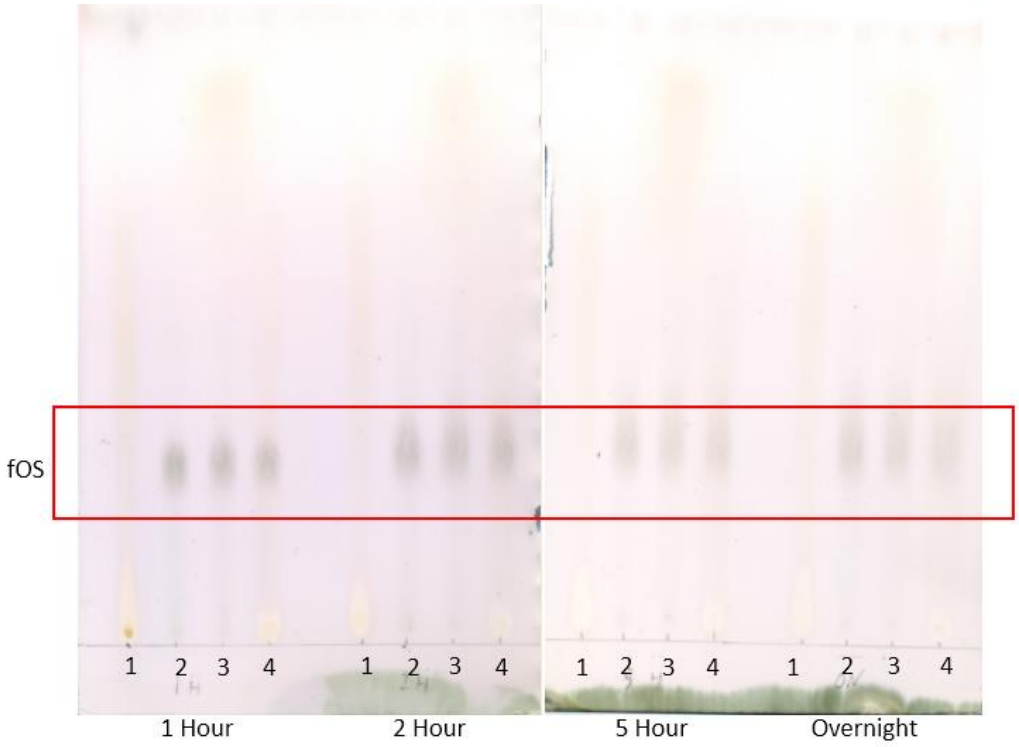


Figure 3.4. Degradation of fOS by reactive oxygen species. Lane 1= fOS + 2M H_2O_2 + 13.6 g/L $FeSO_4$ (Fenton reaction). Lane 2= fOS. Lane 3= fOS + 2M H_2O_2. Lane 4= fOS + 13.6 g/L $FeSO_4$. Reaction was run at room temperature for 1, 2, 5 and 18 hours (overnight). *C. jejuni* fOS has a R_f value of 0.36 and is indicated by the red box.

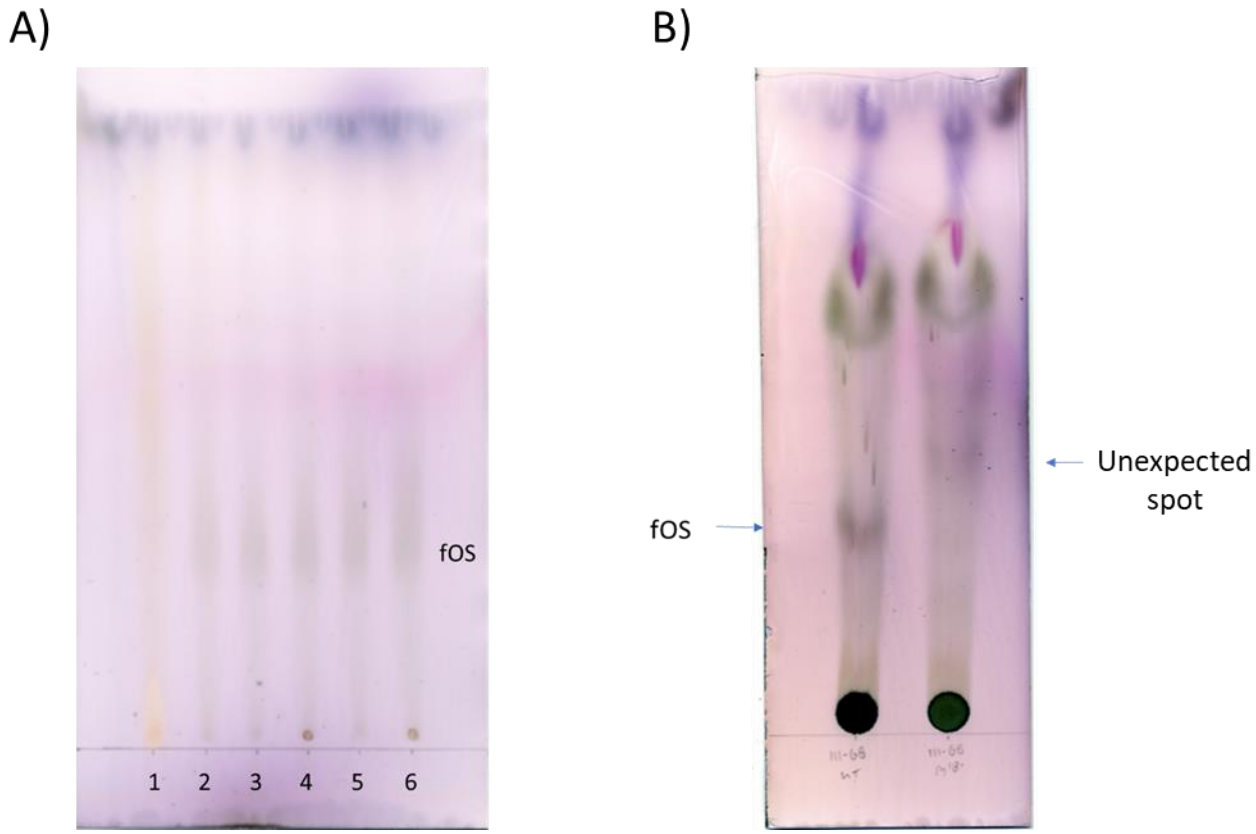


Figure 3.5. A) Degradation of fOS by lowering the concentrations of the Fenton reaction compounds. Lanes are as described in Table 1. Samples were incubated at room temperature for 1 hour. B) TLC plate of extracted fOS from *C. jejuni* 11168 WT and *pglB*- mutant. There is a unexpected TLC spot seen in the *C. jejuni pglB*- mutant. The *C. jejuni* WT fOS has a R_f value of 0.36 and the *pglB*- mutant has a spot with an R_f value of 0.40, indicated by the arrows.

3.3.3 *C. jejuni pglB*- mutant oligosaccharide accumulation analysis.

The TLC fOS spots were extracted from the *C. jejuni* WT and *pglB*- mutant (Fig .3.5B). The TLC showed the wildtype fOS at the regular R_f value of 0.36. The *pglB*- mutant had a spot appear at the R_f value of 0.40 that did not show up on the WT TLC. The analysis of the WT spot by mass spectrometry (Fig 3.6) gave a mass to charge value of 1446 m/z, consistent with the mass of the heptasaccharide of 5 GalNAc, 1 glucose and 1 di-NAcBac. Mass spectrometry of the *pglB*- mutant (Fig. 3.7) produced a mass to charge value of 812. This could be the result of 3 HexNAc residues and 1 Hex residue, totaling 812 m/z.

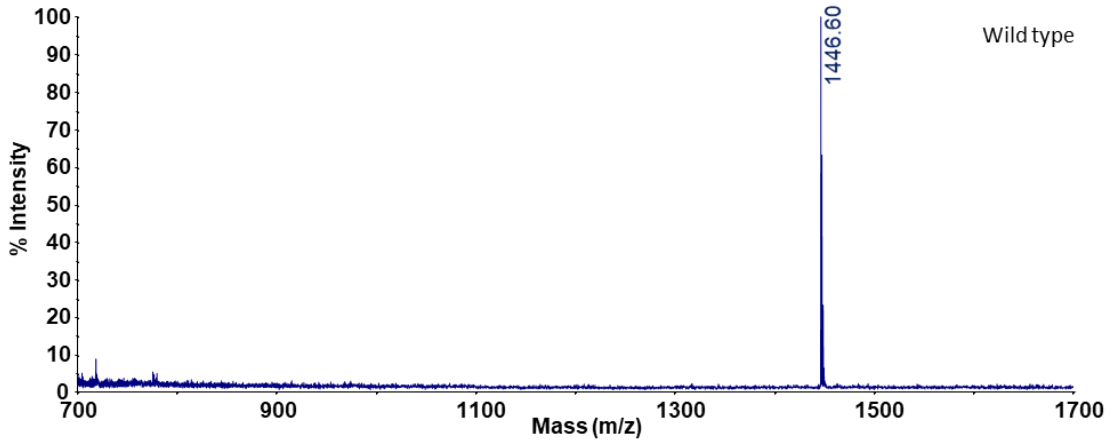


Figure 3.6. Mass spectrometry analysis of *C. jejuni* 11168 WT fOS from the TLC shown in Figure 3.5. Mass to charge (m/z) had a size of 1446 equal to that of the *C. jejuni* heptasaccharide.

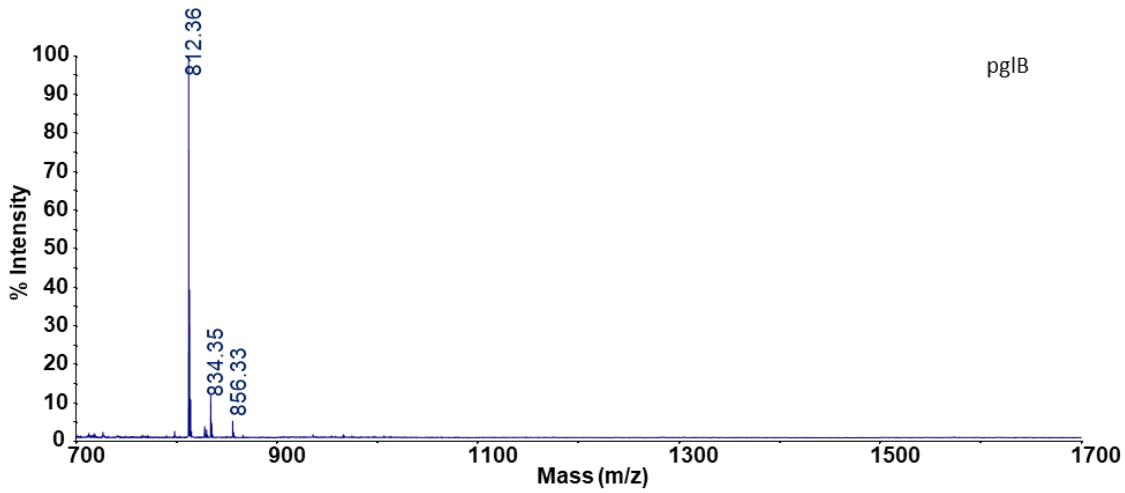


Figure 3.7. Mass spectrometry analysis of *C. jejuni* 11168 *pglB*- mutant fOS from the TLC shown in Figure 3.5. Mass to charge (m/z) of the TLC spot is 812.

3.4 Discussion

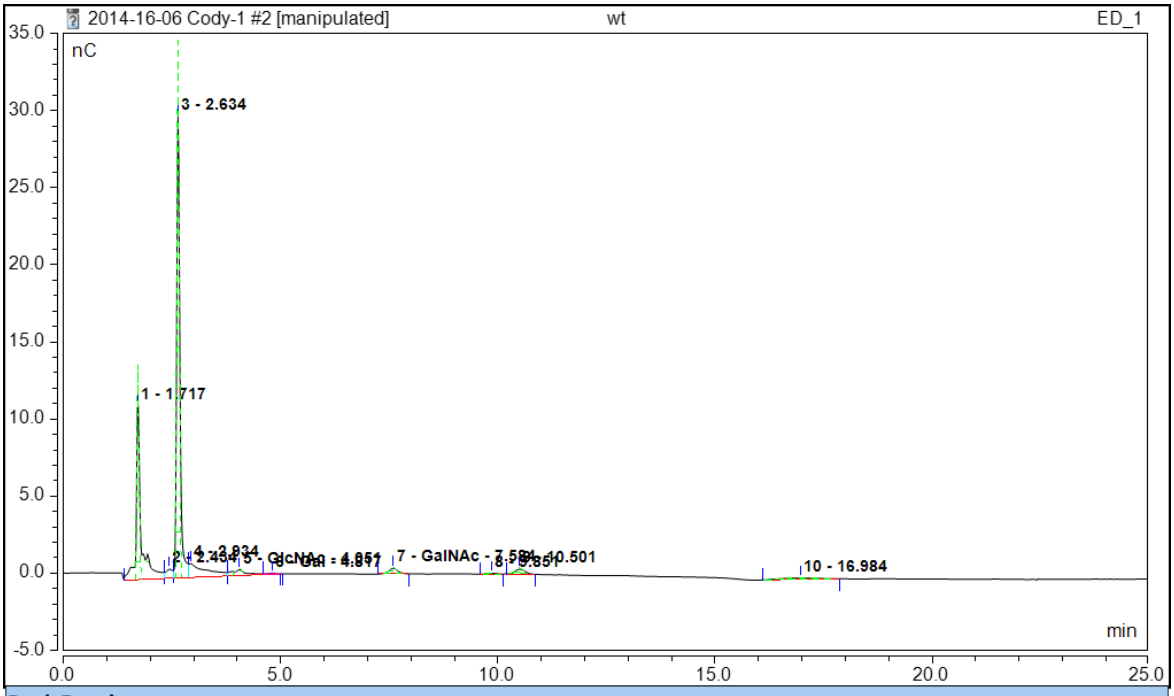
All *Campylobacter* species have a *pgl* locus and produce N-linked glycoproteins and release fOS into the periplasmic space with the same oligosaccharide as their N-glycans [46,47]. Glycoproteins are important to the biology of *C. jejuni* since inactivation of the N-linked protein glycosylation pathway leads to decreased pathogenicity. The ratio of fOS to N-linked glycans in *C. jejuni* under standard laboratory growth conditions has been determined to be 10:1 [48]. Without the OTase (PglB) there is no production of N-glycans or fOS. Bioinformatics analysis of the *C. jejuni* genome does not identify any obvious glycosidases to degrade N-linked glycans or fOS. ROS and RNS might be possible sources for glycan breakdown.

ROS and RNS are very reactive compounds that contain free radicals and have been known to depolymerize cellular structures, including oligosaccharides encountered by immune cells. Analyzing the effects of reactive oxygen species on fOS production in *C. jejuni* can be

done using mutants in oxygen stress pathway to generate persistent ROS for the *C. jejuni* mutants. The SodB enzyme is involved in the conversion of superoxide into hydrogen peroxide that can be further modified by KatA into water and oxygen or used by AhpC to form alcohols and water. Analysis of fOS production by the *sodB* and *ahpC* mutants showed variable amounts but overall, there was a reduction in the levels of fOS produced. The *katA* mutant demonstrated the largest and most consistent reduction of fOS production. This could be because without a method of removing hydrogen peroxide we could see reactive oxygen species accumulating in the cell and causing a general stress response. We know that hydrogen peroxide itself is not enough to degrade fOS, but with intracellular iron, the Fenton reaction could occur and break down fOS in the periplasmic space. Although, a high concentration of Fenton reagents was required to degrade fOS, if the reagents were in close proximity to the fOS the concentration may be appropriate to degrade fOS or even N-glycans non-specifically. These results suggest that ROS can possibly play a role in glycan recycling, however the precise mechanism of glycan degradation remains elusive. Demonstrating that the *pglB*- mutant produces undetectable levels of the fOS heptasaccharide is consistent with there being no OTase to transfer the heptasaccharide from the lipid carrier to proteins or to hydrolyze the LLO to release fOS. This would cause an accumulation of oligosaccharides bound to Und-P on the inner cell membrane. Reducing the levels of Und-P could be detrimental to the cell as this lipid is also needed for peptidoglycan biosynthesis [49]. A bacterial cell may need to find a way of dealing with the oligosaccharide accumulation to allow production of peptidoglycan and cell survival. The analysis of the TLC spot from the *pglB*- mutant that is not seen in the wildtype suggests a tetrasaccharide structure is present. This tetrasaccharide has a mass that could possibly be three GalNAc and 1 glucose residues. It is unknown why the *pglB*- mutant would produce the

tetrasaccharide, but there could be a method of removing the oligosaccharide from the Und-P allowing recycling and some non-specific cleavage that provides Und-P for peptidoglycan production.

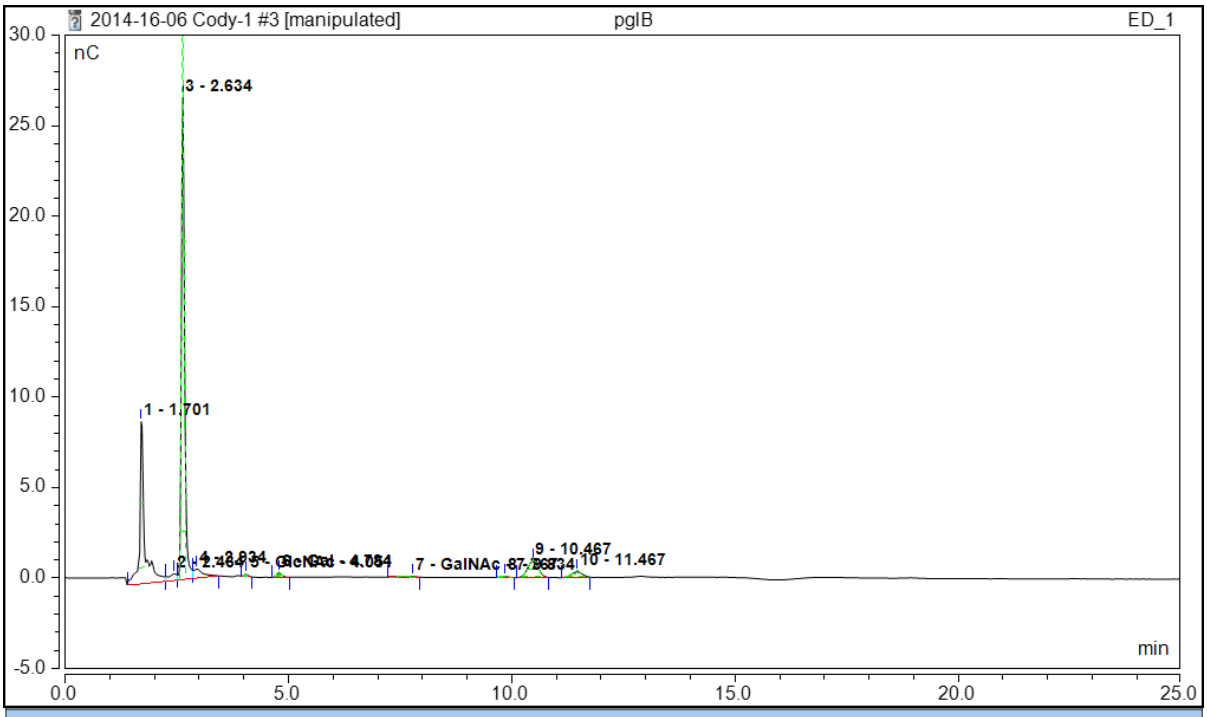
There is further research that is needed for discovering the fate of fOS in *C. jejuni*. Replication of the experiments with the reactive oxygen species mutants followed by HPAEC-PAD analysis and production of an increased amount of the *C. jejuni pglB* mutant TLC spot for mass spectrometry and NMR structure confirmation would help to verify the reported results. Also, developing a method to label the fOS and N-glycans would help follow the glycans through their lifetime, that can be done with stable isotope labeling.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.717	0.075	BM	n.a.	1.34	2934
2		2.434	n.a.	M	n.a.	n.a.	n.a.
3		2.634	0.075	M	n.a.	1.49	6889
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4		2.934	n.a.	M	n.a.	n.a.	n.a.
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5	GlcNAc	4.051	n.a.	MB	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	Gal	4.817	n.a.	Rd	n.a.	n.a.	n.a.
7	GalNAc	7.584	0.213	BMB*	5.83	1.07	7054
8		9.851	0.246	BMB	1.49	1.08	8877
9		10.501	0.268	BMB	5.97	1.01	8531
10		16.984	1.014	BMB	n.a.	0.99	1555

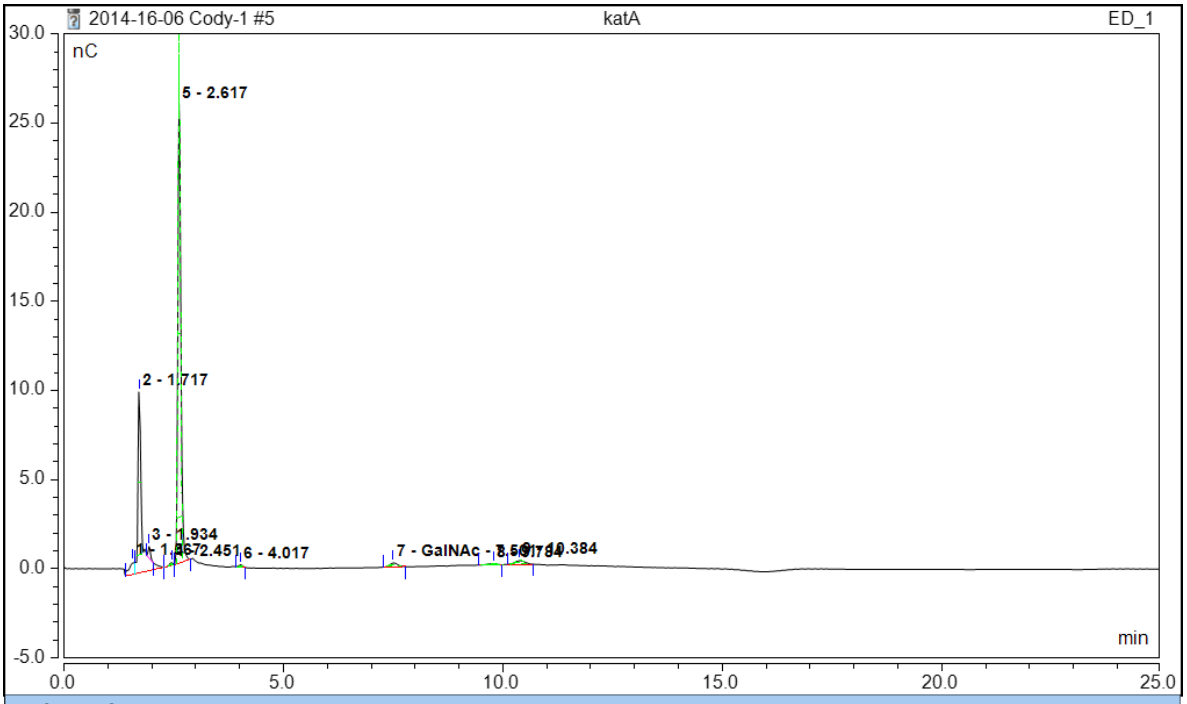
Figure 3.8. HPAEC-PAD supplementary data of *C. jejuni wildtype* grown on BHI agar.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.701	0.070	BM	n.a.	1.35	3263
2		2.434	n.a.	M	n.a.	n.a.	n.a.
3		2.634	0.075	M	n.a.	1.27	6900
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4		2.934	n.a.	MB	n.a.	n.a.	n.a.
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5	GlcNAc	4.051	0.092	BMB	4.08	1.11	10700
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	Gal	4.784	0.120	BMB	9.44	1.19	8787
7	GalNAc	7.767	0.253	BMB*	5.25	0.72	5238
8		9.834	0.212	BMB*	1.57	1.07	11965
9		10.467	0.265	BMB	2.19	1.02	8652
10		11.467	0.275	BMB	n.a.	0.95	9665

