Polyprenyl-dependent glycan assembly pathways in microbial pathogens

by

Meredith Diane Hartley

A.B. Biophysical Chemistry
Dartmouth College, 2005

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 07 2011
LIBRARIES

ARCHIVES

Submitted to the Department of Chemistry in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy at the

Massachusetts Institute of Technology

June 2011

© 2011 Massachusetts Institute of Technology
All rights reserved

Signature of Author _____

Department of Chemistry
May 18, 2011

Certified by _____

Barbara Imperiali
Class of 1922 Professor of Chemistry and Professor of Biology
Thesis Supervisor

Accepted by _____

Robert W. Field
Haslam and Dewey Professor of Chemistry
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Professor Joanne Stubbe ____ _____
Thesis Committee Chair

Professor Barbara Imperiali _ ___
Thesis Supervisor

Professor John Essigmann _ _____

Polyprenyl-dependent glycan assembly pathways in microbial pathogens

by

Meredith Diane Hartley

Submitted to the Department of Chemistry
on May 18, 2011 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Biological Chemistry

ABSTRACT

Polyisoprenyl-dependent glycan assembly pathways form the basis for the biosynthesis of many complex glycoconjugates. This thesis addresses key aspects of undecaprenyl-phosphate related processes; undecaprenol is the linear polyisoprenol that is utilized by virtually all known bacterial species. In the first chapter, a previously identified undecaprenol kinase from *Streptococcus mutans* is adapted for enzymatic phosphorylation of polyprenols as an alternate method to chemical phosphorylation. This chemoenzymatic method is shown to be particularly efficient in the biosynthesis of undecaprenyl-diphosphate linked glycans, since the kinase can be coupled to various glycosyltransferases to generate the complex products starting from undecaprenol and ATP.

A second focus of this thesis involves the biochemical characterization of the undecaprenyl-dependent O-linked protein glycosylation pathway in *Neisseria gonorrhoeae*. The *N. gonorrhoeae* pathway is shown to produce UDP-N,N'-diacetylbacillosamine, which is the UDP-sugar donor in the first membrane-associated step. Furthermore, it is demonstrated that glycosyltransferases in *N. gonorrhoeae* assemble glycans on undecaprenyl-diphosphate prior to transfer to hydroxyl side chains of serine and threonine residues. The final glycan transfer is performed by an oligosaccharyltransferase (OTase), and specificity studies of O-linked and N-linked bacterial OTases suggest that both enzymes prefer their native glycans under in vitro assay conditions. This work represents the first biochemical characterization of a polyprenyl-dependent pathway that produces O-linked glycoprotein in prokaryotes.

The final three chapters present the foundations for a new experimental approach to the study of polyprenyl-dependent pathways by employing a model membrane system, termed Nanodiscs, which comprises a discoidal phospholipid bilayer encircled by a scaffold protein. In this thesis, the glycosyltransferases responsible for the first two committed membrane steps in the *Campylobacter jejuni* N-linked protein glycosylation pathway are co-incorporated into Nanodiscs. Importantly, radioactivity-based assays using a dual isotope-labeling strategy demonstrate the functional reconstitution of both proteins. In addition, efforts are described toward the development of FRET- and LRET-based methodology for future characterization of the protein-protein interactions between PglC and PglA. This work establishes a basis for future studies of polyprenols in the *C. jejuni* glycan assembly process and may be applicable to other essential polyprenyl-dependent processes.

Thesis supervisor: Barbara Imperiali
Title: Class of 1922 Professor of Chemistry and Professor of Biology

Acknowledgements

These acknowledgements must begin with my gratitude towards my advisor Barbara Imperiali. She has been a constant source of inspiration, guidance and support and I feel honored that she gave me the opportunity to work in her laboratory. She has been a great role model for me as an educator and a scientist. In addition, her enthusiasm for all things related to glycosylation has been infectious and I must thank her for introducing me to the complex and challenging world of polyprenyl-dependent protein glycosylation. I also am indebted to her for her tireless editing of every chapter of this thesis. Her input has been invaluable. Finally, I am very grateful that she allowed me the freedom to make mistakes, which has ultimately made me a better and more independent scientist.