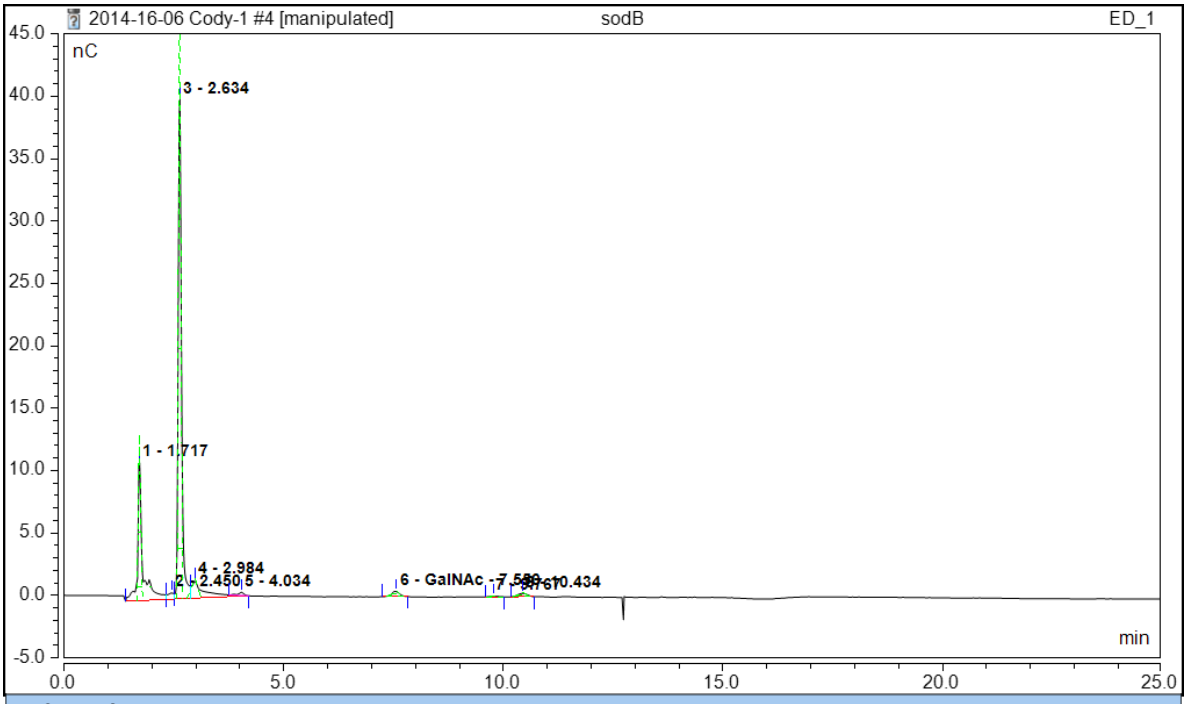
Figure 3.9. HPAEC-PAD supplementary data of *C. jejuni pglB*- grown on BHI agar.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.567	n.a.	BM	n.a.	n.a.	n.a.
2		1.717	0.072	M	5.17	n.a.	3189
3		1.934	n.a.	Rd	n.a.	n.a.	n.a.
4		2.451	0.096	Mb	1.16	0.90	3626
5		2.617	0.074	bMB	9.21	1.36	7014
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6		4.017	0.106	BMB	13.82	1.09	7989
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7	GalNAc	7.501	0.192	BMB	5.22	1.11	8495
8		9.784	0.325	BMB	1.21	0.78	5033
9		10.384	0.260	BMB	n.a.	1.03	8803

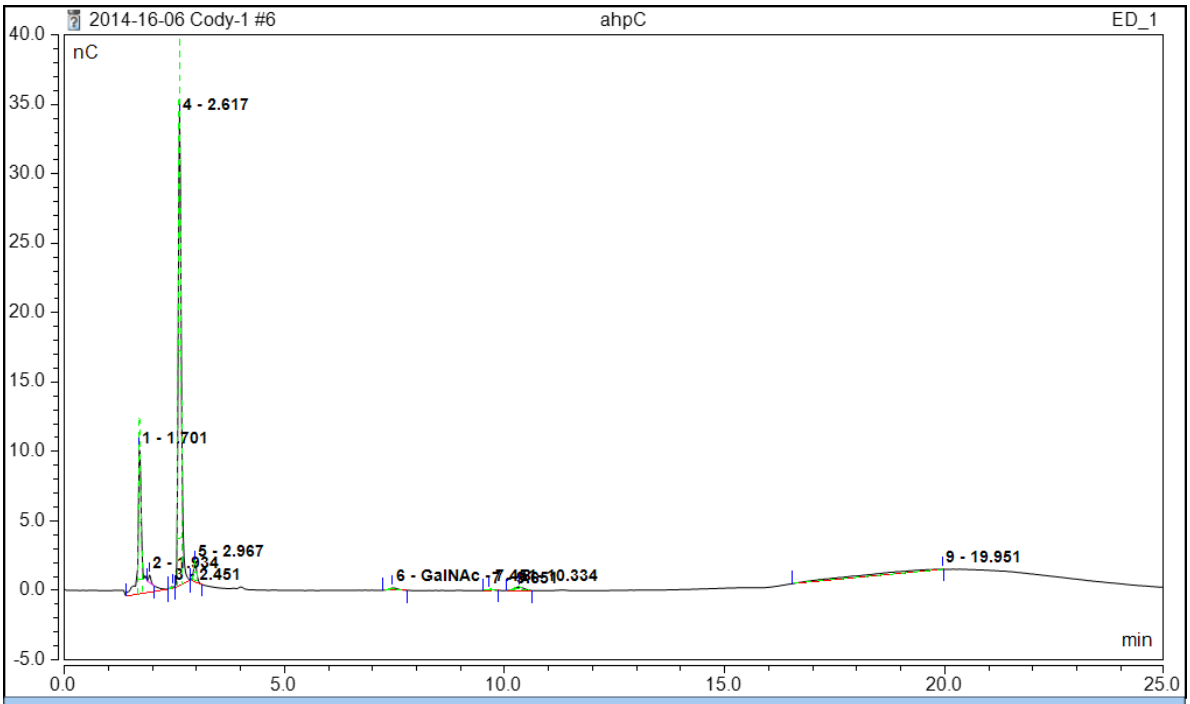
Figure 3.10. HPAEC-PAD supplementary data of *C. jejuni katA*- grown on BHI agar.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.717	0.073	BM	n.a.	1.42	3093
2		2.450	n.a.	M	n.a.	n.a.	n.a.
3		2.634	0.076	M	n.a.	1.25	6724
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4		2.984	n.a.	MB	n.a.	n.a.	n.a.
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5		4.034	n.a.	Rd	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	GalNAc	7.550	0.194	BMB*	6.02	0.99	8356
7		9.767	0.240	BMB*	1.61	1.21	9164
8		10.434	0.249	BMB	n.a.	1.01	9726

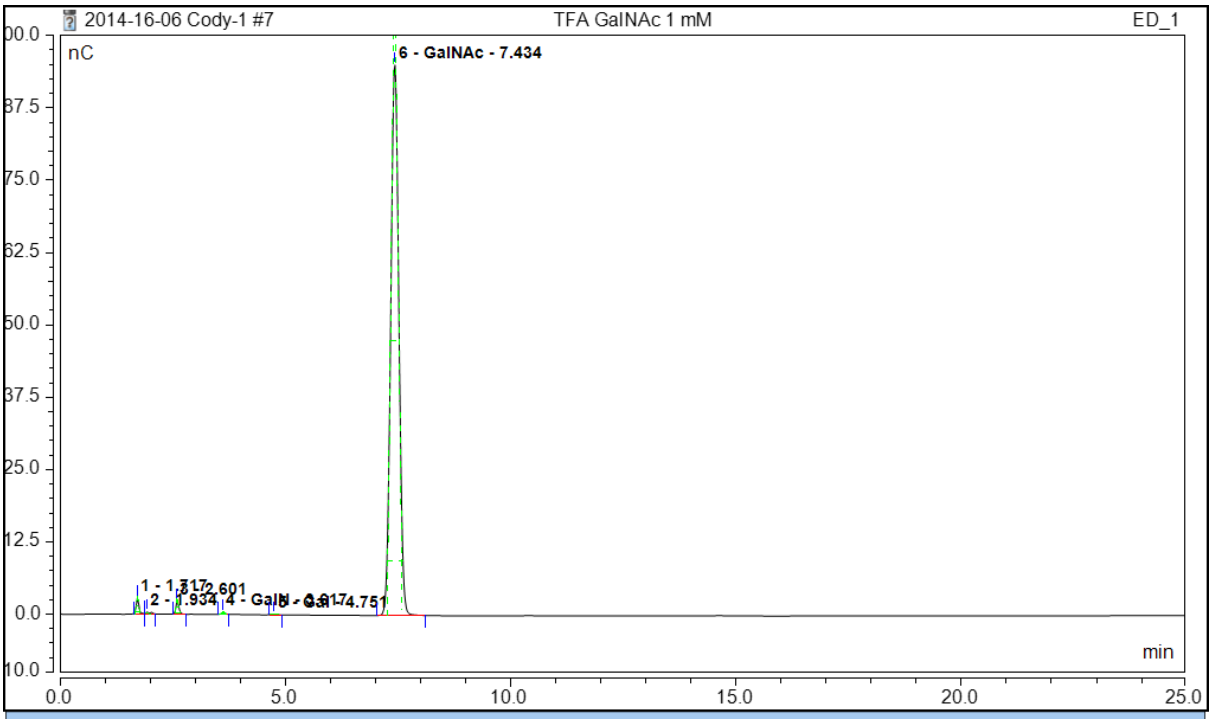
Figure 3.11. HPAEC-PAD supplementary data of *C. jejuni sodB* grown on BHI agar.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.701	0.070	BMb	5.51	1.28	3254
2		1.934	n.a.	Rd	n.a.	n.a.	n.a.
3		2.451	0.090	bMb	1.20	0.82	4072
4		2.617	0.073	bMb	2.66	1.29	7044
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5		2.967	0.082	bMB	17.77	1.10	7290
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	GalNAc	7.451	0.216	BMB	6.85	1.36	6598
7		9.651	0.163	BMB	1.93	1.29	19410
8		10.334	0.256	BMB	40.10	0.99	9056
9		19.951	0.027	BMB	n.a.	0.50	2941649

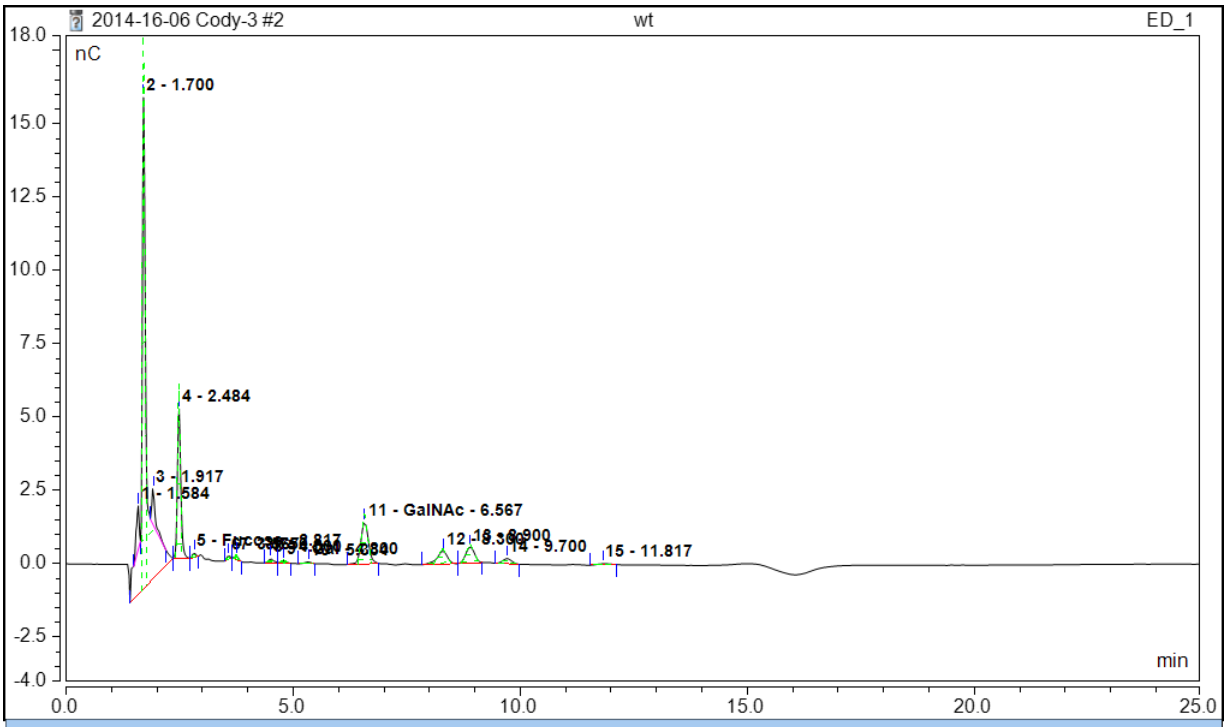
Figure 3.12. HPAEC-PAD supplementary data of *C. jejuni ahpC* grown on BHI agar.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.717	0.065	BMb	2.14	1.17	3844
2		1.934	0.054	bMB	6.30	2.24	7087
3		2.601	0.071	BMB	7.64	1.30	7459
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4	GalN	3.617	0.086	BMB	6.66	1.15	9784
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
5	Gal	4.751	0.115	BMB	10.20	1.18	9484
6	GalNAc	7.434	0.196	BMB	n.a.	1.00	8007

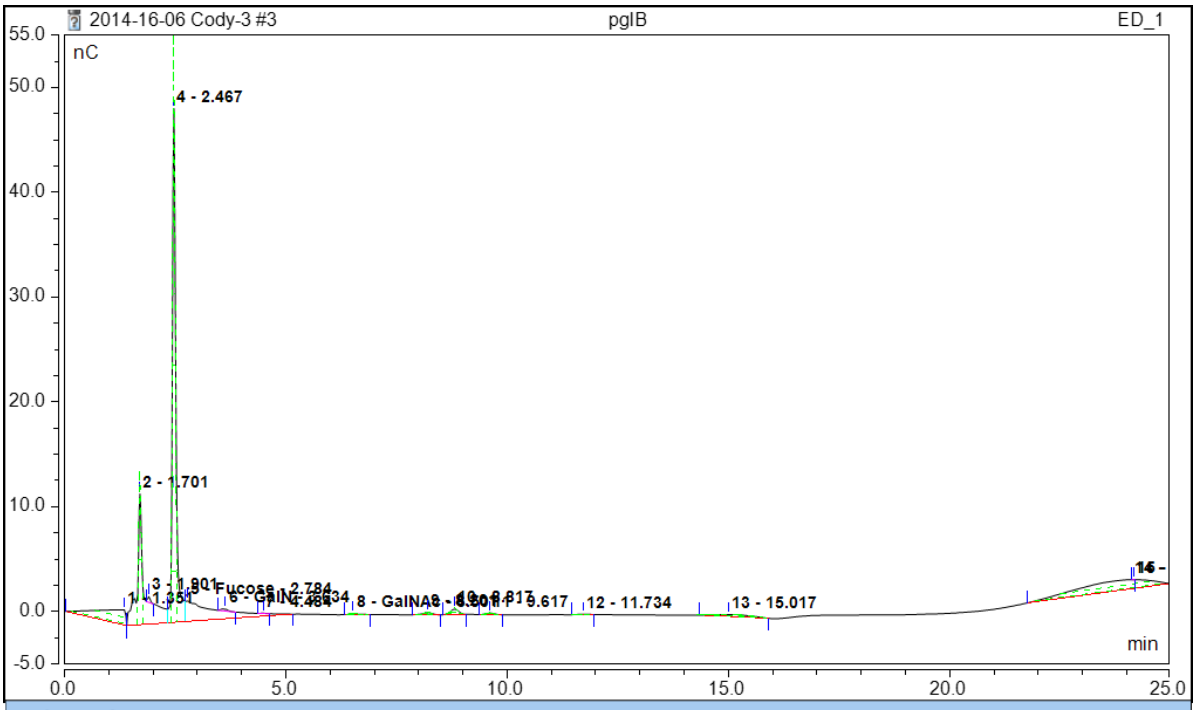
Figure. 3.13. HPAEC-PAD supplementary data of GalNAc for standard.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.584	n.a.	Ru	n.a.	n.a.	n.a.
2		1.700	0.073	BMb	6.15	1.28	3036
3		1.917	n.a.	Rd	n.a.	n.a.	n.a.
4		2.484	0.078	bMB	2.41	1.18	5650
5	Fucose	2.817	0.086	BMB	5.18	1.01	5996
6		3.567	0.085	BMb	1.23	1.03	9708
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7		3.750	0.090	bMB	4.36	0.99	9540
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
8		4.500	0.112	BMb	1.55	1.14	8879
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
9	Gal	4.800	0.116	bMB	2.28	1.00	9533
10		5.334	0.160	BMB	4.31	0.82	6122
11	GalNAc	6.567	0.177	BMB	4.99	1.11	7604
12		8.300	0.233	BMb	1.57	0.80	7036
13		8.900	0.219	bMB	2.12	0.98	9164
14		9.700	0.226	BMB	4.74	1.07	10245
15		11.817	0.302	BMB	n.a.	1.08	8505

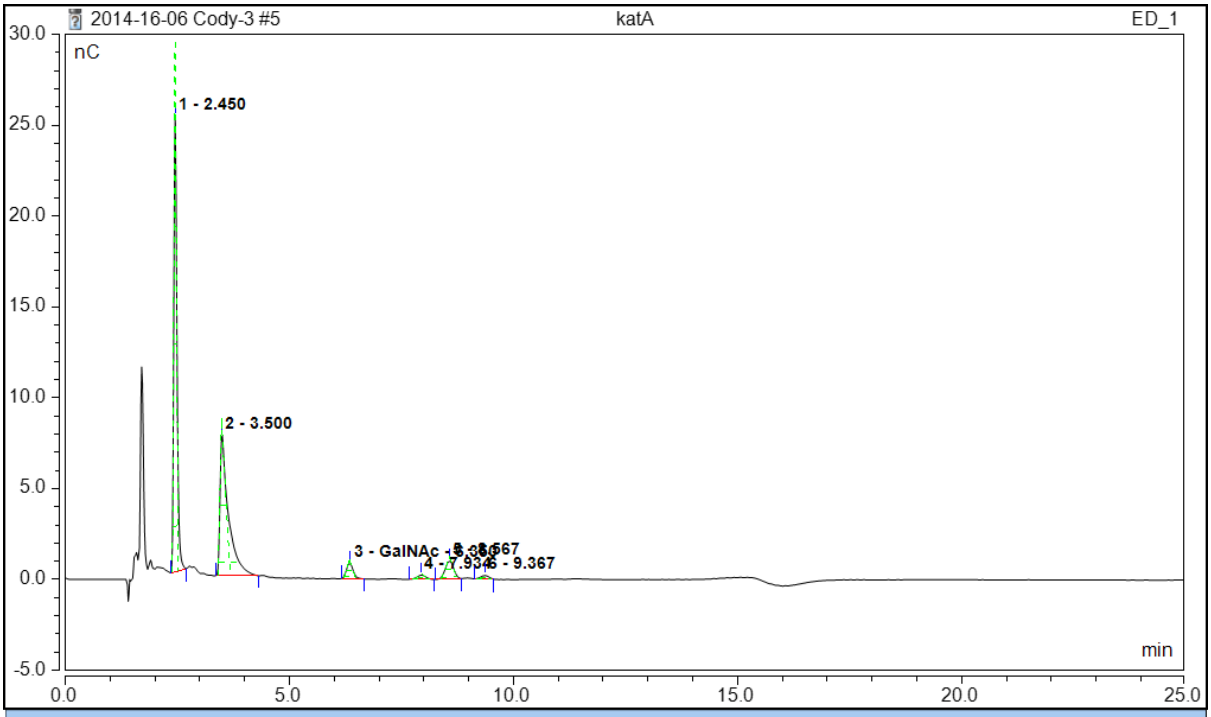
Figure 3.14. HPAEC-PAD supplementary data of *C. jejuni wildtype* grown on BHI broth.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.351	0.718	BMb	0.52	0.52	20
2		1.701	0.080	bM	5.83	n.a.	2524
3		1.901	n.a.	Rd	n.a.	n.a.	n.a.
4		2.467	0.076	M	n.a.	1.56	5907
5	Fucose	2.784	n.a.	MB	n.a.	n.a.	n.a.
6	GalN	3.634	n.a.	Rd	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7		4.484	n.a.	Rd	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
8	GalNAc	6.501	0.186	BMB	4.95	1.69	6776
9		8.201	0.220	BMB	1.68	0.89	7713
10		8.817	0.215	BMB	2.14	0.96	9352
11		9.617	0.226	BMB	5.10	1.01	10041
12		11.734	0.264	BMB	3.14	0.86	10933
13		15.017	0.968	BMB	3.46	1.18	1333
14		24.101	2.129	BMB	n.a.	0.68	710
15		24.167	n.a.	Rd	n.a.	n.a.	n.a.

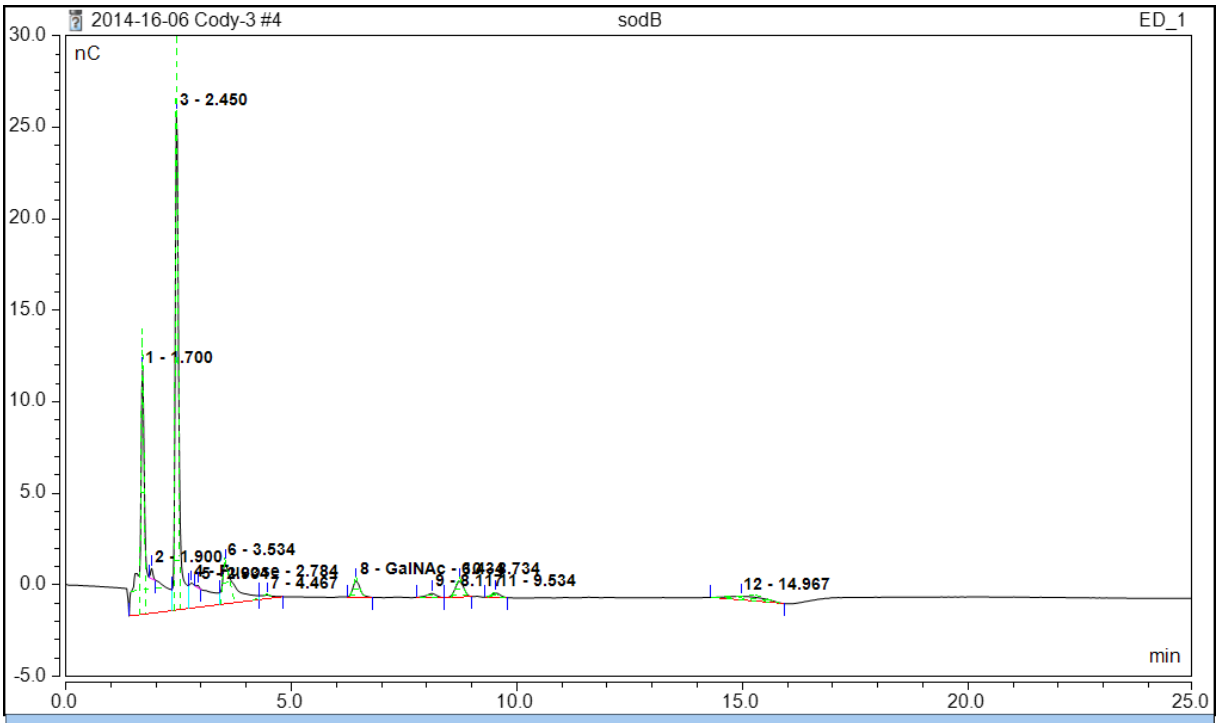
Figure 3.15. HPAEC-PAD supplementary data of *C. jejuni pglB*- grown on BHI broth.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		2.450	0.074	BMB	5.16	1.31	5994
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2		3.500	0.165	BMB	10.23	2.95	2478
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
3	GalNAc	6.350	0.163	BMB	4.95	1.11	8378
4		7.934	0.214	BMB	1.77	1.04	7605
5		8.567	0.209	BMB	2.26	1.00	9287
6		9.367	0.209	BMB	n.a.	0.92	11144

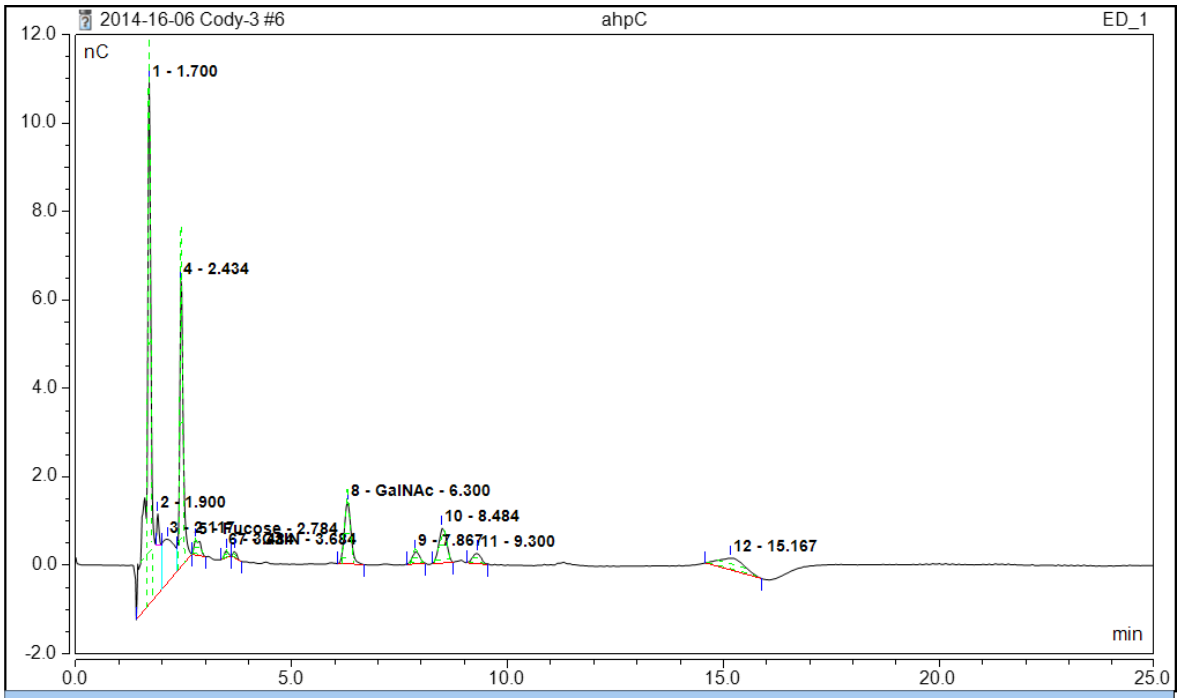
Figure 3.16. HPAEC-PAD supplementary data of *C. jejuni katA*- grown on BHI broth.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.700	0.080	BM	5.63	n.a.	2480
2		1.900	n.a.	Rd	n.a.	n.a.	n.a.
3		2.450	0.077	M	n.a.	1.83	5646
4	Fucose	2.784	n.a.	M	n.a.	n.a.	n.a.
5		2.934	n.a.	Rd	n.a.	n.a.	n.a.
6		3.534	0.212	M	n.a.	n.a.	1532
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
7		4.467	n.a.	MB	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
8	GalNAc	6.434	0.170	BMB	5.03	1.30	7946
9		8.117	0.225	BMb	1.67	0.84	7233
10		8.734	0.212	bMB	2.11	0.94	9413
11		9.534	0.235	BMB	5.08	1.01	9133
12		14.967	1.027	BMB	n.a.	1.21	1176

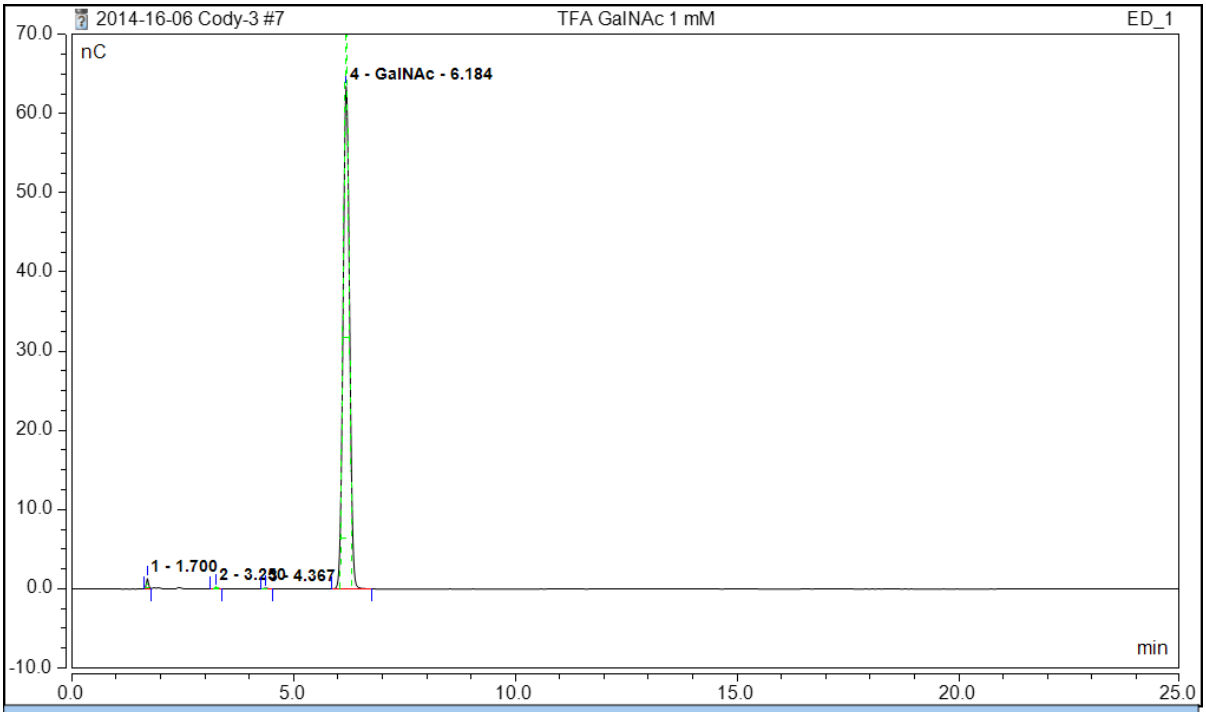
Figure 3. 17. HPAEC-PAD supplementary data of *C. jejuni sodB*- grown on BHI broth.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.700	0.077	BM	n.a.	n.a.	2713
2		1.900	n.a.	Rd	n.a.	n.a.	n.a.
3		2.117	n.a.	M	n.a.	n.a.	n.a.
4		2.434	0.082	Mb	1.61	n.a.	4922
5	Fucose	2.784	0.175	bMB	3.00	1.61	1407
6		3.484	0.101	BMb	1.20	1.15	6594
7	GalN	3.684	0.096	bMB	11.79	1.11	8086
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
8	GalNAc	6.300	0.166	BMB	5.17	1.08	8028
9		7.867	0.192	BMB	1.83	1.08	9269
10		8.484	0.206	BMB	2.28	1.15	9396
11		9.300	0.218	BMB	7.60	0.94	10127
12		15.167	0.693	BMB	n.a.	1.07	2654

Figure 3.18. HPAEC-PAD supplementary data of *C. jejuni ahpC*- grown on BHI broth.



Peak Results

No.	Peak Name	Retention Time min	Width (50%) min	Type	Resolution (EP)	Asymmetry (EP)	Plates (EP)
1		1.700	0.049	BMB	14.78	1.21	6628
n.a.	Fucose	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
2		3.250	0.075	BMB	6.99	1.08	10523
n.a.	GalN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcN	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Ara	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	GlcNAc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
3		4.367	0.114	BMB	7.89	1.32	8151
n.a.	Glc	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
n.a.	Gal	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
4	GalNAc	6.184	0.158	BMB	n.a.	1.05	8517

Figure 3.19. HPAEC-PAD supplementary data of GalNAc standard

3.5 References:

1. Young KT, Davis LM, Dirita VJ. Campylobacter jejuni: Molecular biology and pathogenesis. *Nat Rev Microbiol*. 2007;5(9):665-679. doi: nrmicro1718 [pii].

2. Allos BM. Campylobacter jejuni infections: Update on emerging issues and trends. *Clin Infect Dis*. 2001;32(8):1201-1206. doi: CID001035 [pii].

3. Miller WG, Mandrell RE. Prevalence of campylobacter in the food and water supply: Incidence, outbreaks, isolation and detection. *Campylobacter: Molecular and cellular biology*. 2005:101-163.

4. Luangtongkum T, Jeon B, Han J, Plummer P, Logue CM, Zhang Q. Antibiotic resistance in campylobacter: Emergence, transmission and persistence. *Future Microbiol*. 2009;4(2):189-200. doi: 10.2217/17460913.4.2.189 [doi].

5. Nachamkin I. Chronic effects of campylobacter infection. *Microbes Infect*. 2002;4(4):399-403. doi: S1286457902015538 [pii].

6. Gibreel A, Taylor DE. Macrolide resistance in campylobacter jejuni and campylobacter coli. *J Antimicrob Chemother*. 2006;58(2):243-255. doi: dkl210 [pii].

7. Larsen JC, Szymanski C, Guerry P. N-linked protein glycosylation is required for full competence in campylobacter jejuni 81-176. *J Bacteriol*. 2004;186(19):6508-6514. doi: 10.1128/JB.186.19.6508-6514.2004 [doi].

8. Lin J, Michel LO, Zhang Q. CmeABC functions as a multidrug efflux system in campylobacter jejuni. *Antimicrob Agents Chemother*. 2002;46(7):2124-2131.

9. Wiesner RS, Hendrixson DR, DiRita VJ. Natural transformation of campylobacter jejuni requires components of a type II secretion system. *J Bacteriol*. 2003;185(18):5408-5418.

10. Nothaft H, Szymanski CM. Bacterial protein N-glycosylation: New perspectives and applications. *J Biol Chem*. 2013;288(10):6912-6920.

11. Szymanski CM, Wren BW. Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol*. 2005;3(3):225-237. doi: nrmicro1100 [pii].

12. Scott NE, Parker BL, Connolly AM, et al. 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. *Molecular & Cellular Proteomics*. 2011;10(2):MCP201.

13. Hendrixson DR, DiRita VJ. Identification of campylobacter jejuni genes involved in commensal colonization of the chick gastrointestinal tract. *Mol Microbiol*. 2004;52(2):471-484. doi: 10.1111/j.1365-2958.2004.03988.x [doi].

14. Szymanski CM, Burr DH, Guerry P. Campylobacter protein glycosylation affects host cell interactions. *Infect Immun*. 2002;70(4):2242.

15. Alemka A, Nothaft H, Zheng J, Szymanski CM. N-glycosylation of campylobacter jejuni surface proteins promotes bacterial fitness. *Infect Immun*. 2013;81(5):1674-1682.

16. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the campylobacter genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219. doi: 10.1074/mcp.M112.021519 [doi].

17. Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proc Natl Acad Sci U S A*. 2009;106(35):15019-15024. doi: 10.1073/pnas.0903078106 [doi].

18. Kamel M, Hanafi M, Bassiouni M. Inhibition of elastase enzyme release from human polymorphonuclear leukocytes by N-acetyl-galactosamine and N-acetyl-glucosamine. *Clin Exp Rheumatol*. 1991;9(1):17-21.

19. Kamel M, Alnahdi M. Inhibition of superoxide anion release from human polymorphonuclear leukocytes by N-acetyl-galactosamine and N-acetyl-glucosamine. *Clin Rheumatol*. 1992;11(2):254-260.

20. Needham PG, Brodsky JL. How early studies on secreted and membrane protein quality control gave rise to the ER associated degradation (ERAD) pathway: The early history of ERAD. *Biochim Biophys Acta*. 2013;1833(11):2447-2457. doi: 10.1016/j.bbamcr.2013.03.018 [doi].

21. Aebi M, Bernasconi R, Clerc S, Molinari M. N-glycan structures: Recognition and processing in the ER. *Trends Biochem Sci*. 2010;35(2):74-82. doi: 10.1016/j.tibs.2009.10.001 [doi].

22. Thibault G, Ng DT. The endoplasmic reticulum-associated degradation pathways of budding yeast. *Cold Spring Harb Perspect Biol*. 2012;4(12):10.1101/cshperspect.a013193. doi: 10.1101/cshperspect.a013193 [doi].

23. Suzuki T, Harada Y. Non-lysosomal degradation pathway for N-linked glycans and dolichol-linked oligosaccharides. *Biochem Biophys Res Commun*. 2014;453(2):213-219. doi: 10.1016/j.bbrc.2014.05.075 [doi].

24. Suzuki T. The cytoplasmic peptide:N-glycanase (Ngly1)-basic science encounters a human genetic disorder. *J Biochem*. 2015;157(1):23-34. doi: 10.1093/jb/mvu068 [doi].

25. Hirayama H, Hosomi A, Suzuki T. Physiological and molecular functions of the cytosolic peptide:N-glycanase. *Semin Cell Dev Biol*. 2015;41:110-120. doi: 10.1016/j.semcdb.2014.11.009 [doi].

26. Suzuki T, Funakoshi Y. Free N-linked oligosaccharide chains: Formation and degradation. *Glycoconj J*. 2006;23(5-6):291-302. doi: 10.1007/s10719-006-6975-x [doi].

27. Suzuki T, Park H, Hollingsworth NM, Sternglanz R, Lennarz WJ. PNG1, a yeast gene encoding a highly conserved peptide:N-glycanase. *J Cell Biol*. 2000;149(5):1039-1052.

28. Suzuki T, Kwofie MA, Lennarz WJ. Ngly1, a mouse gene encoding a deglycosylating enzyme implicated in proteasomal degradation: Expression, genomic organization, and chromosomal mapping. *Biochem Biophys Res Commun*. 2003;304(2):326-332. doi: S0006291X03006004 [pii].

29. Need AC, Shashi V, Hitomi Y, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. *J Med Genet*. 2012;49(6):353-361. doi: 10.1136/jmedgenet-2012-100819 [doi].

30. Enns GM, Shashi V, Bainbridge M, et al. Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. *Genet Med*. 2014;16(10):751-758. doi: 10.1038/gim.2014.22 [doi].

31. Suzuki T, Yano K, Sugimoto S, et al. Endo-beta-N-acetylglucosaminidase, an enzyme involved in processing of free oligosaccharides in the cytosol. *Proc Natl Acad Sci U S A*. 2002;99(15):9691-9696. doi: 10.1073/pnas.152333599 [doi].

32. Kato T, Kitamura K, Maeda M, et al. Free oligosaccharides in the cytosol of caenorhabditis elegans are generated through endoplasmic reticulum-golgi trafficking. *J Biol Chem*. 2007;282(30):22080-22088. doi: M700805200 [pii].

33. Caglayan AO, Comu S, Baranoski JF, et al. NGLY1 mutation causes neuromotor impairment, intellectual disability, and neuropathy. *European journal of medical genetics*. 2015;58(1):39-43.

34. Suzuki T, Kwofie MA, Lennarz WJ. Ngly1, a mouse gene encoding a deglycosylating enzyme implicated in proteasomal degradation: Expression, genomic organization, and chromosomal mapping. *Biochem Biophys Res Commun*. 2003;304(2):326-332.

35. Imlay JA. Pathways of oxidative damage. *Annual Reviews in Microbiology*. 2003;57(1):395-418.

36. Imlay JA. Cellular defenses against superoxide and hydrogen peroxide. *Annu.Rev.Biochem.* 2008;77:755-776.

37. van Vliet AH, Ketley JM, Park SF, Penn CW. The role of iron in campylobacter gene regulation, metabolism and oxidative stress defense. *FEMS Microbiol Rev*. 2002;26(2):173-186.

38. Baillon MA, Van Vliet AH, Ketley JM, Constantinidou C, Penn CW. An iron-regulated alkyl hydroperoxide reductase (AhpC) confers aerotolerance and oxidative stress resistance to the microaerophilic pathogen campylobacter jejuni. *J Bacteriol*. 1999;181(16):4798-4804.

39. Purdy D, Cawthraw S, Dickinson JH, Newell DG, Park SF. Generation of a superoxide dismutase (SOD)-deficient mutant of campylobacter coli: Evidence for the significance of SOD in campylobacter survival and colonization. *Appl Environ Microbiol*. 1999;65(6):2540-2546.

40. Palyada K, Sun Y, Flint A, Butcher J, Naikare H, Stintzi A. Characterization of the oxidative stress stimulon and PerR regulon of campylobacter jejuni. *BMC Genomics*. 2009;10(1):481.

41. Parkhill J, Wren BW, Mungall K, et al. The genome sequence of the food-borne pathogen campylobacter jejuni reveals hypervariable sequences. *Nature*. 2000;403(6770):665-668. doi: 10.1038/35001088 [doi].

42. Tanner ACR, Haffer C, Bratthall GT, Visconti RA, Socransky SS. A study of the bacteria associated with advancing periodontitis in man. *J Clin Periodontol*. 1979;6(5):278-307.

43. Etoh Y, Dewhirst FE, Paster BJ, Yamamoto A, Goto N. Campylobacter showae sp. nov., isolated from the human oral cavity. *Int J Syst Bacteriol*. 1993;43(4):631-639. doi: 10.1099/00207713-43-4-631 [doi].

44. Breedveld MW, Benesi AJ, Marco ML, Miller KJ. Effect of phosphate limitation on synthesis of periplasmic cyclic (beta)-(1, 2)-glucans. *Appl Environ Microbiol*. 1995;61(3):1045-1053.

45. Nothaft H, Liu X, McNally DJ, Szymanski CM. N-linked protein glycosylation in a bacterial system. *Methods Mol Biol*. 2010;600:227-243. doi: 10.1007/978-1-60761-454-8_16 [doi].

46. Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proc Natl Acad Sci U S A*. 2009;106(35):15019-15024. doi: 10.1073/pnas.0903078106 [doi].

47. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the campylobacter genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219. doi: 10.1074/mcp.M112.021519 [doi].

48. Liu X, McNally DJ, Nothaft H, Szymanski CM, Brisson J, Li J. Mass spectrometry-based glycomics strategy for exploring N-linked glycosylation in eukaryotes and bacteria. *Anal Chem*. 2006;78(17):6081-6087.

49. Butler EK, Davis RM, Bari V, Nicholson PA, Ruiz N. Structure-function analysis of MurJ reveals a solvent-exposed cavity containing residues essential for peptidoglycan biogenesis in escherichia coli. *J Bacteriol*. 2013;195(20):4639-4649.