In addition, I would like to thank the other professors who have inspired me during my years at MIT. I am very glad that Joanne Stubbe was my thesis chair, because her dedication to rigorous science has been both inspiring and motivating. I am very grateful for her enthusiasm about science and her constant interest in my project, even outside of thesis chair meetings. In addition, I would like to thank Sarah O'Connor and John Essigmann, who have provided support and encouragement during my time in graduate school. Along the same lines, I would like to acknowledge Elizabeth Fong, who holds our lab together in many crucial ways and who always has a smile on her face.

I am excited to have this opportunity to acknowledge the many fantastic colleagues that I have had the pleasure of working with in the Imperiali lab. It starts at the beginning and I am grateful to the enthusiastic, friendly and helpful graduate students and postdocs who welcomed me into the lab. I am thankful that I was able to overlap with so many wonderful people including Bianca, Anne, Dora, Jebrell, Seungjib, Eranthie, Mary, Beth, Elvedin, and Galen. They all played a role in making the Imperiali lab a special place and I have learned from each one of them. I am also indebted to Langdon, Mark, and Guofeng, who whole-heartedly welcomed me into the pod and did more than their share of helping me feel like a part of the lab. In addition, I am thankful for the friendship of Matthieu, who taught me everything I know about volleyball, and whose work ethic I can only admire. Finally, I would be remiss if I did not acknowledge the support and encouragement of Lieutenant-Commander Nelson, who always wanted me to pour my gels faster. He has never failed to surprise me, always has something new to teach me about science or life, and has consistently inspired me with his dogged pursuit of doing or learning anything he wants. I am grateful for both his mentorship and friendship.

Many colleagues have joined the lab since the early days and they have shaped the lab in new and amazing ways. Even though they were on Bioprobes, I am grateful for the opportunity to have known and worked with many of them including Juan Antonio, Melanie, Susana, Katja, Will, Andrew, and Elke. On the Glyco side of the lab, Jay and James were fantastic postdocs, and I am so grateful to have worked and played volleyball alongside them both. In addition, I am thankful for the technical advice that Jay has shared with me during our many benchside conversations and over the past year from his new lab.

As I graduate, I am leaving an excellent crew of Glyco researchers and I am really excited to see what they produce. I am really glad that Michelle joined the lab and I thank her for many wonderful discussions about science and life. I know that she is going to do an excellent job as volleyball co-captain. Marcie has been a wonderful collaborator on a variety of ambitious cloning project and I am grateful for the opportunity to know her and work with her. Mike has

also been a great collaborator and is one of the nicest guys I know, and I thank him for the many ways in which he has helped me out. It is great to have a new postdoc on Glyco and I am thankful that Philipp has taught me so many things about lipids. I wish him all the best as he carries Nanodiscs into the future! In addition, I am glad that Austin has joined the lab, and I wish him the best in graduate school and life.

Importantly, I want to thank the members of my awesome and dynamic pod, Cliff, Vinita, and Thaís, who have been extremely supportive during the many stressful moments of the past few years. Cliff has been a great friend and I am so excited for his new adventure with Marilyne as Professors Stains and Stains. In addition, I am thankful for the boundless optimism and constant cheer that Thaís has brought to my life. I don't know what she will do in the future, but I know that it will be amazing. Finally, I was honored and delighted to mentor Vinita (LukNAz), and I look forward to a lifelong friendship with her. I thank her for insightful advice on a number of key decisions and for teaching me more than she realizes.

I have had the honor of attending graduate school with many other wonderful graduate students not in the Imperiali lab and I am thankful for the lasting friendships I have made. I am grateful to Christy and Sarah, who have been supportive and wonderful friends and who always reminded me to be appreciative of the Imperiali lab. I am also thankful for the friendship of Laura, Pam, and Alisha, who have always had time for essential non-lab activities that have helped make the years of graduate school more enjoyable.