Chapter IV

Conclusions and future directions

4.1 Research purpose

My research project involved characterizing the generic serine-protease inhibitor known as ecotin from two oral *Campylobacter* species (*C. rectus* and *C.* showae). The objectives of this research were to identify if these ecotin homologues are function protease inhibitors and if they can inhibit a range of serine-proteases. One approach is to test if the ecotin homologues have an effect *in vivo* in promoting bacterial fitness in the oral cavity by protecting against serine proteases produced by the host organism. The second research project entails investigating and enhancing our current understanding of the underlying mechanisms of N-glycosylation and fOS turnover and recycling in the pathogen *C. jejuni*. Understanding how N-glycans are synthesized and recycled can allow for increased knowledge that can be used for improving glycoprotein engineering.

4.2 Summary and future directions

At the start of this thesis project, all the *Campylobacter* species were shown to contain the *pgl* loci and capable of N-glycosylation of proteins and production of fOS. It was noticed that when identifying the *pgl* loci of oral *Campylobacter* species that a homologue of the *E. coli*

serine-protease inhibitor ecotin was found upstream of the N-glycosylation machinery. It has been recently published in the Szymanski laboratory that N-glycosylation promotes bacterial fitness in the chicken cecum, protecting the bacteria from gut serine-proteases, but it is unknown whether N-glycosylation is sufficient for protection of proteolysis in the oral cavity. [1].

The first objective of my Master's project was to identify if these ecotin homologues are functional protease inhibitors (see chapter 2). I first aligned all the *Campylobacter* species ecotin homologues against the *E. coli* ecotin. As there was low sequence homology between the ecotin homologues, I used the protein homology/analogy recognition engine (Phyre2) to develop a theoretical tertiary structure, there was high structural homology between the homologues. I first only wanted to test two of the ecotin homologues before testing others. The first homologue I selected was *C. rectus* ecotin, because this is the known periodontal pathogen of *Campylobacter* species [2,3] and is found in the periodontal pockets that are heavily attacked by neutrophils. The *C. rectus* ecotin was also of interest because it contains a leucine at the P1 active site where *E. coli* ecotin contains a methionine. This is of interest because the methionine at the P1 active site is one of the defining features of the serine-protease inhibitor ecotin [4]. The second ecotin homologue that I selected was *C. showae,* a more recently discovered oral *Campylobacter* species and its full role in periodontal disease is unknown [5]. *C. showae* ecotin contains a methionine at the P1 active site like the *E. coli* ecotin. I first cloned multiple ecotin homologues into the expression vector pET22B for overexpression in *E. coli*. I removed their natural signal sequence for transport to the periplasmic space and replaced it with a *pelB* sequence that would be recognized by the *E. coli* machinery. After purification and a western-blot, I noticed that *C. rectus* and *C. showae* ecotins purified with the highest yield. An *in vitro* protease protection assay was first completed using the serine-protease trypsin. I showed that the *C. rectus* ecotin

and *C. showae* ecotin were effective at inhibiting trypsin. One interesting feature about *E. coli* ecotin is that is a generic serine protease inhibitor and it can inhibit a wide range or serine-proteases. I then modified a FRET assay that we developed to show protection of proteins from factor Xa by addition of N-glycans, to be used for ecotin protection. The FRET assay showed that the ecotin homologues were able to inhibit factor Xa, but the *C. rectus* Ecotin was not as effective as the *E. coli* or *C. showae.* The most important serine protease was tested next, neutrophil elastase which is one of the most prominent proteases in the periodontal disease because of the innate immune system response to bacterial infections. The ecotin homologues showed high affinity for binding and inhibiting neutrophil elastase. An interesting aspect of ecotins is that they contain a signal peptide for their transport into the periplasm, but it is unknown if they are secreted out of the cells. To understand if the ecotin homologues are for an internal process or for exogenous protection, they were tested against self-serine proteases DegP from *E. coli* and HtrA from *C. jejuni.* HtrA and DegP were not inhibited by the ecotin homologues; they still showed proteolytic activity in the presence of ecotins. This bolsters the idea that ecotins are purely for protection against exogenous serine-protease that the bacteria would encounter.

Multiple attempts have been made at the construction of an ecotin mutant in *C. rectus* and *C. showae* but were never successful. An alternative approach I used was to construct an ecotin mutant in *E. coli* and complement this strain with the ecotin homologues from the *Campylobacter* species. The construction of the *E. coli* ecotin mutant and the complementation of ecotin homologues was successful. In collaboration with Dr. Balaz Rada (University of Georgia), I tested survival of the *E. coli* constructs against whole cell neutrophil killing. We found that the *E. coli* ecotin mutant had a major decrease in bacterial fitness and was killed

almost immediately in contrast to wildtype *E. coli*. Complementation of the ecotin homologues restored bacterial fitness back to wildtype levels. The complemented strains were then tested against purified NETs of neutrophils showing comparable results to the survival by whole cell neutrophils, therefore all ecotins tested were active against proteases embedded in neutrophil NETs. *C. rectus* and *C. showae* ecotin proteins are generic serine-protease inhibitors with the ability to inhibit trypsin, factor Xa, neutrophil elastase and increase bacterial fitness against neutrophil cells.

Future work is still required to create mutants in these *Campylobacter* species to fully understand the complete role and effects ecotin plays in these bacteria and their roles in periodontitis. Crystallization of the ecotin homologues could be very useful to ascertain in whether the structure is similar to the *E. coli* ecotin. Crystallization could also show the effects of the methionine or leucine in the P1 active site of the ecotin homologues. Once fully characterized, the protease inhibitor could potentially be engineered as a therapeutic for overzealous immune systems that damage "self-cells" with constant attack of serine-proteases. We are currently testing if addition of ecotin influences *C. jejuni* colonization levels in an avian model. I think it is important to understand if ecotins are effective inhibitors in cystic fibrosis models. However, it seems that they may not be able to diffuse through the thick mucus [6]. This can be tested quite easily by using cystic fibrosis patient mucus samples and incubating the samples with ecotin homologues followed by testing the amount of active neutrophil elastase in the samples. Once initial efficacy is shown then therapeutic trials of ecotin can be tested in cystic fibrosis murine models [7], to determine if ecotin can reduce host-tissue damage.

The second objective of my research project was to study the degradation and recycling of products of the N-linked protein glycosylation pathway, N-glycans and free oligosaccharides

(fOS) in *C. jejuni* (see Chapter 3). Development of methods to label and follow N-glycans and fOS are important tools necessary to examine this process. The problem with labelling N-glycans in *C. jejuni* is that the organism has a limited metabolism for consumption of carbohydrates [8,9], therefore it is not possible to simply feed labelled sugars and follow the precursors throughout the system. While studying a *C. jejuni pglB* oligosaccharyltransferase mutant that cannot produce fOS, we noticed a unique spot by thin-layer chromatography (TLC). Analysis of this TLC spot by mass spectrometry suggested that it is potentially a product of the N-glycosylation machinery producing a tetrasaccharide of 3 GalNAc and 1 glucose residues. It is possible that the *pglB* mutant needs to recycle its undecaprenyl phosphate (Und-P) to allow building and restructuring of peptidoglycan. If PglB is not active, all the Und-P could be used, and the glycan would need to be removed to allow survival. There may be a mechanism that exists in wildtype cells that allows the *pglB* mutant to remove the glycan from Und-P. either one or both of the unidentified glycanases could remove the GalNAc and glucose residues signaling the rest of the glycan to be recycled. Another unidentified glycanase could hydrolyse between the GalNAc and GalNAc releasing the tetrasaccharide of GalNAcs and glucose identified by MS. But, it is difficult to imagine that so many unidentified enzymes, including one for release of the diNAcBac from the Und-PP would exist. Future work is required to understand this process. Isotopic detection of amino sugars with glutamine (I-DAWG) [10] and stable isotope labeling with amino acids in cell culture (SILAC) [11] are methods that can be used to study glycan recycling. I-DAWG includes the feeding of heavy labelled glutamine would result in labelling of amino sugars on N-glycans and fOS. SILAC labeling includes the feeding of heavy labelled lysine into a *C. jejuni* auxotrophic mutant that requires lysine. With completion of a pulse-chase experiment, we would be able to determine if *C. jejuni* does turn over and recycle glycans or if the sugars are not broken down

and are simply diluted with cell division. If N-glycans are recycled, there will be a decrease in the amount of labelled glycans once the heavy labelled amino acids are removed, but if there is no decrease in labelled N-glycans and fOS after switching to non-labelled amino acids this alludes to the idea that *C. jejuni* does not have a method for recycling and turnover of N-glycans and fOS.

The third objective of my research project was to characterize the DGGK sequence of *Campylobacter* PglB enzymes (see Appendix I)[12]. We analyzed multiple sequence alignments of *Campylobacter* PglB enzymes identifying a second conserved motif, DGGK, located 12 amino acids away from the catalytic WWDYG motif [13]. The DGGK motif is conserved in among all *Campylobacter* species. *Campylobacter gracilis* and *Campylobacter curvus* each have a second inactive PglB enzyme, missing the DGGK motif. I added the DGGK motif into the PglB homologues missing the sequence. However, N-glycosylation was not restored when examined by western blotting suggesting that other changes to the protein sequence in these two PglB enzymes are also important for activity. I then wanted to test if altering the DGGK sequence on the *C. curvus* active PglB resulted in loss of function of the OTase. I switched the DGGK motif with the inactive PglB, putting in a DPGR motif into the active *C. curvus* PglB. This did not cause complete loss of function, but activity was heavily reduced indicating the importance of the DGGK motif for the PglB enzyme. Dr. Harald Nothaft completed the limited proteolysis experiments to demonstrate that the DGGK motif plays a role in binding the donor glycan LLO [14]. The structure of the yeast STT3 OTase bound to LLO [15] shows similar grooves where the dolichol tail of the LLO most likely bind. This suggests that the PglB and STT3 OTase share a common substrate recognition even though having different substrate specificities. Aebi and

Locher (2017) also completed the first x-ray crystallography of a PglB in complex with a LLO and acceptor peptide, showing that the DGGK motif interacts with the LLO holding it in place

4.3 Conclusions and perspectives

Previous work in our laboratory demonstrated that N-linked protein glycosylation enhances *C. jejuni* growth in the presence of gastrointestinal proteases [1]. Analyses of the N-glycan pathways of all known *Campylobacter* species identified ecotin homologues within the protein glycosylation loci of oral *Campylobacters* [16]. The key objective of this thesis was to characterize the serine-protease inhibitors (ecotins) of two oral *Campylobacter* species: *C. rectus* and *C. showae*. The second objective of my study was to understand how glycoproteins and the fOS generated through the N-linked protein glycosylation pathways are broken down and what role PglB may play in this process.

Analysis of the ecotin homologues can shed light on the bacterial fitness and survival in the human host. Ecotins could potentially be antimicrobial targets to reduce fitness of certain bacterial strains that are able to evade our immune systems while resulting in damage to host tissues. Ecotins should be further studied and potentially be used as therapeutics to control the immune system in patients where it is overstimulated and damaging to self-tissue. Proteases are important in our vital processes such as digestion [17], blood coagulation [18], apoptosis [19], immunity [20] and tissue remodeling [21]. Homeostasis disruption and dysregulation of protease activity can lead to numerous pathologic disorders such as cardiovascular and inflammatory diseases [22], cancer [23] and neurological disorders [24]. Thus, targeting proteases is a promising therapeutic approach for these disorders. Analysis of the OTase activity as well as the N-glycan

structures in other organisms can help shed light on the protection and importance of these carbohydrates in the bacterial world. The disruption of the N-glycosylation pathway in *C. jejuni* does not affect its viability, but disturbs many biological functions [25]. The conserved N-glycosylation pathway in all *Campylobacter* species supports a key role in the physiology of these bacteria [26].

4.4 References:

1. Alemka A, Nothaft H, Zheng J, Szymanski CM. N-glycosylation of *campylobacter jejuni* surface proteins promotes bacterial fitness. *Infect Immun*. 2013;81(5):1674-1682.

2. Henne K, Fuchs F, Kruth S, Horz H, Conrads G. Shifts in *campylobacter* species abundance may reflect general microbial community shifts in periodontitis progression. *Journal of oral microbiology*. 2014;6(1):25874.

3. Moore W, Moore LV. The bacteria of periodontal diseases. *Periodontol 2000*. 1994;5(1):66-77.

4. McGrath ME, Hines WM, Sakanari JA, Fletterick RJ, Craik CS. The sequence and reactive site of ecotin. A general inhibitor of pancreatic serine proteases from *escherichia coli. J Biol Chem*. 1991;266(10):6620-6625.

5. Etoh Y, Dewhirst FE, Paster BJ, Yamamoto A, Goto N. *Campylobacter showae* sp. nov., isolated from the human oral cavity. *Int J Syst Bacteriol*. 1993;43(4):631-639.

6. Quinton PM. Cystic fibrosis: Impaired bicarbonate secretion and mucoviscidosis. *The Lancet*. 2008;372(9636):415-417.

7. Rosen BH, Chanson M, Gawenis LR, et al. Animal and model systems for studying cystic fibrosis. *Journal of Cystic Fibrosis*. 2018;17(2):S34.

8. Parkhill J, Wren BW, Mungall K, et al. The genome sequence of the food-borne pathogen *campylobacter jejuni* reveals hypervariable sequences. *Nature*. 2000;403(6770):665-668.

9. Velayudhan J, Kelly DJ. Analysis of gluconeogenic and anaplerotic enzymes in *campylobacter jejuni*: An essential role for phosphoenolpyruvate carboxykinase. *Microbiology*. 2002;148(3):685-694.

10. Orlando R, Lim J, Atwood III JA, et al. IDAWG: Metabolic incorporation of stable isotope labels for quantitative glycomics of cultured cells. *Journal of proteome research*. 2009;8(8):3816-3823.

11. Sun S, Shah P, Eshghi ST, et al. Comprehensive analysis of protein glycosylation by solid-phase extraction of N-linked glycans and glycosite-containing peptides. *Nat Biotechnol*. 2016;34(1):84.

12. Barre Y, Nothaft H, Thomas C, et al. A conserved DGGK motif is essential for the function of the PglB oligosaccharyltransferase from *campylobacter jejuni*. *Glycobiology*. 2017;27(10):978-989.

13. Wacker M, Linton D, Hitchen PG, et al. N-linked glycosylation in *campylobacter jejuni* and its functional transfer into *E. coli*. *Science*. 2002;298(5599):1790-1793.

14. Barre Y, Nothaft H, Thomas C, et al. A conserved DGGK motif is essential for the function of the PglB oligosaccharyltransferase from *campylobacter jejuni*. *Glycobiology*. 2017;27(10):978-989.

15. Wild R, Kowal J, Eyring J, Ngwa EM, Aebi M, Locher KP. Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science*. 2018;359(6375):545-550.

16. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *campylobacter* genus. *Molecular & Cellular Proteomics*. 2012;11(11):1203-1219.

17. Goodman BE. Insights into digestion and absorption of major nutrients in humans. *Adv Physiol Educ*. 2010;34(2):44-53.

18. Walsh DM, Klyubin I, Fadeeva JV, Rowan MJ, Selkoe DJ. Amyloid-β oligomers: Their production, toxicity and therapeutic inhibition. *Amyloid-β oligomers: their production, toxicity and therapeutic inhibition*. 2002.

19. Tummers B, Green DR. Caspase-8: Regulating life and death. *Immunol Rev*. 2017;277(1):76-89.

20. Heutinck KM, ten Berge IJ, Hack CE, Hamann J, Rowshani AT. Serine proteases of the human immune system in health and disease. *Mol Immunol*. 2010;47(11-12):1943-1955.

21. Shrock E, Güell M. CRISPR in animals and animal models. In: *Progress in molecular biology and translational science.* Vol 152. Elsevier; 2017:95-114.

22. Puddu P, Puddu GM, Cravero E, Muscari S, Muscari A. The involvement of circulating microparticles in inflammation, coagulation and cardiovascular diseases. *Can J Cardiol*. 2010;26(4):140-147.

23. Koblinski JE, Ahram M, Sloane BF. Unraveling the role of proteases in cancer. *Clinica chimica acta*. 2000;291(2):113-135.

24. Holsinger RD, McLean CA, Beyreuther K, Masters CL, Evin G. Increased expression of the amyloid precursor β-secretase in alzheimer's disease. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*. 2002;51(6):783-786.

25. Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proceedings of the National Academy of Sciences*. 2009;106(35):15019-15024.

26. Nothaft H, Szymanski CM. Protein glycosylation in bacteria: Sweeter than ever. *Nature Reviews Microbiology*. 2010;8(11):765.

Appendix I

A CONSERVED DGGK MOTIF IS ESSENTIAL FOR THE FUNCTION OF THE PGLB OLIGOSACCHARYLTRANSFERASE FROM *CAMPYLOBACTER JEJUNI* [1]

Barre, Y., Nothaft, H., Thomas, C., Liu, X., Li, J., Ng, KKS, and Szymanski CM. 2017. *Glycobiology*. (27) 10:978-989.

Abstract

In *Campylobacter jejuni*, the PglB oligosaccharyltransferase catalyzes the transfer of a heptasaccharide from a lipid donor to asparagine within the D/E-X1-N–X2-S/T sequon (X1,2 ≠ P) or releases this heptasaccharide as free oligosaccharides (fOS). Using available crystal structures and sequence alignments, we identified a DGGK motif near the active site of PglB that is conserved among all *Campylobacter* species. We demonstrate that amino acid substitutions in the aspartate and lysine residues result in loss of protein glycosylation in the heterologous *Escherichia coli* system. Similarly, complementation of a *C. jejuni pglB* knock-out strain with mutated *pglB* alleles results in reduced levels of N-linked glycoproteins and fOS in the native host. Analysis of the PglB crystal structures from *Campylobacter lari* and the soluble C-terminal domain from *C. jejuni* suggests a particularly important structural role for the aspartate residue and the two following glycine residues, as well as a more subtle, less defined role for the lysine residue. Limited proteolysis experiments indicate that conformational changes of wildtype PglB that are induced by the binding of the lipid-linked oligosaccharide are altered by changes in the DGGK motif. Related to these findings, certain *Campylobacter* species possess two PglB orthologues and we demonstrate that only the orthologue containing the DGGK motif is active. Combining the knowledge gained from the PglB structures and mutagenesis studies, we propose a function for the DGGK motif in affecting the binding of the undecaprenylpyrophosphate glycan donor substrate that subsequently influences N-glycan and fOS production.

5.1 Introduction

Asparagine-(N)-linked protein glycosylation is one of the most prevalent protein modifications in eukaryotes. It involves transfer of sugars from nucleotide-activated sugar donors onto the dolicholpyrophosphate lipid at the cytoplasmic side of the endoplasmic reticulum (ER) [1]. The partially assembled glycan is flipped to the lumenal side of the ER membrane by an adenosine triphosphate (ATP)-independent flippase where additional modification of the glycan occurs before it is transferred from the lipid carrier to asparagine residues in the N–X–S/T (X ≠ P) sequon of protein acceptors by the oligosaccharyltransferase (OTase) complex [2,3].

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 [4]. In lower eukaryotes, such as *Giardia* and kinetoplastids, the OTase is composed of a single polypeptide membrane protein that consists of *STT3* alone [5]. The genomes of *Trypanosoma brucei* and *Leishmania major* reveal the presence of three and four *STT3* paralogous genes, respectively [6,7]. Homologous processes of N-linked protein glycosylation have been found outside of the Eukaryotic domain, in both Archaea and Bacteria [8,9]. 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[10].

The bacterial process of N-glycosylation shares many features with its eukaryotic and archaeal counterparts, but significant differences exist. In *Campylobacter jejuni*, N-glycosylation is coupled to general secretion-mediated translocation to the periplasm [11]. In this process, nucleotide-activated sugars are synthesized in the cytoplasm and assembled onto an undecaprenylpyrophosphate-linked carrier forming a heptasaccharide at the cytoplasmic side of the inner membrane, where they are then translocated to the periplasm [12,13] 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 acids except proline [14]. The central enzyme responsible for the transfer reaction is PglB, the OTase encoded by the protein glycosylation (*pgl*) locus [9]. The *C. jejuni pgl* gene cluster contains all of the genes necessary for the N-glycosylation process, and can be functionally reconstituted into *Escherichia coli* to produce recombinant glycoproteins [15]. In addition to N-glycosylation, PglB also releases free oligosaccharides (fOS) into the periplasm in response to changes in the osmotic environment and bacterial growth phase [16].

The N-glycosylation pathway has been identified in an increasing number of bacteria [17,18]. 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 [18], the delta-proteobacterium *Desulfovibrio desulfuricans* [19], and in the epsilon-proteobacteria *Wolinella succinogenes* [20], *Sulfurovum* sp., *Nitratiruptor* sp. [21] and certain *Helicobacter* species [22]. The N-glycosylation pathway has not yet been demonstrated to be functional in all of these organisms, but they each possess the key enzyme, PglB [23,24]. 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 [25]. It comprises 13 predicted hydrophobic N-terminal transmembrane domains spanning the periplasmic membrane and a soluble C-terminal domain oriented toward the periplasmic space [26].

All OTase orthologues in all three domains of life have the signature WWDXG motif (where X is Y or W) shown to be necessary for *Cj*-PglB activity [15]. The crystal structure of the full length

PglB of *Campylobacter lari* (*Cl*-PglB) has been solved and has 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 [27]. A crystal structure of the soluble C-terminal domain was previously determined for the closely related homolog *Cj*-PglB, and structural comparisons with
the archaeal AglB revealed a conserved Ile 571 within an MXXI motif [28]. Based on the proximity of the isoleucine residue to the bound peptide or protein acceptor substrate, the isoleucine residue in homologous proteins has been proposed to provide contact with the +2 Thr to aid in protein binding and/or recognition [27].

Here, we identified and analyzed amino acids that are important for the function and/or structure of *Cj*-PglB. Analyses of available crystal structures and sequence alignments identified a DGGK sequence located in close proximity to the WWDXG motif. The DGGK motif is highly conserved in all campylobacter PglB proteins and is present in at least one PglB orthologue for species that contain two PglB enzymes. We show that specific PglB mutations in D and K influence N-glycosylation and fOS hydrolytic activity and that predicted conformational changes induced by the binding of the lipid-linked oligosaccharide (LLO) are altered by introducing amino acid exchanges in the DGGK motif. We further demonstrate that orthologues containing a DGGK motif are functional in N-glycan transfer but that the presence of DGGK alone is not sufficient to establish PglB activity in orthologues that are lacking the motif.