I owe my largest thanks to my three fellow grad students who arrived in the Imperiali lab not before or after me, but with me. Angelyn, Brenda and Wendy have been my classmates, colleagues, labmates, confidants, friends and most importantly my lab family. Together, we have experienced every high and low of graduate school and life over the past six years. I am grateful for their presence in my life, as I know it has made each day of graduate school a little bit better. I will miss them a great deal as we all take the next step. I am thankful to Wendy for her constant encouragement and support and for always being willing to listen. Angelyn has been an incredible scientific collaborator and her dedication to her science has always inspired me. More importantly, she has been a kind and supportive friend and I am appreciative of the many ways in which she has been there for me through everything. Finally, I am grateful for the friendship of Brenda, who is always willing to eat lunch, go for a snack, talk through an experiment, commiserate over lab failures/celebrate triumphs, or discuss any random topic that comes to mind. I am thankful for everything she has taught me and for the daily (and hourly) support and encouragement she has provided.

Finally, these acknowledgements would be meaningless if I did not thank my mom and Jeremiah. Their support has been endless, and I know that is how graduate school has seemed at times. I am thankful to my mom (the first Dr. Hartley) for always letting me make my own decisions, and for always believing in what I choose to do. I am grateful to Jeremiah for grounding me in reality and for always reminding what is important. I would be lost without his presence in my life.

I would like to dedicate this thesis to my dad, who inspires me on a daily basis in ways he will never know.

Table of Contents

Abstract 3

Acknowledgements 4

Table of Contents 6

List of Figures 10

List of Tables 13

List of Abbreviations 14

Chapter 1: Examining the role of polyisoprenols and polyisoprenyl-phosphates in biology 16

Introduction 17
Polyprenol and polyprenyl-phosphate abundance in eukaryotic membranes 25
Undecaprenol and undecaprenyl-phosphate abundance in microbial membranes 29
Why polyprenols? Regulation of polyprenyl-dependent pathways 32
Why polyprenols? Importance in glycan "flipping" 37
Why polyprenol? Models of macromolecular complexes 44
Conclusions 48
Acknowledgements 49
References 50

Chapter 2: Chemoenzymatic synthesis of polyprenyl-phosphates and polyprenyl-phosphate derivatives 58

Introduction 59
Results and Discussion 64
Substrate specificity of *C. jejuni* BacA, *S. mutans* DGK, and *C. jejuni* DGK 64
Development of a one-pot enzymatic route to undecaprenyl-diphosphate disaccharide 66
Quantification of kinase enzymes by luminescence 69
Optimization of Und-PP-diNAcBac-[^{3}H]-GalNAc synthesis for assays of PglB 72
General utilization of *S. mutans* DGK 74
Conclusions 76
Acknowledgements 77
Experimental Procedures 78
Materials 78
Cloning of kinase genes into expression vector 78
Expression of enzymes in *E. coli* BL21 cells 79
Isolation of kinases in cell envelope fractions 79
Assay of kinases with [^{32}P]-ATP 80
Purification of kinases from cell envelope fraction 80
Purification of polyprenyl-phosphates by normal phase HPLC 81
Synthesis of radiolabeled undecaprenyl-diphosphate disaccharide 81
Quantification of the dLBT-kinase in a cell envelope fraction using luminescence 82
References 84

Purification of polyprenyl-phosphates by normal phase HPLC 81
Synthesis of radiolabeled undecaprenyl-diphosphate disaccharide 81
Quantification of the dLBT-kinase in a cell envelope fraction using luminescence 82
References **84**

Chapter 3: Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of undecaprenyl-diphosphate-linked glycan intermediates **87**

Introduction **88**
Results and Discussion **92**
Determination of UDP-DATDH stereochemistry by NMR 92
Characterization of PglB-ATD and PglC 95
Biochemical characterization of the *N. gonorrhoeae* glycosyltransferases 97
Characterization of PglH, an alternative glycosyltransferase 104
Significant glycan diversity in *N. gonorrhoeae* 107
Conclusions **109**
Acknowledgements **109**
Experimental Procedures **111**
Materials 111
Expression of proteins 112
Protein purification 112
Acetyltransferase (PglB-ATD) Activity Assay 115
Aminotransferase (PglC) Activity Assay 116
Biosynthesis and stereochemical assignment of UDP-DATDH 116
Preparation of radiolabeled Und-PP-linked substrates 117
Normal phase HPLC purification of Und-PP-linked substrates 119
Preparation and analysis of 2-AB labeled oligosaccharides 119
Glycosyltransferase substrate specificity assays 120
References **122**

Chapter 4: Characterization of the O-linked oligosaccharyltransferase from the protein glycosylation pathway in *Neisseria gonorrhoeae* **125**

Introduction **126**
Results and Discussion **130**
Functional characterization of oligosaccharyltransferase, PglO 130
Glycan donor specificity of PglO(*Ng*) 133
Pilin glycosylation by PglB(*Cj*) and PglO(*Ng*) 136
Protein substrate specificity of PglO(*Ng*) 137
Conclusions **141**
Acknowledgements **142**
Experimental Procedures **143**
Preparation of DNA constructs 143
Expression of proteins 143
Protein purification 144
Oligosaccharyltransferase assays 145
References **148**

Chapter 5: Adaptation of Nanodisc technology to the study of the N-linked protein glycosylation pathway in *Campylobacter jejuni* **150**

Introduction **151**
Results and Discussion **157**
Purification of PglC 157
Purification of PglA 162
Incorporation of PglC into Nanodiscs 163
PglA-containing Nanodiscs 169
Interaction of PglC and PglA in Nanodiscs 170
Conclusions **171**
Acknowledgements **172**
Experimental Procedures **172**
Materials 172
DNA constructs 173
Protein expression 173
Purification of PglC 174
Purification of PglA 175
PglA TEV cleavage 176
Purification of MSP 177
Lipid preparation and quantification 177
Incorporation of PglC into Nanodiscs 179
Gel densitometry analysis of Nanodiscs 181
References **182**

Chapter 6: Biophysical approaches to studying protein-protein interactions in Nanodiscs **185**

Introduction **186**
Results and Discussion **190**
Fluorescent labeling of PglC 190
Fluorescent labeling of PglA 194
Generation of PglA-LBT constructs 199
Preparation of Nanodiscs with LRET partners 200
Use of Amphipols to solubilize PglA 203
Conclusions **205**
Acknowledgements **206**
Experimental Procedures **207**
Preparation of DNA constructs encoding PglC Cys mutants, LBT-PglA, and PglA-LBT 207
DTNB and iodoacetamide labeling of PglA 208
Fluorescent labeling of PglC and PglA 210
Solubilization of PglA with amphipols 210
References **211**

Chapter 7: Studies towards the functional reconstitution of the PglC and PglA glycosyltransferases into Nanodiscs and incorporation of undecaprenyl-phosphate-linked substrates **213**

Introduction **214**
Results and discussion **218**
Preliminary studies on the incorporation of undecaprenyl-linked substrates 218
Preliminary activity assays of the PglA glycosyltransferase 222
Preliminary examination of PglC function in Nanodiscs 225
Nanodiscs prepared with Und-^{33}P 227

Conclusions **231**
Acknowledgements **232**
Experimental Protocols **233**
Materials 233
Incorporation of undecaprenyl-linked substrates into Nanodiscs 233
Activity analysis of Nanodisc samples 234
References **236**

Appendix **237**

Attempts towards the purification of PglC for X-ray crystallography 237
^{1}H NMR spectrum of UDP-diNAcBac. 239
^{31}P NMR spectrum of UDP-diNAcBac. 240
^{13}C NMR spectrum of UDP-diNAcBac. 241
^{1}H-^{1}H COSY NMR spectrum of UDP-diNAcBac. 242