5.2 Results

5.2.1 Functional analysis of *C. jejuni* DGGK points mutants in *E. coli*

The crystal structures of *C. lari* PglB and the soluble C-terminal domain of *C. jejuni* PglB, together with a homology model of the full-length *C. jejuni* PglB reveal that the ^{475}DGGK^{478} motif is located directly next to the conserved ^{457}WWDXG^{461} sequon (Figure 1) [15]. In silico analysis further demonstrated that the DGGK motif is conserved in all *Campylobacter* species (Figure 2). To analyze if these amino acids are important for PglB function, we substituted the D and K residues in ^{475}DGGK^{478} to generate the following variants: D475A, D475V, D475K, K478A, K478V, ^{475}AGGA^{478}, ^{475}VGGV^{478} and ^{475}KGGD^{478}. Western blotting with HA-tag specific antibodies showed that the introduction of the point mutations had no effect on PglB expression levels when compared to the wildtype PglB protein (Figure 3A). Since whole cell lysates of *E. coli* expressing PglB-HA proteins were examined, the HA signal in the higher molecular weight range that is present in all preparations, including the vector control, is due to unspecific binding to an *E. coli* protein. To analyze PglB activities, we expressed the various *pglB* alleles in *E. coli* CLM24 in the presence of pACYC184(*pglmut*) and pWA2 encoding the soluble form of the His6-tagged *C. jejuni*-N-glycosylation acceptor protein *Cj*-CmeA. Wildtype PglB (expressed from pMAF10) or inactive PglB (WWDYG to WAAYG, expressed from pWA1) served as positive and negative controls, respectively. PglB activities were determined by analyzing the glycosylation status of purified *Cj*-CmeA-His6 by western blotting using His6-Tag-specific and *C. jejuni*-N-glycan-specific antibodies (Figure 3A) followed by densitometric analysis of the *Cj*-CmeA-specific signals of the anti-

His6 western blot (Figure 3B). Loss of *Cj*-CmeA glycosylation was observed with the double mutations AGGA, VGGV, KGGD and the single amino acid substitutions D475K and D475V. The single point mutations D475A, K478V and K478A resulted in reduced *Cj*-CmeAHis6 glycosylation when compared to the glycosylation pattern observed in the presence of wildtype PglB. This series of mutants shows that substitutions of D475 are less tolerated than K478.

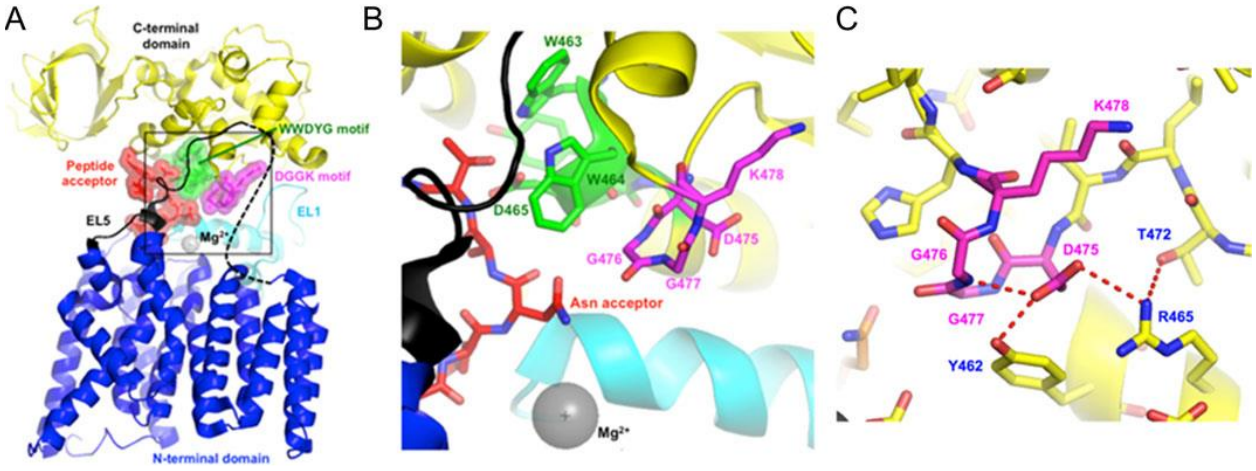


Fig. A.1. Location of DGGK motif within the crystal structure of the *C. lari* PglB:acceptor peptide complex (PDB 3RCE). (A) Ribbon diagram of PglB, showing the N-terminal transmembrane domain (blue), C-terminal soluble domain (yellow), peptide substrate acceptor (red), WWDYG motif (green), DGGK motif (magenta) and catalytic divalent cation (modeled as magnesium ion, gray). Extended loops EL1 (cyan) and EL5 (black) are also highlighted; the disordered N-terminal portion of EL5 missing from the crystal structure is denoted by a dashed line. (B) Close-up view of the region denoted as a black box in panel (A) showing the atomic positions of

the atoms in the peptide acceptor substrate (red), WWDYG motif (green), DGGK motif (magenta) and divalent cation (modeled as magnesium ion, gray). (C) Hydrogen-bonding interactions (dashed red lines) involving the side chain carboxylate group of Asp^{475} are shown,

indicating some of the structural constraints imposed by residues surrounding the DGGK motif. The numbers for the residues in the DGGK motif are given for *C. jejuni*, even though the structure that is shown is from *C. lari*. Note that the three-dimensional structure of PglB from *C. jejuni* is predicted to be very similar to that from *C. lari* due to the high level of sequence identity between the two homologous proteins. This figure is available in black and white in print and in color at *Glycobiology* online.

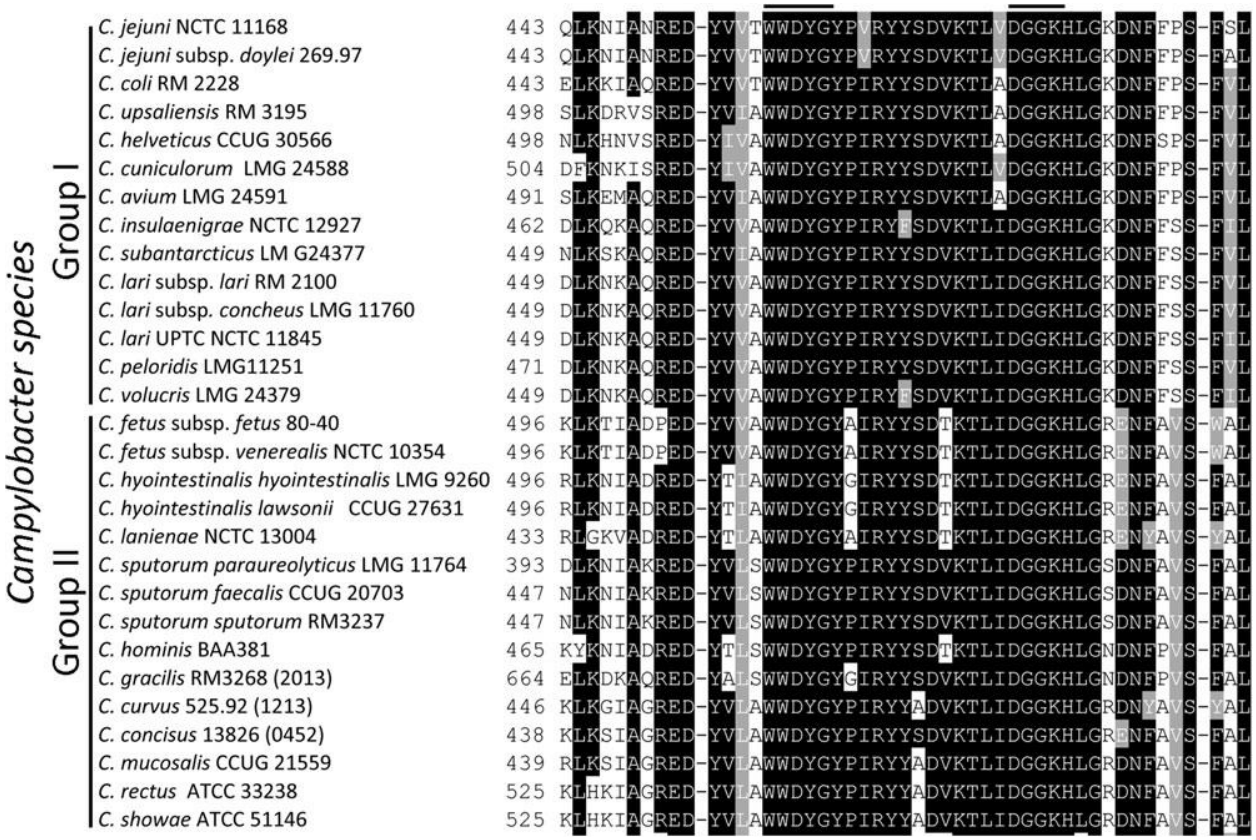


Fig. A.2. Sequence alignment of *Campylobacter* PglB protein sequences. The ^{475}DGGK^{478} sequence in *C. jejuni* NCTC 11168 is located close to the highly conserved ^{457}WWDYG^{461} motif indicated by black bars above the sequence that are conserved in all *Campylobacter* species of the thermophilic (I) and nonthermophilic (II) group. Highlighted in black are amino acids that are conserved in at least 50% of the sequences, gray highlighting indicates the presence of an

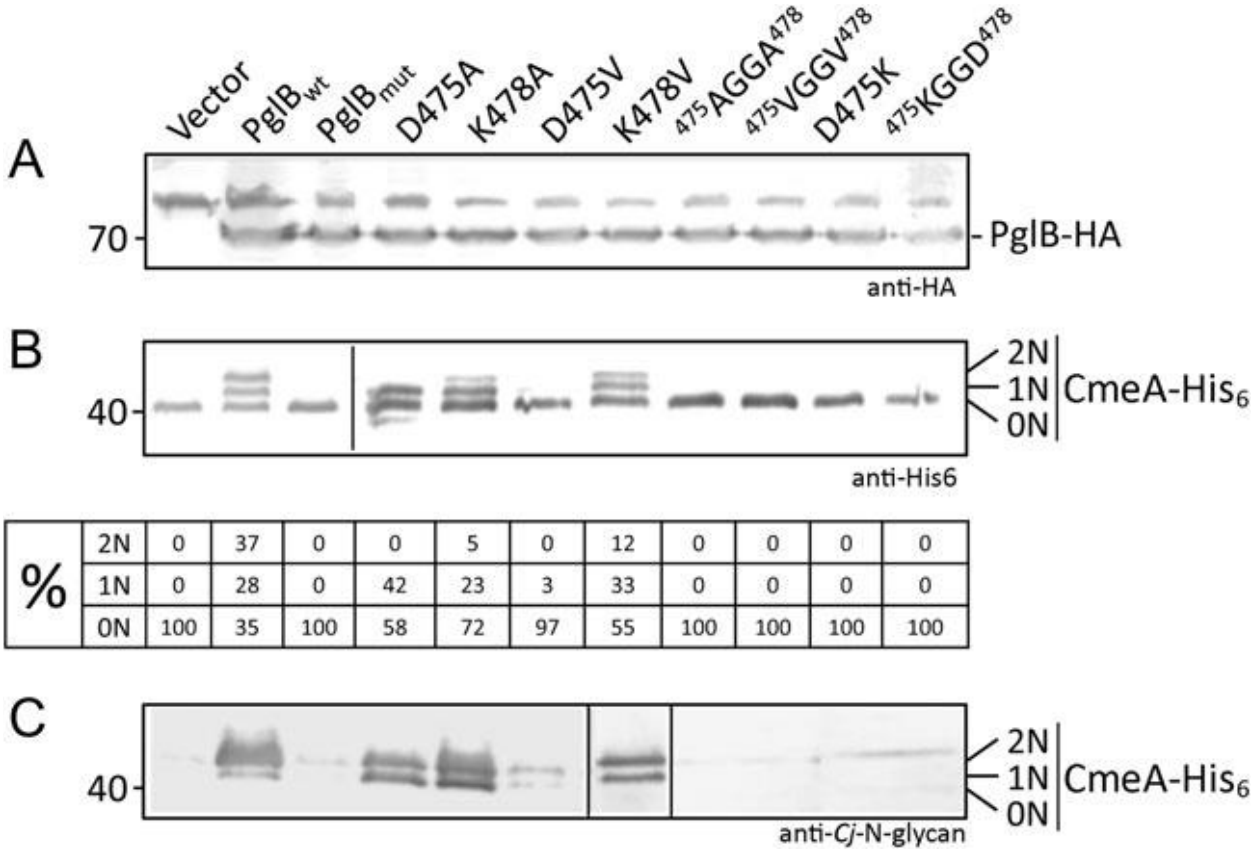
amino acid with similar charge, white indicates no amino acid conservation at this position. Multiple sequence alignments and figures were generated with ClustalW and BOXshade; http://www.ch.embnet.org/.

5.2.2 Activity of PglB DGGK point mutants in *C. jejuni* and analysis of fOS production

Next, we investigated PglB activities in the native host. We expressed the various *pglB* alleles in the *C. jejuni pglB* knock-out strain and analyzed the glycosylation status of native *Cj*-CmeA by western blotting using *Cj*-CmeA-specific antibodies (Figure 4). Similar to what was observed in the heterologous *E. coli* system, expression of PglB K478A and K478V resulted in a reduction of *Cj*-CmeA glycosylation that was further reduced upon expression of PglB D475A. Expression of the PglB double mutations AGGA, VGGV, KGGD and the single amino acid substitutions D475K and D475V resulted in complete loss of native *Cj*-CmeA glycosylation. Complementation of *pglB* with wildtype *Cj*-PglB (positive control) led to restoration of *Cj*-CmeA glycosylation when compared to the wildtype strain whereas only non-glycosylated *Cj*-CmeA was observed upon expression of inactive PglB (WAAYG, negative control).

To investigate whether the DGGK motif is also required for the hydrolytic activity of PglB, we followed the release of the N-glycan structure as fOS into the periplasm of *C. jejuni*. We applied semi-quantitative mass spectrometry to determine the fOS levels in normalized whole cell lysates of *C. jejuni* wildtype, *pglB* and *pglB* complemented with PglB D475V and K478V single point mutants (Figure 5). The fOS amount obtained with the wildtype strain was used as the reference and set to 100%. No fOS could be detected in the *pglB* knock-out strain itself as well as in the presence of PglB WAAYG. Expression of the *Cj*-wildtype *pglB* allele in the *pglB* knock-out resulted in partial complementation of the fOS-negative phenotype, with levels at 50%

when compared to the wildtype strain. Significantly reduced amounts of fOS levels were observed upon expression of PglB D475V and K478V. The fOS levels reached 11% and 20% when compared to the fOS levels produced in the *pglB* mutant complemented with the wildtype *Cj-pglB* allele. The decrease in the formation of fOS and therefore the reduction in PglB hydrolase activity was comparable with the reduction in PglB N-glycan transfer activity as determined above.



%	2N	0	37	0	0	5	0	12	0	0	0	0
	1N	0	28	0	42	23	3	33	0	0	0	0
	0N	100	35	100	58	72	97	55	100	100	100	100

Fig. A.3. In vivo activity of *C. jejuni* PglBs with specific amino acid changes in D and K within the DGGK motif. (A) Expression of PglB proteins in *E. coli* whole cell lysates was followed with HA-specific antiserum. The glycosylation pattern of *Cj*-CmeA-His expressed in the presence various PglB alleles was determined by western blotting with (B) His-Tag and (C) anti-N-glycan specific antiserum. The glycosylation status of *Cj*-CmeA-His is indicated on the right

(0N, 1N, 2N refer to none-, mono- and di-glycosylated *Cj*-CmeA-His, respectively). Western blot signals corresponding to *Cj*-CmeA-His (from (B)) were further quantified by densitometry. The results are expressed as percentage normalized to the values obtained after co-expression with *Cj*-wildtype PglB (set to 100%). Relevant molecular weight markers (in kDa) are indicated on the left.

Fig. A.4. In vivo complementation of *C. jejuni pglB* with the DGGK *pglB* point mutants. The presence and the glycosylation pattern of native *Cj*-CmeA protein in whole cell lysates of *C. jejuni* wildtype, *pglB* mutant and the *pglB* mutant expressing various *pglB* alleles was determined by western blotting with *Cj*-CmeA-specific antiserum. The glycosylation status of native *Cj*-CmeA is indicated on the right (0N, 1 N, 2 N refer to none-, mono- and diglycosylated *Cj*-CmeA, respectively). Relevant molecular weight markers (in kDa) are indicated on the left.

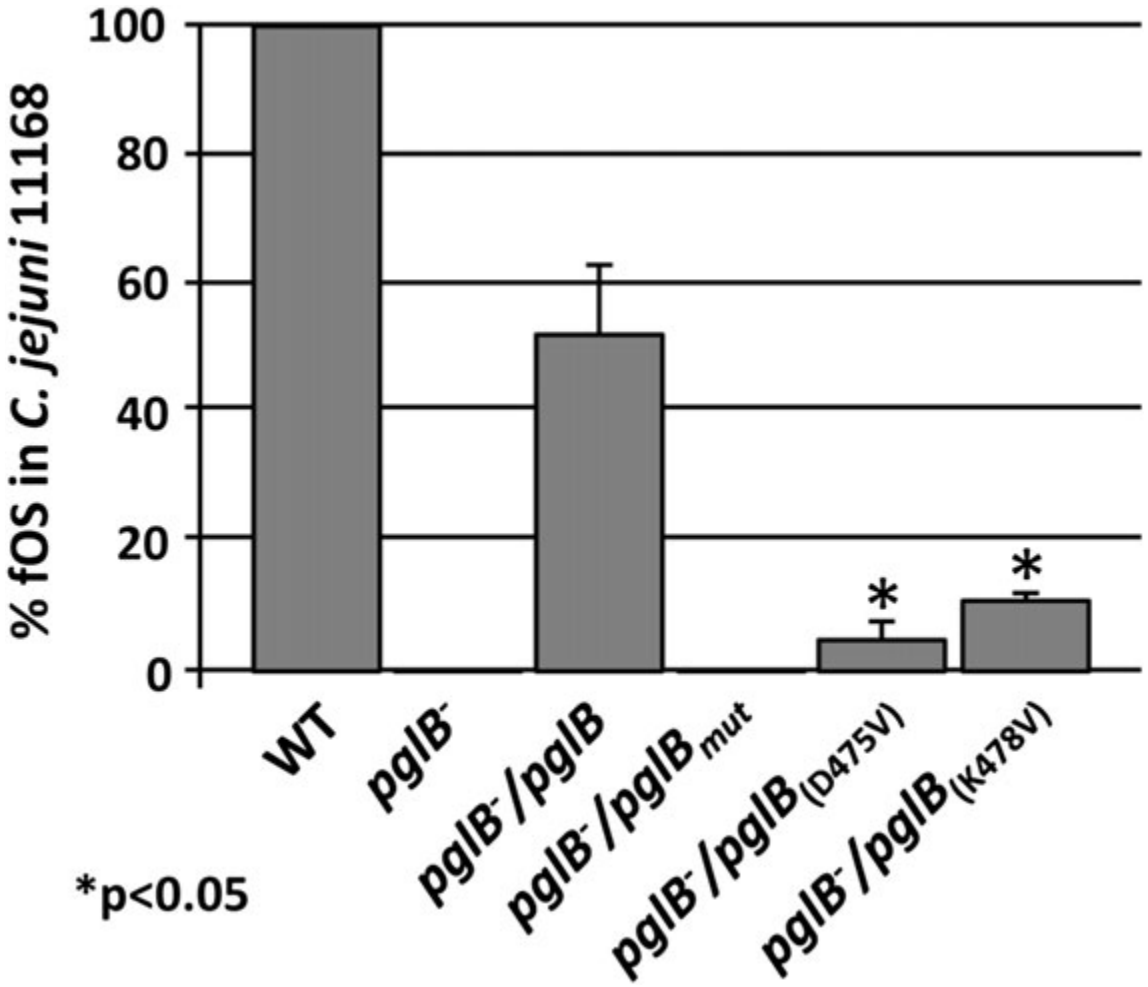


Fig. A.5. Mutation in DGGK results in reduced fOS release in *C. jejuni*. Amounts of fOS released in *C. jejuni* wildtype, *pglB* and *pglB* complemented with various *pglB* mutant alleles (as indicated) was determined by semiquantitative mass spectrometry. Bars represent the mean calculated from values obtained from three independent biological replicates. Variations are indicated by error bars, the asterisks indicate a statistically significant difference in fOS amounts ($P < 0.05$) when compared to *C. jejuni pglB* mutant complemented with the wt-*pglB* allele.

5.2.3 Limited proteolysis of PglB proteins suggests involvement of ^{475}DGGK^{478} in substrate binding

To further investigate the possible involvement of the DGGK motif in either directly mediating substrate binding or affecting protein conformational changes related to substrate binding, we performed limited proteolysis experiments in the presence or absence of *Cj*-LLOs (Figure 6). *Cj*-LLO and control extracts were prepared and characterized from *E. coli* expressing *the C. jejuni pgl* operon as described in Supplementary data, Figure S1. Full length PglB proteins that migrate at around 70 kDa were readily degraded over the 60 min timeframe of the assay. Specific proteolytic fragments containing the C-terminal HA-tag were observed by western blotting after 10, 20, 30, 45 and 60 min of incubation with proteinase K followed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. For wildtype PglB, the greatest change in the proteolytic digestion pattern in the presence of *Cj*-LLOs is an increase in resistance to proteolytic digestion for a fragment of ~26 kDa, at longer digestion times of 45–60min. In mutants K478A, K478V and D475A, the proteolytic digestion patterns are similar to the patterns seen for wildtype PglB. In contrast, in mutants D475V, D475K, AGGA, VGGK and KGGD, the degradation of the 26kDa fragment at all time points did not appear to be affected by the presence of *Cj*-LLOs. Notably, for the mutant in which the highly conserved WWDYG motif is replaced by WAAYG, the 26 kDa fragment appears to form more readily at early time points of 0–10min in the presence of *Cj*-LLOs. No degradation of PglB proteins was observed over a 60min timeframe in the absence of proteinase K (Supplementary data, Figure S2). Inspection of the three-dimensional structure of PglB suggests that the most likely cleavage site resulting in the formation of a 26 kDa C-terminal fragment lies in the loop immediately following the DGGK motif (Supplementary data, Figure S3). Cleavage between L480 and G481 would result in the

formation of a 27.8 kDa fragment expected to migrate near the location of the 26 kDa band seen in SDSPAGE. The other major C-terminal fragment (~45 kDa) produced by proteinase K is predicted to be produced by a cleavage event in EL5. For example, cleavage between F308 and N309 yields a 47.5 kDa fragment expected to migrate near the location of the 45 kDa band seen in SDS-PAGE.

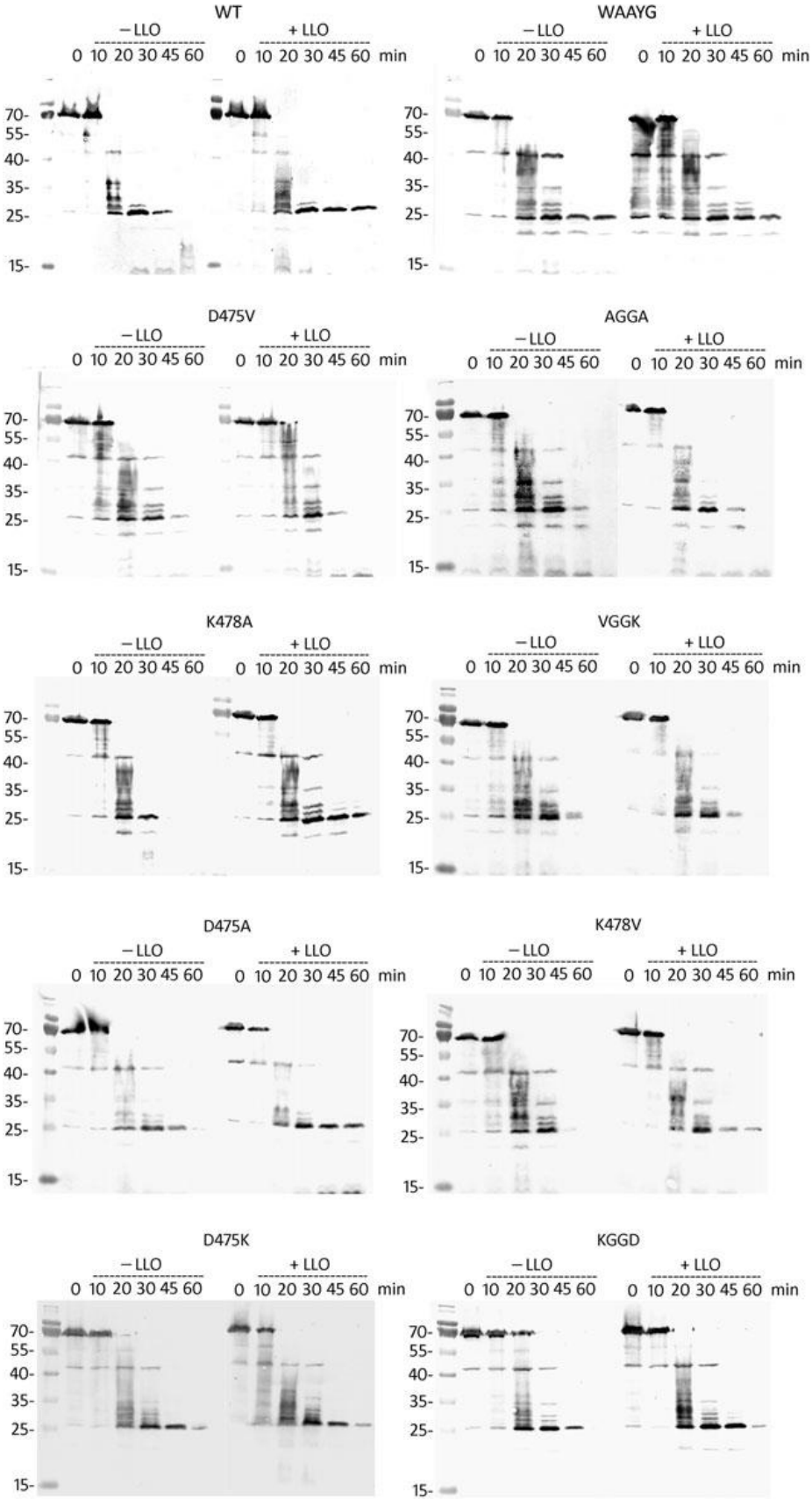

WT
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
WAAYG
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
D475V
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
AGGA
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
K478A
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
VGGK
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
D475A
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
K478V
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
D475K
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-
KGGD
– LLO
+ LLO
0 10 20 30 45 60 0 10 20 30 45 60 min
70-
55-
40-
35-
25-
15-

Fig. A.6. Limited proteolysis of PglB proteins in the presence and absence of LLO. *C. jejuni* PglB proteins (indicated above each panel) expressed in *E. coli* and enriched in CEF were subjected to proteinase K digestion after incubation with *Cj*-LLOs (+LLO) or a control preparation that does not contain LLOs (−LLO). Formation of proteolytic fragments over time (as indicated above each panel) was followed by western blotting with HA-specific antiserum. Full length

PglB migrates at 70 kDa, relevant sizes of the molecular weight marker (MWin kDa) are indicated on the left of each figure panel.

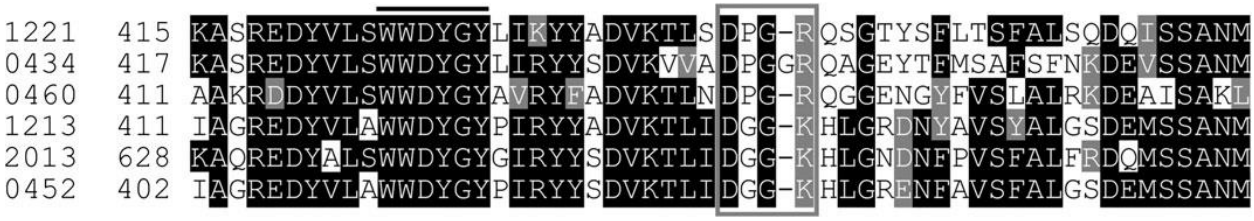


Fig. A.7. *Campylobacter* species with two PglB orthologues. Sequence alignment of PglB protein sequences from *Campylobacter* species that possess two PglB orthologues. The highly conserved WWDYG motif is indicated by a black bar; the region containing the DGGK sequence motif (either present or absent) is boxed. Highlighted in black are amino acids that are conserved in at least 50% of the sequences, grey highlighting indicates the presence of an amino acid with

similar charge, white indicates no amino acid conservation at this position. Ccu, *Campylobacter curvus* (CCV52592_1221, A7GWV7; CCV52592_1213, A7GWW5); Ccon, *Campylobacter concisus* (CCC13826_0460, A7ZEU3; CCC13826_0452, A7ZET6); Cgr, *Campylobacter gracilis* (CAMGR0001_2013, C8PLK3; CAMGR0001_0434, C8PHI9). Multiple sequence

alignments and figures were generated with ClustalW and BOXshade: Http://www.ch.embnet.org/.

5.2.4 Certain *Campylobacter* species possess two PglB orthologues

In a previous study, we identified three *Campylobacter* species that possess two PglB orthologues in their chromosome; *C. curvus* 525.92, *C. concisus* 13826 and *C. gracilis* RM3266 (Nothaft and Szymanski 2010). A similar observation was made for certain *Helicobacter* species (Jervis et al. 2010). Our *in silico* analysis revealed that in campylobacter, only one of these PglB orthologues possesses the conserved DGGK motif, whereas the WWDYG motif is conserved in both copies (Figures 7 and 8). To functionally analyze these *pglB* alleles we cloned and expressed the corresponding proteins from *C. curvus* and *C. gracilis* in the heterologous *E. coli* system in the presence of pACYC184(*pg*lmut) encoding the *Cj*-glycosylation operon with an inactive PglB. The glycosylation status of the soluble version of the *Cj*-CmeA-His6, expressed from plasmid pWA2 served as a read-out for PglB activity (Figure 9A). Expression of the PglB-HA proteins was confirmed in *E. coli* whole cell lysates after western blotting with HA-tag specific antiserum. All PglB proteins migrate at approximately 70 kDa, although their calculated mass is >20% larger. The anomalous migration behavior of proteins containing multiple transmembrane domains when analyzed by SDS-PAGE has previously been described and a more compact, incompletely unfolded conformation for full-length PglB-HA may explain in part a rate of migration that is higher than expected for a typical soluble protein (Rath et al. 2009). The additional HA-specific bands in the higher MW range observed for Cgr2013, Cgr0434 and Ccu1213 could either represent fully unfolded PglB protein that migrates at the correct theoretical mass or might be due to a combination of oligomerization and incomplete

unfolding of the respective protein.

Western blot analysis of purified His6-tagged *Cj*-CmeA coexpressed with pACYC184(*pgl*mut) and pWA1 (PglB WAAYG) led to the formation of purely non-glycosylated *Cj*-CmeA while expression of the *Cj*-wildtype *pglB* allele from pMAF10 resulted in the formation of three bands that corresponded to non- mono- and diglycosylated *Cj*-CmeA (Figure 9A upper panel); only the two higher molecular weight signals reacted with the *Cj*-N-glycan-specific antiserum (Figure 9A middle panel). Expression of the PglB proteins from *C. curvus* and *C. gracilis* only resulted in the formation of glycosylated *Cj*-CmeA when the orthologue with the DGGK motif was co-expressed (i.e. Ccu1213 and Cgr2013) while only non-glycosylated *Cj*-CmeA could be observed upon co-expression of the protein variant that does not possess the DGGK motif (i.e. Ccu1221 and Cgr0434). In parallel, western blot analyses carried out with anti-HA (Figure 9A, lower panel) showed similar levels of expression of the various PglB proteins, indicating that the absence of glycosylation activity is indeed the result of the missing N-glycan transfer function of the tested PglB proteins.

Next, we investigated the N-glycosylation potential of these PglB proteins in the *Campylobacter* system for their ability to complement the *Cj*-pglB mutant. Western blot analyses of whole cell lysates probed with N-glycan-specific antiserum (Figure 9B) revealed that N-glycosylation in whole cell lysates could only be observed in samples prepared from cells that express Cgr2013 or Ccu1213 in the *Cj-pglB* mutant background. No glycosylation activity could be determined upon expression of Cgr0434 and Ccu1221 (that do not have the DGGK motif) as evidenced by the lack of N-linked glycoproteins similar to the complementation with the inactive *Cj-pglB* allele. It is worth mentioning that we could not detect fOS in extracts of the *Cj-pglB* mutant complemented with active Cgr2013 or Ccu1213 (results not shown) indicating that the levels of

fOS were either below our detection limit or that PglB enzymes from other *Campylobacter* species do not hydrolyze the *Cj*-glycan from the lipid-linked-N-glycan LLO intermediates, although fOS is naturally produced by these species [29].

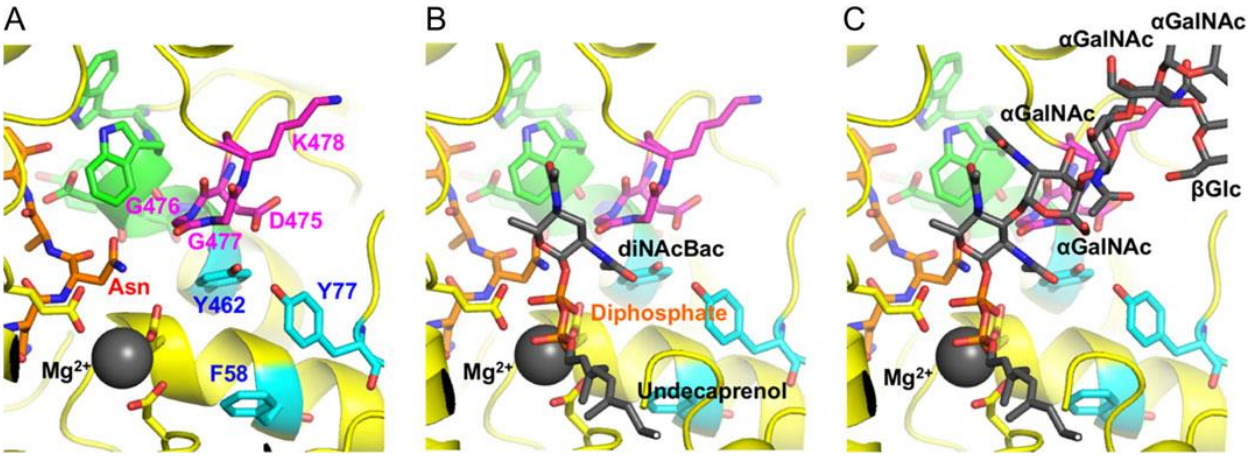


Fig. A.8. Docking model of undecaprenyl-pyrophosphate-glycan donor bound to *C. lari* PglB. Close-up views are drawn from the same orientation as shown in Figure 1B for the active-site regions of (A) the crystal structure of the *C. jejuni* PglB peptide complex (PDB 3RCE) and (B) a model of the undecaprenylpyrophosphate-glycan donor that was manually docked into the active site by aligning the diphosphate moiety with the magnesium ion and the anomeric carbon of the diNAcBac residue with the side chain amide nitrogen atom of the Asn residue in the peptide acceptor. (C) The full GalNAc-αl,4-GalNAc-αl,4-[Glcβ1,3]GalNAc-αl,4-GalNAc-αl,4-GalNAc-αl,3-diNAcBAc glycan structure is drawn. In panel (A), the residue numbering follows that for *C. jejuni* PglB, even though the structure that is shown is from the closely related PglB homolog from *C. lari*. This figure is available in black and white in print and in color at Glycobiology online.

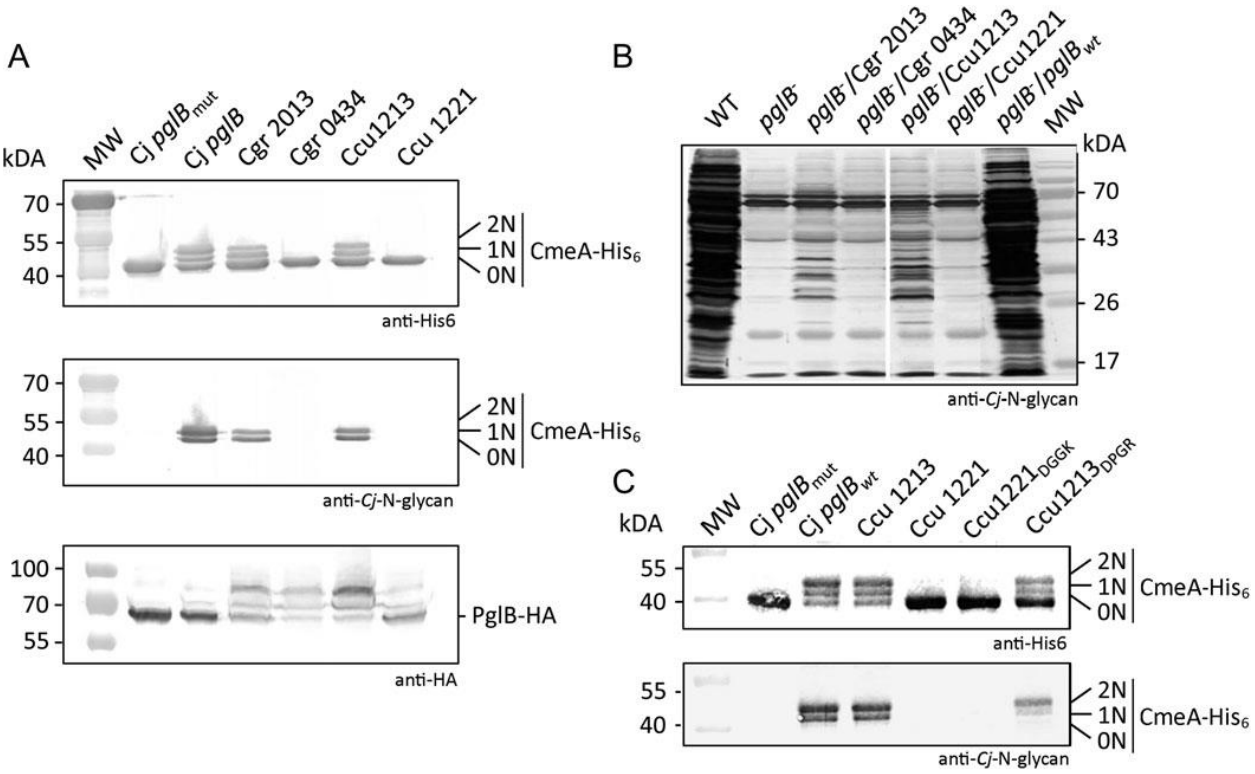


Fig. A.9. Functional analyses of PglB proteins from species with two PglB orthologues. (A) Functional analysis of PglB proteins in the heterologous *E. coli* system. The glycosylation pattern of purified *Cj*-CmeA proteins co-expressed in the presence of the *C. jejuni* protein glycosylation operon with an inactive PglB and various PglB proteins was determined by western blotting using His-tag specific (upper panel) and *Cj*-N-glycan-specific (middle panel) antisera. The glycosylation status of the *Cj*-CmeA protein is indicated on the right; 0N, 1N, 2N refer to non-, mono-, and di-glycosylated *Cj*-CmeA. In parallel, western blots using HA-tag specific antisera were performed to follow expression of the various HA-tagged PglB proteins (lower panel). Relevant molecular weight markers (in kDa) are indicated on the left. *Cj-pglB*mut, *Cj*-PglB with mutation in WWDYG; *Cj-pglB*, native *Cj* PglB; Cgr2013, PglB orthologue form *C. gracilis* with DGGK; Cgr0434, PglB orthologue form *C. gracilis* without DGGK; Ccu1213, PglB, orthologue form *C. curvus* with DGGK; Ccu1213, PglB, orthologue form *C. curvus*

without DGGK. (B) Functional analysis of PglB proteins in the *C. jejuni* pglB mutant. The overall protein N-glycosylation pattern in whole cell lysates of *C. jejuni* wildtype (*Cj*-wt), *pglB* mutant (*pglB*-) and the *pglB* mutant expressing various PglB alleles (as indicated) was determined by western blots with *Cj*-N-glycan-specific antiserum. Relevant molecular weight markers (in kDa) are indicated on the left. (C) A functional DGGK is required but not sufficient for PglB activity. Functional analysis of PglB proteins in the heterologous *E. coli* system as described above using His-tag specific or *Cj*-N-glycan-specific antisera. Ccu1213, PglB orthologue form *C. curvus* with DGGK; Ccu1221, PglB orthologue form *C. curvus* without DGGK; Ccu1221 DGGK PglB, orthologue form *C. curvus* with DPGR to DGGK; Ccu1213 DGPR PglB, orthologue form *C. curvus* with DGGK to DPGR.

5.2.5 Replacement of DGGK alone is not sufficient to restore PglB activity in *C. curvus* 1221

We next wanted to investigate if PglB functionality of Ccu1221 can be restored by introducing the DGGK motif into the correct position within the protein. We also did the opposite experiment replacing the functional DGGK motif in Ccu1213 with the Ccu1221 DPGR motif. Functional analysis of the Ccu1221 (DPGR to DGGK) and the Ccu1213 (DGGK to DPGR) PglB proteins after co-expression with the inactive *Cj-pgl* operon (pACYC184(*pgl*mut)) and the soluble N-glycan acceptor *Cj*-CmeA-His6 in *E. coli* (Figure 9C) indicated that Ccu1213 (DGGK to DPGR) is still able, albeit with lower efficiency, to N-glycosylate the *Cj*-CmeA-His6 protein (Figure 9C), whereas for Ccu1221 (DPGR to DGGK) only one band that corresponded to non-glycosylated *Cj*-CmeA-His6 was observed similar to the inactive *Cj*-PglB (WAAYG) allele, as well as the native form of Ccu1221. In contrast, non-, mono- and di-glycosylated *Cj*-CmeA-

His6 proteins were detected in western blots with anti-His6-specific antiserum (Figure 9C, upper panel) and N-glycan-specific antiserum (Figure 9C, lower panel) upon co-expression of the *Cj*-PglB wildtype allele, Ccu1213 and Ccu1213 (DGGK to DPGR).

5.3 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 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 [4,30].

A structure-guided sequence alignment that included the yeast *STT3* and the archaeal *P. furiosus* AglB sequences revealed a ^{571}DXXK^{574} sequence (DINK in yeast *STT3*, and DWAK in *P. furiosus*) rather than the more defined ε-proteobacterial DGGK motif that was demonstrated through mutagenesis studies to be essential for OTase activity in yeast [10]. 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 DWAK sequence in AglB actually corresponds to ^{568}MXXI^{571} in PglB, and that OTase activity in vitro was reduced resulting from an I571A substitution [28]. 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) and alanine mutations of residue D54 in *Cj*-PglB resulted in nearly complete loss of PglB activity in vitro [28]. The crystal structure of the full-length *Cl*-PglB 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 [27]. The peptide bound structure of PglB reveals that the β-hydroxyl group of the +2Thr forms hydrogen bonds with each of the side

chains of the WWD residues. The structure also shows that I571 provides contact with the +2 Thr, but since the Ile within the MXXI motif is not conserved in all *STT3* homologs, the interactions seen with Ile are either not essential or can be replaced by interactions with other residues.

Other studies have investigated that the importance of specific amino acid residues in the PglB protein with the goal to learn more about the function of the protein or to change and potentially improve PglB function with respect to glycoengineering in the heterologous *E. coli* system [31-34]. In this study, analyses of available crystal structures and alignments of PglB sequences from multiple *Campylobacter* species revealed a second conserved ^{475}DGGK^{478} sequence. In the *Cj*-PglB, this sequence motif is located 12 amino acids downstream of the conserved WWDXG motif. Site-directed mutagenesis of the D and K amino acids was carried out to show the importance of these residues with respect to OTase and hydrolase function. Alanine substitutions of D475 and K478 in the *Cj*-PglB showed a minor decrease in glycosylation activity compared to wildtype PglB.