List of Figures

Chapter 1

Figure 1-1. Structures of fully unsaturated polyprenols and dolichols.....17
Figure 1-2. Mechanism of isoprene biosynthesis. 19
Figure 1-3. Dolichyl-linked glycan substrates in eukaryotes and archaea. 21
Figure 1-4. Undecaprenyl-linked glycans in bacterial plasma membranes. 22
Figure 1-5. Phospholipid structures in presence of polyprenols.....29
Figure 1-6. Cellular sources of dolichyl-phosphate.....34
Figure 1-7. Structures of dolichol, dolichyl-phosphate and peptide-bound dolichol. 43
Figure 1-8. Independent and sequential models of polyprenol-dependent glycan biosynthesis. .47

Chapter 2

Figure 2-1. Polyprenol phosphorylation reactions.....59
Figure 2-2. Undecaprenyl-dependent protein glycosylation pathways. 60
Figure 2-3. Two cellular sources of Und-P. 62
Figure 2-4. The structure of diacylglycerol. 65
Figure 2-5. Biosynthesis of the *C. jejuni* undecaprenyl-diphosphate-linked disaccharide.....67
Figure 2-6. Time course assay in which the kinase candidates are coupled to PglC and PglA. ..68
Figure 2-7. Structure of the 2-AB-labeled diNAcBac-GalNAc. 68
Figure 2-8. ClustalW sequence alignment of various DGKs. 71
Figure 2-9. TMHMM prediction analysis of various DGKs. 72
Figure 2-10. Assay of the OTase PglB. 73
Figure 2-11. A time course of a PglB reaction. 74
Figure 2-12. Und-PP-linked glycan substrates generated by kinase coupled reaction.....76
Figure 2-13. A standard curve of dLBT-ubiquitin luminescence.....83

Chapter 3

Figure 3-1. O-linked protein glycosylation pathway in *Neisseria gonorrhoeae*. 89
Figure 3-2. Schematic representations of bacterial protein glycosylation pathways: 90
Figure 3-3. SDS-PAGE and Western blot analysis of purified Pgl proteins. 93
Figure 3-4. Structure of UDP-diNAcBac. 94
Figure 3-5. Normal phase HPLC separation of radiolabeled glycosyltrasnferase products. 99
Figure 3-6. Normal phase HPLC with fluorescence detection of 2-AB labeled glycans. 100
Figure 3-7. The UDP-specificity preferences of PglB, PglA, and PglE. 102
Figure 3-8. PglB and PglA specificity assay with UDP-4-amino. 103
Figure 3-9. The Und-PP-glycan specificity of PglE. 104
Figure 3-10. SDS-PAGE and anti-MBP Western blot analysis of purified MBP-PglH. 105
Figure 3-11. UDP-sugar specificity of PglH. 105
Figure 3-12. NP-HPLC trace of 2-AB labeled diNAcBac-Glc. 106
Figure 3-13. Normal phase HPLC separation of the product of PglH.....107

Chapter 4

Figure 4-1. Enzymatic reactions mediated by PglB(*Cj*) and PglO(*Ng*).....126
Figure 4-2. Transmembrane topology predictions for PglB(*Cj*) and PglO(*Ng*).....127
Figure 4-3. A pilin monomer from *N. gonorrhoeae* modified at Ser 63 by Gal-α-(1,3)-diNAcBac.....129
Figure 4-4. Anti-His_4 Western blot and SDS-PAGE analysis of PglO(*Ng*) and pilin.....131
Figure 4-5. Purification of glycosylated pilin by Ni-NTA resin.....131
Figure 4-6. Western blot analysis of unmodified pilin and PglO-glycosylated pilin.....132
Figure 4-7. Und-PP-linked glycans assayed with PglO(*Ng*) and PglB(*Cj*).....133
Figure 4-8. Glycan donor specificity of PglB(*Cj*) and PglO(*Ng*).....135
Figure 4-9. Verification of glycosylated residues in pilin protein substrate.....137
Figure 4-10. Activity of PglO(*Ng*) with pilin, Ng1225, and Ng1769.....139
Figure 4-11. Western blot analysis of PglO(*Ng*) protein substrate specificity.....139
Figure 4-12. PglB(*Cj*) and PglO(*Ng*) activity with a pilin peptide substrate.....140

Chapter 5

Figure 5-1. N-linked protein glycosylation in *C. jejuni*.....152
Figure 5-2. Generic integral membrane protein in lipid bilayer or detergent micelle.....153
Figure 5-3. Representations of full-length ApoA-1, truncated ApoA-1 and a Nanodisc.....155
Figure 5-4. A model of Und-PP-heptasaccharide biosynthesis at the membrane interface.....156
Figure 5-5. Western blot analysis of original PglC purification attempts.....158
Figure 5-6. SDS-PAGE analysis of PglC constructs.....159
Figure 5-7. Western blot analysis of successful His_6-GB1-PglC purification.....160
Figure 5-8. Transmembrane topology prediction of PglC and MraY (*E. coli*).....161
Figure 5-9. SEC analysis of detergent-solubilized PglA sample.....163
Figure 5-10. Schematic of Nanodisc preparation.....165
Figure 5-11. Purification and quantification of PglC-containing Nanodiscs.....166
Figure 5-12. Standard curves from the gels depicted in Figure 5-11.....166
Figure 5-13. Size exclusion chromatography of PglC-containing Nanodiscs.....168
Figure 5-14. SDS-PAGE analysis of PglA-containing Nanodiscs.....169
Figure 5-15. SDS-PAGE analysis of PglC-PglA-Nanodisc quantification by gel densitometry.....171

Chapter 6

Figure 6-1. Co-incorporation of PglC and PglA into Nanodiscs.....187
Figure 6-2. Depiction of LRET between an LBT and a Bodipy fluorophore.....189
Figure 6-3. Secondary structure prediction of PglC and identification of PglC Cys mutant candidates.....191
Figure 6-4. Activity of PglC and three Cys-containing PglC mutants.....192
Figure 6-5. SDS-PAGE analysis of PglC labeled with Bodipy-FL.....193
Figure 6-6. Comparison of PglC activity with and without fluorophore labeling.....194
Figure 6-7. The I-TASSER predicted structure of PglA.....195
Figure 6-8. AlexaFluor 568 and Bodipy FL as FRET partners.....196
Figure 6-9. SDS-PAGE analysis of fluorophore labeled PglA.....197

Figure 6-10. SDS-PAGE analysis of Nanodisc samples containing fluorescently-labeled versions of PglC and PglA. 198
Figure 6-11. Diagram of the N-terminal and C-terminal single-LBT constructs. 200
Figure 6-12. Experimental design for analysis of PglA and PglC-containing Nanodiscs by LRET. 201
Figure 6-13. SDS-PAGE analysis of PglC-Nanodiscs with PglA, LBT-PglA and PglA-LBT. 203
Figure 6-14. Amphipol-solubilized PglA. 205

Chapter 7

Figure 7-1. The chemical structure of bacterial undecaprenyl-phosphate (Und-P) 215
Figure 7-2. Independent and sequential models of glycosyltransferase activity in *C. jejuni*. 216
Figure 7-3. Phosphoglycosyl- and glycosyltransferase reactions carried out by PglC and PglA, respectively. 218
Figure 7-4. PglB(*Ng*) produces Und-PP-diNAcBac. 219
Figure 7-5. Preparation of "unloaded" Nanodiscs containing radiolabeled Und-PP-glycans. 220
Figure 7-6. Incorporation of Und-PP-[^{3}H]diNAcBac into "unloaded" Nanodiscs 221
Figure 7-7. Incorporation of both Und-PP-[^{3}H]diNAcBac and PglC into Nanodiscs. 222
Figure 7-8. Activity assay of PglA in Nanodiscs. 223
Figure 7-9. PglA activity in detergent-based micelles versus Nanodiscs 224
Figure 7-10. Ni-NTA purification of Nanodiscs reacted with PglA. 225
Figure 7-11. Coupled enzyme assay of PglC in Nanodiscs. 227
Figure 7-12. Preliminary assay of PglC in Nanodiscs coupled to PglA activity. 227
Figure 7-13. A schematic of a Nanodisc preparation with PglC, PglA, and Und-[^{33}P]P. 229
Figure 7-14. Analysis of Nanodiscs containing PglC, PglA, and Und-[^{33}P]P 229
Figure 7-15. Assay of Nanodiscs pre-loaded with PglC, PglA, and Und-[^{33}P]P 230
Figure 7-16. Activity of Nanodiscs containing PglC, PglA, and Und-P. 230
Figure 7-17. Assay of Nanodiscs containing PglC, PglA and Und-P. 231