Single valine substitutions of the D475 and K478 residues in PglB resulted in reduced glycosylation efficiency of D475V, but not K478V in the heterologous *E. coli* host, indicating that introducing a branched and larger hydrophobic residue in place of D475 was more disruptive than replacing K478 with a branched, hydrophobic residue. The effects of alanine and glutamate substitutions at the D and K residues in the DGGK motif of *Cj*-PglB were previously analyzed using an in vitro glycosylation assay [31]. D475 was implicated in PglB function, since a significant reduction in activity was observed when D475 was replaced with alanine. Our results indicate that the double alanine or double valine mutations at both the D and K sites result in the

complete loss of *Cj*-CmeA glycosylation, an effect more drastic than the single D475 substitution. Our data also suggest that both, the D and the K, residues are important for the N-glycosylation activity of PglB. The introduction of hydrophobic residues in place of both D475 and K478 potentially destabilizes the structure of the DGGK loop more than a single-site mutation, with the effect being more pronounced for the D475V mutation. In addition, we show that the same mutations that adversely affect the N-glycosylation activity of PglB in the heterologous *E. coli* system and in the native *C. jejuni* host also led to a decrease in the hydrolase activity of *C. jejuni* PglB and subsequently in the production of lower amounts of fOS [16].

The PglB crystal structures from *C. lari* and the soluble C-terminal domain of *C. jejuni*, together with a homology model of full-length PglB from *C. jejuni* reveal that the ^{475}DGGK^{478} motif is located directly next to the highly conserved WWDYG motif (Figure 1). In addition, the DGGK motif is near the active-site divalent cation modeled as magnesium ion according to the prediction of Lizak et al. (2011) and the side chain of the Asn residue of the peptide acceptor substrate. The carboxylate group in the side chain of D475 accepts hydrogen bonds from the backbone amide groups of the highly conserved G476 and G477 residues; the carboxylate group also appears to accept hydrogen bonds from the side chain hydroxyl group and guanidino group of the highly conserved Y462 and R465 residues, respectively. The D475V and D475A mutants of *C. jejuni* PglB showed nearly complete loss of activity, presumably in part because of a disruption of the hydrogen-bonding network formed by the carboxylate group in the side chain of D475. Replacing D475 with lysine in the D475K single and KGGD double mutants also led to a loss of activity that is consistent with charge repulsion between the positively charged lysine side chain and R465, as well as the loss of other stabilizing interactions normally formed by the carboxylate side chain.

Even though the specific functional role of the DGGK motif in PglB catalysis is not known at this time, the proximity of the motif to the WWDYG motif, the active-site divalent ion and the peptide acceptor asparagine side chain all suggested a possible role in binding the donor glycan. Limited proteolysis data provide additional support for the involvement of this motif in donor glycan binding. Most significantly, the protection of the 26 kDa C-terminal fragment to proteolytic digestion by LLOs in all of the enzymatically active PglB mutants indicates that LLO binding likely stabilizes interactions between the watersoluble C-terminal domain and the N-terminal transmembrane domain. The fact that the in vivo PglB activities for the various DGGK mutants can be correlated to the stability of this 26 kDa fragment further suggest that certain DGGK mutants are no longer able to bind the substrate efficiently. Somewhat related to these observations, the cleavage site leading to the formation of the 26 kDa fragment appears to be more accessible to proteinase K in the WAAYG mutant relative to the wildtype enzyme containing the WWDYG motif. In the folded structure of PglB, the WWDYG motif is adjacent to the loop consisting of residues 474–485, which includes the ^{475}DGGK^{478} motif and the putative cleavage site between residues L480 and G481 leading to the formation of the 26 kDa fragment. Disruption of the highly conserved WWDYG motif's interactions with the adjacent loop may account for the increased sensitivity of the loop to proteinase K digestion. To help further explore the structural basis of LLO-dependent effects on protease sensitivity, we have constructed a model of the undecaprenylpyrophosphate glycan and docked it into the active sites of *C. lari* PglB and the homology model of *C. jejuni* PglB. These models indicate that the di-N-acetylbacillosamine (diNAcBAc) residue lies within 3–4Å of G477, and the side chain of Y462 that forms a hydrogen bond with the carboxylate side chain of D475 is also positioned close enough to directly interact with both the diNAcBac residue and the α(1-4)-linked

GalNAc residue. These models suggest that the DGGK motif either directly interacts with residues in the donor glycan or affects the conformation of nearby residues that in turn interact with the donor glycan. These models are also consistent with previously reported data from Jaffee and Imperiali [31] indicating that increasing the concentration of the polyprenyldiphosphate glycan donor partially compensated for the loss of activity seen in the D475A and D475E mutants, as well as recently reported molecular dynamics simulations of substrate complexes formed with PglB [35]. These models may also help to explain why all members of the *Campylobacter* genus form diNAcBac-GalNAc at the reducing terminus of their oligosaccharide and show divergence of structure beyond the second sugar [29].

Our observation that PglB orthologues in which DGGK is replaced by DPGR (*C. curvus* and *C. concisus*) and DPGGR (*C. gracilis*) lack glycosylation activity prompted us to investigate how changes in this motif may affect the structure of PglB. Homology models of these orthologues suggest that the aspartate residue adopts the same position, and the replacement of the first glycine with proline is not expected to significantly affect the structure of the motif or nearby residues (Supplementary data, Figure S4). The homology models also suggest that the replacement of the lysine residue with arginine or the insertion of an extra glycine and replacement of lysine with arginine are unlikely to perturb the global structure or folding of PglB significantly, because this motif is mostly exposed to solvent and these changes are thus not expected to affect the folding or conformation of other parts of the structure. Instead of these variations in structure and sequence affecting the overall PglB fold, it seems more likely that the compact structural motif created by these four or five residues forms a portion of the binding site for the donor glycan, in combination with residues from the extended EL1 loop and perhaps also the extended EL5 loop. Although EL5 is partially disordered in the crystal structure of PglB

that was determined in the absence of the donor glycan, it is positioned near where the donor glycan is expected to be, and therefore is predicted to interact with and stabilize the donor glycan prior to glycosyl transfer [34-36]. Since the DGGK motif likely combines with nearby residues from loops EL1 and EL5 to form the majority of the binding site for the acceptor glycan, we predict that the activity of the inactive PglB enzymes from *C. curvus, C. concisus* and *C. gracilis* may only be rescued if adjacent portions of loops EL1 and EL5 were also modified in addition to the sequence of the DGGK motif.

In summary, our results indicate that the DGGK motif plays an important role for PglB activity. However, the presence of the DGGK motif alone, although in combination with other necessary sequence motifs (like WWDYG) that have been shown to be required for PglB activity, is not sufficient for the generation of a functional OTase. This and other studies have used the known structural and sequence requirements of bacterial OTases not only to identify and study additional PglB proteins [23,24], but also (using targeted and non-targeted approaches in combination with novel high throughput assays) to identify and to study amino acid residues and motifs necessary for the interaction of PglB with the protein acceptor, its glycosylation site preference or the LLO donor [32,34,37]. These studies have already helped to understand protein function in more detail and demonstrated that it is possible to change PglB substrate specificities [34] for the generation of novel glycoconjugate vaccines. We now describe a conserved DGGK structural motif in the PglB OTases of the *Campylobacter* genus that is necessary for N-linked glycoprotein production and fOS release. We predict this motif is involved in binding of the donor undecaprenyl-pyrophosphate glycan substrate and is therefore required for OTase activity. Unlike in eukaryotes where duplicate STT3 enzymes play a supporting role for N-

glycosylation, the reason inactive PglB enzymes are maintained in several *Campylobacter* and *Helicobacter* species remains to be resolved.

5.4 Materials and Methods

5.4.1 Bacterial strains, plasmids and growth conditions

The bacterial strains and plasmids used in this study are listed in Table I. *Escherichia coli* strains were grown in Luria-Bertani (LB) broth or on LB agar plates. *Campylobacter jejuni* NCTC 11168 strains were grown on Mueller Hinton (MH, Difco) agar plates under microaerobic conditions (10% CO2, 5% O2, 85% N2) at 37°C for 18 h. *Campylobacter curvus* and *Campylobacter gracilis* were grown on BHI agar supplemented with 5% horse blood under anaerobic conditions. Ampicillin (100 μg/mL), chloramphenicol (25 μg/mL), kanamycin (25 μg/mL), trimethoprim (25 μg/mL) and tetracycline (25 μg/mL) were added to the medium as needed for selection.

Table A.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
CLM24	W3110, Δ*waaL*	Feldman et al. (2005)
C600 (RK212.2)	*leu thr thi lacy supE44 tonA;* pRK212.2, Amp^R, Tet^R	Figurski and Helinski (1979)
K12 *wzy*::*kan*	*E. coli* K12 O-antigen polymerase mutant, Km^R	Baba et al. (2006)
K12 *waaL*::*kan*	*E. coli* K12 O-antigen ligase mutant, Km^R	Baba et al. (2006)
K12 Δ*wzy*	*E. coli* K12 O-antigen polymerase deletion mutant	This study
K12 Δ*wzy*/*waaL*::*kan*	*E. coli* K12 O-antigen polymerase, O-antigen ligase double mutant, Km^R	This study
C. jejuni		
11168 NCTC	Clinical isolate used for genome sequencing	Parkhill et al. (2000)
11168-*pglB*::*kan*	*pglB* mutant, Km^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* containing point mutations W458A and D459A in PglB, Cm^R	Wacker et al. (2002)
pWA2	Soluble periplasmic *C. jejuni CmeA*His_6 under control of Tet promoter, in pBR322, Amp^R	Feldman et al. (2005)
pMLBAD	Cloning vector, arabinose-inducible, Tmp^R	Lefebre et al. (2005)
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
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
pCcu1213-EC	*C. curvus* 1213 *pglB*-$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCcu1221-EC	*C. curvus* 1221 *pglB*-$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCgr0434-EC	*C. gracilis* 0434 *pglB*-$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCgr2013-EC	*C. gracilis* 2013 *pglB*-$_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCcu1221DGGK-EC	*C. curvus* 1221 *pglB*(DPGR to DGGK) $_{HA}$ cloned into pMLBAD, Tmp^R	This study
pCcu1213DGPR-EC	*C. curvus* 1213 *pglB*(DGGK to DPGR) $_{HA}$ 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
pCcu1213-Cj	*C. curvus* 1213 *pglB* cloned in pCE111-28, Cm^R	This study
pCcu1221-Cj	*C. curvus* 1221 *pglB* cloned in pCE111-28, Cm^R	This study
pCgr0434-Cj	*C. gracilis* 0434 *pglB* cloned in pCE111-28, Cm^R	This study
pCgr2013-Cj	*C. gracilis* 2013 *pglB* cloned in pCE111-28, Cm^R	This study

5.4.2 Construction of plasmids

Amino acid substitutions in the *pglB* genes were generated using oligonucleotides listed in Table II. 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 *Cj-pglB* mutations were introduced following a modified three-piece ligation method

described by Kato et al. (2009). Two polymerase chain reaction (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*$_{\text{BamHIF}}$ with each of the oligonucleotides carrying the amino acid substitutions within the DGGK motif: D475A, K478A, D475K, AGGA, VGGV and KGGD. The C-terminal fragment of *pglB*$_{Cj}$ was amplified using oligonucleotides *pglB*-C1-F and *pglB*HAKpnI-R. The N-terminal and C-terminal *pglB* fragments were digested with *Eco*RI and *Kpn*I, respectively, to generate blunt-ended PCR products that were cloned in the same sites of the pMLBAD empty vector.

Genes encoding *C. gracilis* and *C. curvus* PglB orthologous proteins were PCR amplified from chromosomal DNA of the respective strain with oligonucleotides (also see Table II) CCV52592_1213-*Nco*I-F and CCV52592_1213-*Hind*III-R, for *C. curvus* 1213; CCV52592_1221-*Nco*I-F, CCV52592_1221-*Hind*III-R for *C. curvus* 1221; Cg0434-F-*Nco*I, Cg0434-R-*Pst*I, for *C. gracilis* 0434 and Cg2013-F-*Nco*I, Cg2013-R-*Pst*I for *C. gracilis* 0434, introducing an HA-Tag sequence at the 3′end of each *pglB* gene. For expression in *E. coli* obtained PCR products were digested with *Nco*I and *Hind*III (for *C. curvus pglB* alleles) or with *Nco*I and *Pst*I (for *C. gracilis pglB* alleles) and ligated into plasmid pMLBAD digested with the appropriate enzymes. Positive candidates containing the desired *pglB* allele were verified by restriction analysis and DNA sequencing. To generate the DPGR to DGGK mutation in Ccu1221, two PCR fragments were generated using the Ccu1221-pMLBAD derivative as a template with the oligonucleotide combinations CCV52592 1221-*Nco*I, 1221_DGGK-R-P and 1221_DGGK-F-P (introducing the DGGK mutation) plus CCV52592_1221-*Hind*III-R. To generate the DGGK to DPGR mutation in Ccu1213, two PCR fragments were generated using

the Ccu1213-pMLBAD derivative as a template with the oligonucleotide combinations CCV52592 1213-*Nco*IF with 1213_DGPR-R-P and 1213_DGPR-F-P (introducing the DGPR mutation) with CCV52592_1213-*Hind*III-R. The PCR fragments were generated using *Pfx* polymerase and the subsequent three arm ligation [38] was carried out using the *Nco*I digested N-terminal and the *Hind*III-digested C-terminal PCR products together with pMLBAD cut with *Nco*I-*Hind*III. Plasmids from clones obtained after transformation were verified by restriction analysis and DNA sequencing.

For the *C. jejuni* NCTC 11168 *pglB* mutant complementation assays, *pglB* alleles were amplified by PCR with oligonucleotides *pglB*BamHI-F and *pglB*XhoI-HA-R using Vent polymerase (New England Biolabs) and the respective pMLBAD derivative as a template. PCR products were digested with *Bam*HI and *Xho*I (or by digestion with *Bam*HI only if an internal *Xho*I site was present in the *pglB* allele) and subcloned into the *C. jejuni-E. coli* shuttle vector pCE111-28 treated with the same enzymes (or with *Bam*HI-*Eco*RV if an internal *Xho*I site was present in the *pglB* allele). Plasmids obtained after transformation of *E. coli* DH5alpha (Table I) were confirmed by restriction analysis. *E. coli* pRK212.2 was transformed with positive shuttle-plasmid candidates. Shuttle plasmids expressing the various *pglB* alleles were mobilized from *E. coli* pRK212.2 into *C. jejuni* NCTC 11168 as described [39,40].

Table A.2. Oligonucleotides used in this study

Oligonucleotide	Sequence 5′-3′	Application
D475V-F	CGATGTGAAAACTTTAGTAGTTGGTGGAAAGCATTTAGGTAAGG	Mutagenesis
D475V-R	CCTTACCTAAATGCTTTCCACCAACTACTAAAGTTTTCACATCG	Mutagenesis
K478V-F	CTTTAGTAGATGGTGGAGTACATTTAGGTAAGGATAATTTTTTCC	Mutagenesis
K478V-R	GGAAAAAATTATCCTTACCTAAATGTACTCCACCATCTACTAAAG	Mutagenesis
D475A-R	CCTTACCTAAATGTTTTCCACCCGCTACTAAAGTTTTCACATCG	Mutagenesis
K478A-R	CCTTACCTAAATGCGCTCCACCATCTACTAAAGTTTTCACATCG	Mutagenesis
AGGA-R	CCTTACCTAAATGCGCTCCACCCGCTACTAAAGTTTTCACATCG	Mutagenesis
VGGV-R	CCTTACCTAAATGTACTCCACCAACTACTAAAGTTTTCACATCG	Mutagenesis
D475K-R	CCTTACCTAAATGTTTTCCACCTTTTACTAAAGTTTTCACATCG	Mutagenesis
KGGD-R	CCTTACCTAAATGATCTCCACCTTTTACTAAAGTTTTCACATCG	Mutagenesis
pglB-C1-F	/5Phos/ATAATTTTTTCCCTTCTTTTGCTTTAAGCAAAGATG	Mutagenesis
pglBHA*Kpn*I-R	ATATACTCGAG*GTAC*CATGGTTAAGCGTAATCTGGAACATCG, *Kpn*I	Mutagenesis
pglB*Bam*HI-F	ATTAGC*GGATC*CTACCTGACGCTTTTTATCGC, *Bam*HI	Mutagenesis PCR, cloning
pglB*Xho*I-R	TATA*CTCGAG*TTAAGCGTAATCTGGAACATCGTATGG, *Xho*I	PCR, cloning
CCV52592_1213-NcoI-F	ATAA*CCATGG*ATAAAGAGGGTTTTAGCTCATATTTTAAAAATTTTCACCTCTCAAAGC, *Nco*I	PCR, cloning
CCV52592_1213-HindIII-R	TA*AAGCTT*CAAGCGTAATCTGGAACATCGTATGGGTACCTCTTTAGTCTATAAACCTTTG, *Hind*III	PCR, cloning
CCV52592_1221-NcoI-F	AAATC*CATGG*ATAGTCAAGCCTTAAATTCAAATTTTCGCTCATCGAGGCAAATTTTAGC, *Nco*I	PCR, cloning
CCV52592_1221-HindIII-R	AAA*AGCTT*CAAGCGTAATCTGGAACATCGTATGGGTACTTTAGCAACCTGTAAATTTTAG, *Hind*III	PCR, cloning
Cg0434-F-NcoI	ATAAAC*CATGG*AAAAATTTAGAATTACAAACGGCGAATGCTCACGTATGG, *Nco*I	PCR, cloning
Cg0434-R-PstI	ATC*TGCAG*TCAAGCGTAATCTGGAACATCGTATGGGTATCTTAATAACTTGTAAATCG, *Pst*I	PCR, cloning
Cg2013-F-NcoI	ATAAAC*CATGG*GGTGAAGAAAAAAAGGCTTTTAGCGTCAGAGA GATCGACG, *Nco*I	PCR, cloning
1221_DGGK-F-P	/5Phos/GATGGTGGCAAGCAATCAGGCACATACAGCTTTCTAACGAGC	PCR, cloning Mutagenesis
1221_DGGK-R-P	/5Phos/GCTCAGCGTTTTGACATCGGCATAATACTTGATC	PCR, cloning Mutagenesis PCR, cloning
1213_DGPR-F-P	/5Phos/GATCCGGGGAGGCACCTTGGACGTGATAACTACGCCGTTAGC	Mutagenesis PCR, cloning
1213_DGPR-R-P	/5Phos/TATTAAGGTCTTGACATCGGCGTAATATCGTATCGGATAGC	Mutagenesis PCR, cloning

Underlined, mutated nucleotides within *pglB*; nucleotides in italics indicate the presence of restriction enzyme sites; /5Phos/indicates phosphorylation of the 5′ end.

5.4.3 Expression of PglB proteins in *E. coli* and purification of *Cj*-CmeA-His$_6$

Escherichia coli CLM24 expressing the *C. jejuni* protein glycosylation locus from ACYC184(*pglmut*), the soluble form of the *C. jejuni* glycoprotein *Cj*-CmeA-His$_6$, and each of the *pglB* alleles from the pMLBAD derivative were grown at 37°C to an OD_{600} of 0.5–0.6. After induction with 0.2% L-arabinose (wt/vol) for 4 h, cells were harvested by centrifugation (3696 × g, 15 min, 4°C), and washed twice with chilled phosphate-buffered saline (PBS) pH 7.2. Whole cell lysates were prepared by re-suspending the cells in 1/50 of the original culture volume of PBS, pH 7.2 followed by sonication (3 × 1 min) with a Branson sonicator equipped

with a microtip. After removal of unbroken cells by low-spin centrifugation (16,260 × g, 20 min, 4°C) PglB proteins in the supernatant were visualized by western blotting using HA-tag specific antibodies.

Cj-CmeA was purified from the same whole cell lysates by nickel affinity chromatography (NTA agarose, Qiagen) as follows. After equilibration of the resin with PBS, whole cell lysates containing *Cj*-CmeA-His_6 protein were loaded. The resin was then washed with 10 column volumes of 1 × PBS with 20mM imidazole to remove unspecific bound protein. *Cj*-CmeA proteins were eluted with 1 × PBS with 500mM imidazole, dialyzed against 1 × PBS and stored at 4°C. The glycosylation status of purified *Cj*-CmeA-His_6 protein was analyzed by western blotting using *Cj*-CmeA-specific and *C. jejuni*-N-glycan-specific antibodies.

5.4.4 Preparation of whole cell lysates and western blotting

Whole cell lysates of *C. jejuni* NCTC11168 were prepared as described [40]. The overall glycosylation profile in whole cell lysates of the *C. jejuni* wildtype, the *C. jejuni pglB* mutant and the *C. jejuni pglB* mutant expressing various *pglB* alleles was analyzed by western blotting after separation on 12.5% protein gels using *C. jejuni*-N-glycan-specific (R1) antibodies [29]. The expression and glycosylation status of proteins was analyzed after separation by 10% SDS-PAGE and transfer onto PVDF membranes as described previously [40]. Briefly, HA-Tag and His-Tag-specific (1:1000) (Santa Cruz) or *C. jejuni*-Nglycan-specific (R1) (1:10,000) polyclonal sera served as the primary antibody, alkaline phosphatase conjugated anti-rabbit or anti-mouse IgG (Santa Cruz) served as the secondary antibodies at a 1:2000 dilution. Immuno-reactive bands were visualized using the NBTBCIP development reagents (Promega) according to the instructions of the manufacturer.

5.4.5 Densitometry analysis

Western blotting was quantified by densitometry. The dry PVDF membrane was scanned at 600 dpi, 16 bit, followed by quantification of the band intensity using ImageQuant version 5.1 software (Molecular Dynamics). Results were expressed as percentage relative to wildtype signals.

5.4.6 Semi-quantitative MS-based fOS analysis

Free glycans (fOS) from the *C. jejuni* NCTC 11168 *pglB* mutants were prepared and analyzed as described previously (Nothaft et al. 2009). Sample preparation and fOS analysis were repeated in triplicate and the average relative quantities were used.

5.4.7 Homology modeling and docking

Homology models of PglB from *C. jejuni*, *C. curvus*, *C. gracilis* and *C. concisus* were prepared using the crystal structure of *C. lari* PglB in complex with acceptor peptide (PDB 3RCE) [27] as the template. The Swiss-Model server was used to generate sequence alignments and to generate the homology models [41]. A model of the undecaprenyl-diphosphate glycan donor was constructed by modeling the structure of the glycan using the GLYCAM-WEB carbohydrate-builder website [42] and building stereochemically reasonable structures for the undecaprenol lipid and pyrophosphate moieties. The donor glycan lipid was manually docked to align the pyrophosphate group with the divalent metal ion at the active site and to position the anomeric carbon atom of the diNAcBac residue within ~2.5 Å of the side chain amide nitrogen atom of the acceptor peptide asparagine residue in the *C. lari* PglB complex crystal structure.

5.4.8 Limited proteolysis

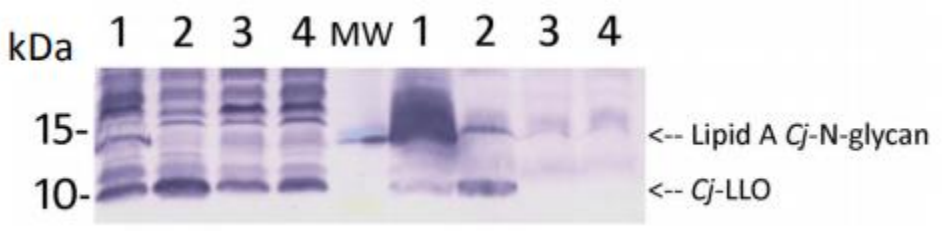
Cell envelope fractions (CEFs) proteins were prepared from whole cell lysates containing PglB (grown and expressed as described above) according to [31]. *Cj*-LLOs were prepared from *E. coli* K12 (*Δwzy/waaL*::kan) expressing the *Cj-pgl* locus (Supplementary data, Figure S1). Then, 160 µL of CEF containing PglB proteins were first incubated with heat treated *Cj*-LLO for 30 min on ice followed by addition of 2 µL proteinase K (0.1mg/mL)
and 12.5 µL $MgSO_4$ (50mM) and incubation at room temperature. Aliquots of 20 µL were taken at different time points, immediately mixed with 5 µL of 5 × protein inhibitor cocktail mix (Roche) and 6 µL of 6 × SDS loading dye and frozen at −80°C until further use. Proteolytic profiles were analyzed by western blotting using the HA-Tag antibody as described above, after separation of the reaction mixtures by 12.5% SDS-PAGE.

5.5 Supplementary data

Supplementary Methods and Results Expression and preparation of *C. jejuni* lipid-linked oligosaccharides (*Cj*-LLOs) in *E. coli* The heterologous Cj-LLO expression strain was constructed by eliminating the kanamycin cassette from the *E. coli* K12 *wzy::kan* mutant (KEIO collection (Baba et al. 2006)) according to Datsenko et al (Datsenko and Wanner 2000). The generated Δ*wzy* strain was used to introduce the kanamycin cassette into the O-antigen ligase (*waaL*) locus. To do so, a 2.5 kbp *waaL::kan* PCR fragment was generated with oligonucleotides rfaL_out_1 (5'- AGAATTGATTGGTTTACTGGCTGATTCAGG-3') and rfaL_out_2 (5'-TAAATTTTCCACCAACTACGACTTGCATTTACC-3') from chromosomal DNA of *E. coli waaL::kan* (KEIO collection (Baba et al. 2006)) and introduced into the *Δwzy* strain as described (Datsenko and Wanner 2000). Expression of *Cj*-LLOs after introduction of the *pgl* operon into *E. coli Δwzy/waaL::kan* was followed by Western Blotting with *Cj*-N-glycan-specific antiserum (R1) in whole cell lysates of the respective strain after treatment with proteinase K for 24 hr at 37°C and heat inactivation for 15 min at 95°C (Supplementary Figure S1). A *Cj*-N-glycan positive signal corresponding to *Cj*-LLOs migrates at approximately 10 kDa after separation on a 15% SDS gel. This signal is absent in *E. coli Δwzy* and *E. coli waaL::kan* (negative controls). The N-glycan specific signal in the *Δwzy/Cj-pgl* strain (used as a positive control) results from the transfer of the *Cj*-N-glycan to the *E. coli* lipid A core that migrates at approximately 15 kDa (Nothaft et al. 2016). LLOs were heat treated for 15 min at 95°C and stored at -20°C until further use.

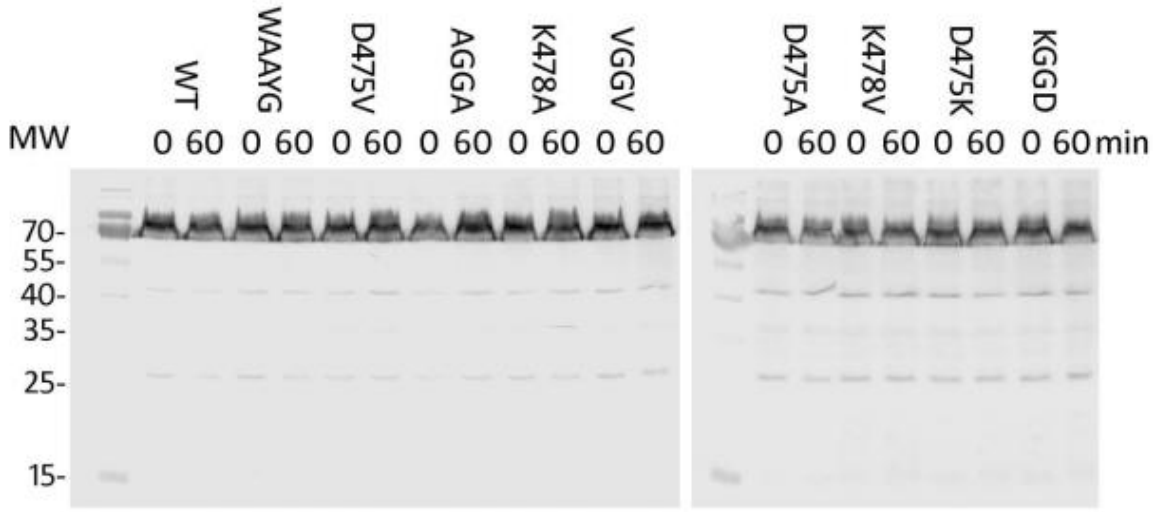
Supplementary Figures

Supplementary Figure S1



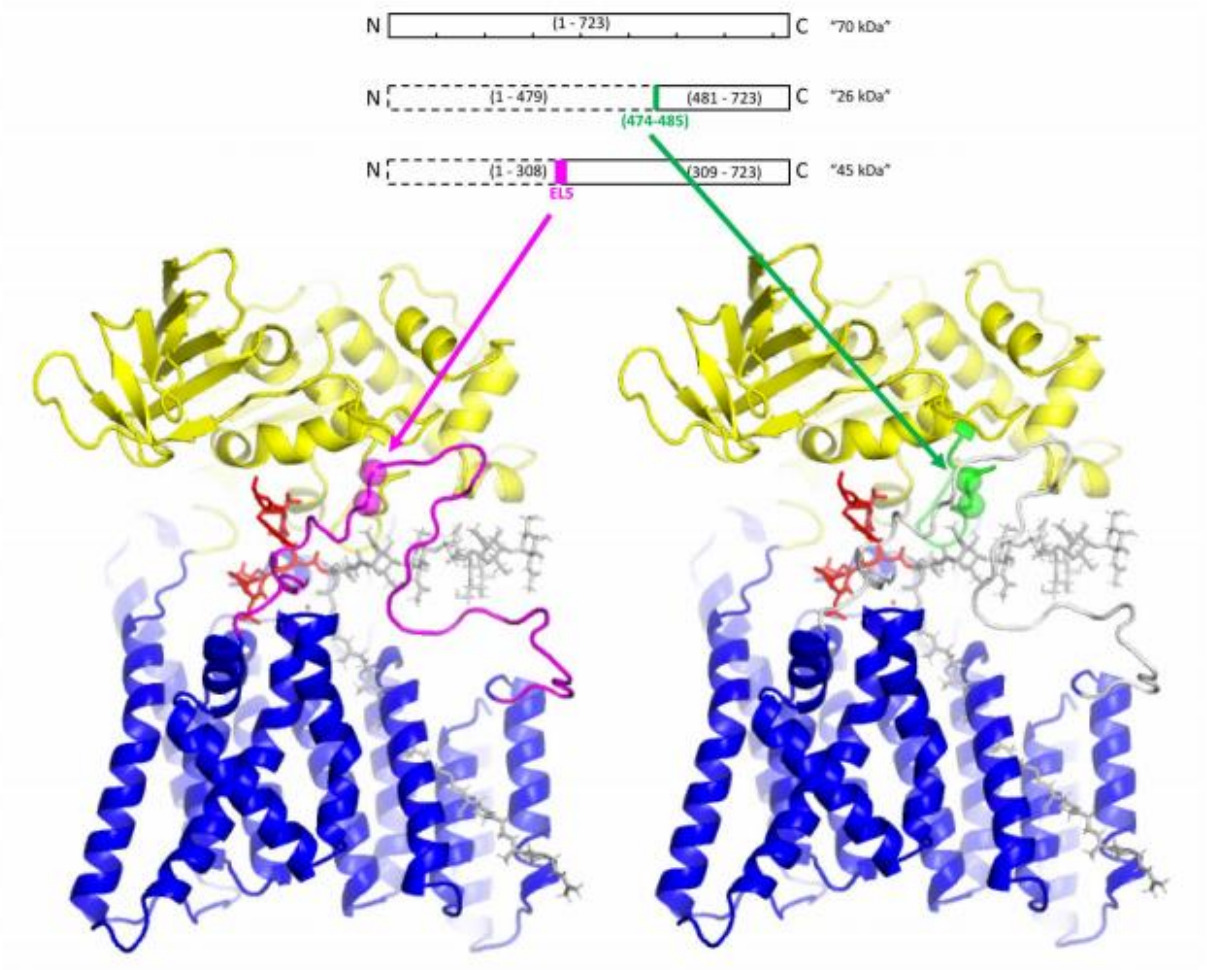
Supplementary Figure A.S1. *Cj*-LLO preparation and analysis. The formation of *Cj*-LLOs verified by Western blotting with *Cj*-N-glycan-specific antiserum in *E. coli* whole cell lysates before (left panel) and after proteinase K treatment and heat inactivation (right panel). In *E. coli* Δ*wzy/waaL::kan* that expresses the *Cj*-pgl operon (lane 2) *Cj*-LLO migrates at ~10 kDa, indicated by an arrow. Lipid A N-glycan formation in *E. coli wzy::kan* expressing the *Cj*-pgl locus was used as a positive control for antiserum reactivity (lane 1). Lipid A-N-glycan compounds migrate at ~15 kDa, indicated by an arrow. No *Cj*-LLOs were formed in the absence of the *Cj*-pgl genes in strains *E. coli waaL::kan* (lane 3) and *E. coli wzy::kan* (lane 4). Molecular weight markers (in kDa) are indicated on the left. MW indicates the markers used to separate the two portions of the SDS-PAGE gel.

Supplementary Figure S2



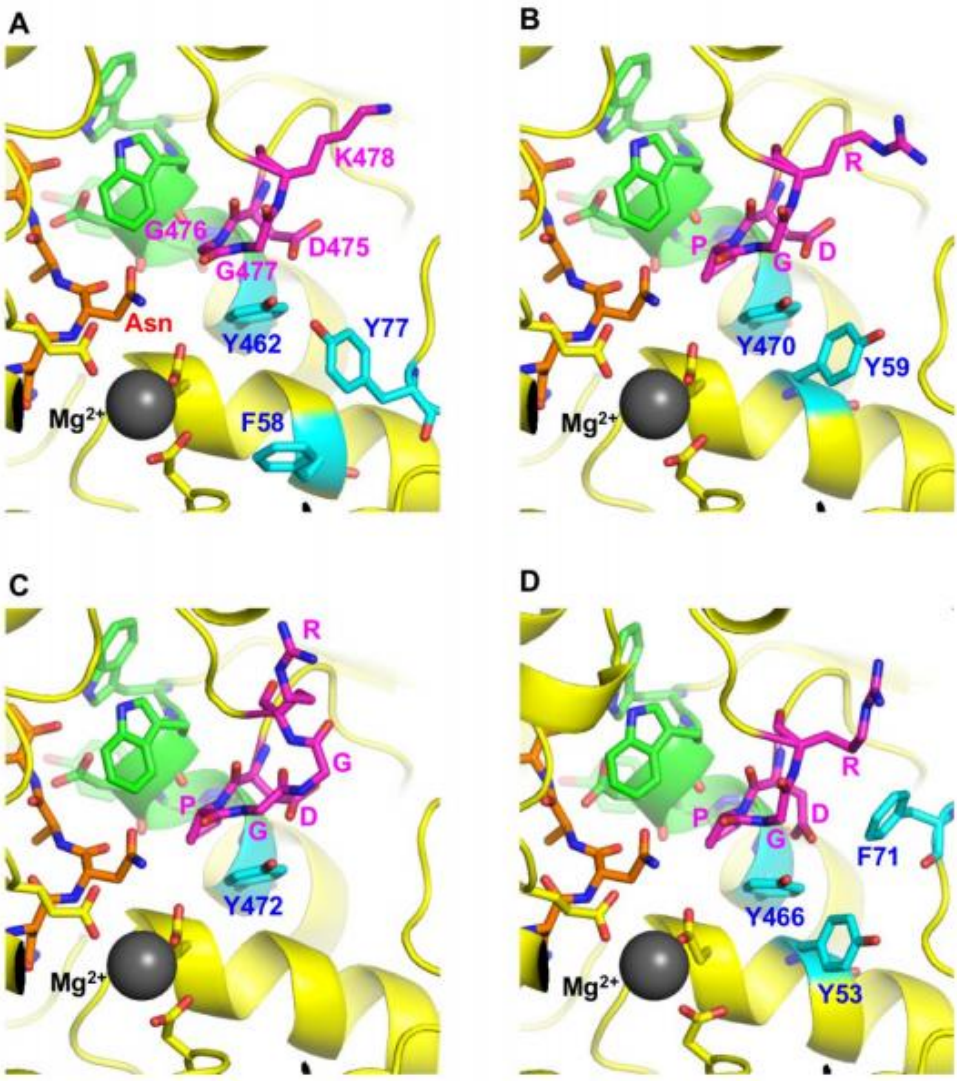
Supplementary Figure A.S2. PglB stability in the absence of proteinase K. Detection of PglB proteins (as indicated) by Western blotting with HA-specific antibodies at t=0 and after t=60 min incubation under limited proteolysis assay conditions, but in the absence of proteinase K. PglB migrated at 70 kDa. Molecular weight markers (in kDa) are indicated on the left.

Supplementary Figure S3



Supplementary Figure A.S3. Putative sites of proteolysis in PglB. Spheres indicate the alpha-carbon positions of residues F308 and N309 near the C-terminal end of EL5 (magenta) and L480 and G481 near the C-terminal end of the 474-485 loop (green). The N-terminal, transmembrane domain of PglB (blue), C-terminal soluble domain (yellow), acceptor peptide (red) and LLO donor (gray) are shown as in the same orientation as in Figure 1.

Supplementary Figure S4



Supplementary Figure A.S4. Homology models of PglB orthologues. Close-up views are drawn from the same orientation as shown in Fig. 1B for the active-site regions of (A) the crystal structure of the *C. jejuni* PglB peptide complex (PDB 3RCE), and homology models of PglB orthologues from (B) *C. curvus* 1221, (C) *C. gracilis* 0434, and (D) *C. concisus* 0460. In panel (A), the residue numbering follows that for *C. jejuni* PglB, even though the structure that is shown is from the closely related PglB homologue from *C. lari*.

5.6 References:

1. Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta (BBA)-General Subjects*. 1999;1426(2):239-257.

2. Helenius J, Ng DT, Marolda CL, Walter P, Valvano MA, Aebi M. Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature*. 2002;415(6870):447.

3. Marshall RD. Glycoproteins. *Annu Rev Biochem*. 1972;41:673-702.

4. Yan Q, Lennarz WJ. Oligosaccharyltransferase: A complex multisubunit enzyme of the endoplasmic reticulum. *Biochem Biophys Res Commun*. 1999;266(3):684-689.

5. Kelleher DJ, Gilmore R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*. 2005;16(4):62R.

6. Berriman M, Ghedin E, Hertz-Fowler C, et al. The genome of the african trypanosome *Trypanosoma brucei*. *Science*. 2005;309(5733):416-422.

7. Ivens AC, Peacock CS, Worthey EA, et al. The genome of the kinetoplastid parasite, *Leishmania major*. *Science*. 2005;309(5733):436-442.

8. Mescher MF, Strominger JL. Purification and characterization of a prokaryotic glucoprotein from the cell envelope of halobacterium salinarium. *J Biol Chem*. 1976;251(7):2005-2014.

9. Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol Microbiol*. 1999;32(5):1022-1030.

10. Igura M, Maita N, Kamishikiryo J, et al. Structure guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J*. 2008;27(1):234-243.

11. Silverman JM, Imperiali B. Bacterial N-glycosylation efficiency is dependent on the structural context of target sequons. *J Biol Chem*. 2016;291(42):22001-22010.

12. Alaimo C, Catrein I, Morf L, et al. Two distinct but interchangeable mechanisms for flipping of lipid linked oligosaccharides. *EMBO J*. 2006;25(5):967-976.

13. Linton D, Dorrell N, Hitchen PG, et al. Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. *Mol Microbiol*. 2005;55(6):1695-1703.

14. Kowarik M, Young NM, Numao S, et al. Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J*. 2006;25(9):1957-1966.

15. Wacker M, Linton D, Hitchen PG, et al. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science*. 2002;298(5599):1790-1793.

16. Nothaft H, Liu X, McNally DJ, Li J, Szymanski CM. Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proc Natl Acad Sci U S A*. 2009;106(35):15019-15024.

17. Jervis AJ, Butler JA, Lawson AJ, Langdon R, Wren BW, Linton D. Characterization of the structurally diverse N-linked glycans of *Campylobacter* species. *J Bacteriol*. 2012;194(9):2355-2362.

18. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *Campylobacter* genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219.

19. Ielmini MV, Feldman MF. *Desulfovibrio desulfuricans* PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements. *Glycobiology*. 2010;21(6):734-742.

20. Baar C, Eppinger M, Raddatz G, et al. Complete genome sequence and analysis of *Wolinella succinogenes*. *Proceedings of the National Academy of Sciences*. 2003;100(20):11690-11695.

21. Nakagawa S, Takaki Y, Shimamura S, Reysenbach AL, Takai K, Horikoshi K. Deep-sea vent epsilon-proteobacterial genomes provide insights into emergence of pathogens. *Proc Natl Acad Sci U S A*. 2007;104(29):12146-12150.

22. Jervis AJ, Langdon R, Hitchen P, et al. Characterization of N-linked protein glycosylation in *Helicobacter pullorum*. *J Bacteriol*. 2010;192(19):5228-5236.

23. Mills DC, Jervis AJ, Abouelhadid S, et al. Functional analysis of N-linking oligosaccharyl transferase enzymes encoded by deep-sea vent proteobacteria. *Glycobiology*. 2016;26(4):398-409.

24. Ollis AA, Chai Y, Natarajan A, et al. Substitute sweeteners: Diverse bacterial oligosaccharyltransferases with unique N-glycosylation site preferences. *Sci Rep*. 2015;5:15237.

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

26. Li L, Woodward R, Ding Y, et al. Overexpression and topology of bacterial oligosaccharyltransferase PglB. *Biochem Biophys Res Commun*. 2010;394(4):1069-1074.

27. Lizak C, Gerber S, Numao S, Aebi M, Locher KP. X-ray structure of a bacterial oligosaccharyltransferase. *Nature*. 2011;474(7351):350-355.

28. Maita N, Nyirenda J, Igura M, Kamishikiryo J, Kohda D. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J Biol Chem*. 2010;285(7):4941-4950.

29. Nothaft H, Scott NE, Vinogradov E, et al. Diversity in the protein N-glycosylation pathways within the *Campylobacter* genus. *Mol Cell Proteomics*. 2012;11(11):1203-1219.

30. Wacker M, Linton D, Hitchen PG, et al. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science*. 2002;298(5599):1790-1793.

31. Jaffee MB, Imperiali B. Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound N-linked oligosaccharyl transferases. *Biochemistry (N Y)*. 2011;50(35):7557-7567.

32. Ihssen J, Kowarik M, Wiesli L, Reiss R, Wacker M, Thony-Meyer L. Structural insights from random mutagenesis of *Campylobacter jejuni* oligosaccharyltransferase PglB. *BMC biotechnology*. 2012;12(1):67.

33. Kämpf MM, Braun M, Sirena D, Ihssen J, Thöny-Meyer L, Ren Q. *In vivo* production of a novel glycoconjugate vaccine against *Shigella flexneri* 2a in recombinant *Escherichia coli*: Identification of stimulating factors for *in vivo* glycosylation. *Microbial cell factories*. 2015;14(1):12.

34. Ihssen J, Haas J, Kowarik M, et al. Increased efficiency of *Campylobacter jejuni* N-oligosaccharyltransferase PglB by structure-guided engineering. *Open biology*. 2015;5(4):140227.

35. Sun Lee H, Im W. Transmembrane motions of PglB induced by LLO are coupled with EL5 loop conformational changes necessary for OST activity. *Glycobiology*. 2017.

36. Lizak C, Gerber S, Zinne D, et al. A catalytically essential motif in external loop 5 of the bacterial oligosaccharyltransferase PglB. *J Biol Chem*. 2014;289(2):735-746.

37. Ollis AA, Zhang S, Fisher AC, DeLisa MP. Engineered oligosaccharyltransferases with greatly relaxed acceptor-site specificity. *Nat Chem Biol*. 2014;10(10):816-822.

38. Kato Y, Arakawa N, Masuishi Y, Kawasaki H, Hirano H. Mutagenesis of longer inserts by the ligation of two PCR fragments amplified with a mutation primer. *Journal of bioscience and bioengineering*. 2009;107(1):95-97.

39. Labigne-Roussel A, Harel J, Tompkins L. Gene transfer from *Escherichia coli* to *campylobacter* species: Development of shuttle vectors for genetic analysis of *Campylobacter jejuni*. *J Bacteriol*. 1987;169(11):5320-5323.

40. Nothaft H, Liu X, McNally DJ, Szymanski CM. N-linked protein glycosylation in a bacterial system. *Methods Mol Biol*. 2010;600:227-243.

41. Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014;42(W1):W258.

42. Kirschner KN, Yongye AB, Tschampel SM, Daniels CR, Foley BL, Woods RJ. GLYCAM06: A generalizable biomolecular force field. carbohydrates. *Journal of computational chemistry*. 2008;29(4):622-655.