List of Tables

Chapter 1

Table 1-1. Structures of polyprenols in different representative species. 19
Table 1-2. Intracellular amounts of dolichol and dolichyl-phosphate in mammalian sources. 26
Table 1-3. The percentages of undecaprenol, undecaprenyl-phosphate (Und-P), and undecaprenyl-diphosphate (Und-PP) in various bacterial species. 31

Chapter 2

Table 2-1. Specificity of *S. mutans* DGK and *C. jejuni* DGK for undecaprenol, diacylglycerol, and dolichol. 66
Table 2-2. Results from studies of dLBT-kinase activity and luminescence. 69
Table 2-3. Percent turnover of *C. jejuni* DGK and *S. mutans* DGK. 73
Table 2-4. Primers for the amplification of the candidate kinase genes from genomic DNA. 79

Chapter 3

Table 3-1. Percent sequence identity (%) between Pgl proteins. 91
Table 3-2. Comparison of *C. jejuni* and *N. gonorrhoeae* UDP-diNAcBac ^{1}H chemical shifts and coupling constant assignments. 94
Table 3-3. Steady-state parameters for PglC and PglB-ATD. 96

Chapter 4

Table 4-1. Primers for generation of pilin alanine mutants. The mutated residues are italicized. 143

Chapter 5

Table 5-1. PglC constructs screened for improved solubility characteristics. 159
Table 5-2. Membrane scaffold protein (MSP) constructs and Nanodisc sizes. 164
Table 5-3. PglC and MSP quantification in 12 nm Nanodiscs using gel densitometry. 167
Table 5-4. PglC and MSP quantification in 13 nm Nanodiscs using gel densitometry. 167
Table 5-5. PglC, PglA and MSP quantification in 13 nm Nanodiscs using gel densitometry. 171
Table 5-6. Primers used for preparation of PglC DNA constructs. 173

Chapter 6

Table 6-1. Primers used to generate PglC Cys mutants and PglA-LBT DNA constructs 207
Table 6-2. DTNB quantification of PglA thiol content. 209

List of Abbreviations

2-AB	2-aminobenzamide
AcCoA	acetyl coenzyme A
ApoA-1	Apolipoprotein A1
BacA	*Escherichia coli* undecaprenyl pyrophosphate phosphatase
CDG	congential glycosylation disorder
CEF	Cell envelope fraction
CEF	cell envelope fraction
DDM	n-dodecyl-β-D-maltopyranoside
DGK	Diacylglycerol kinase
dLBT	Double Lanthanide Binding Tag
DLS	dynamic light scattering
Dol-P	dolichyl-phosphate
Dol-PP	dolichyl-diphosphate
DRS	dolichol recognition sequence
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid) or Ellman's reagent
ER	endoplasmic reticulum
FRET	fluorescence resonance energy transfer
GDP-Man	Guanine diphosphate mannose
GlPglB(*Ng*)	Glycosyltransferase domain of PglB(*Ng*)
GPI	glycosylphosphatidylinositol
GT-B	glycosyltrasnferase-B protein fold
LBT	Lanthanide Binding Tag
LBT-PglA	N-terminal LBT-PglA(*Cj*)
Lipid I	Und-PP-MurNAc(pentapeptide)
Lipid II	Und-PP-MurNAc(pentapeptide)-GlcNAc
LRET	luminescence resonance energy transfer
MALDI MS	matrix-assisted laser desorption ionization mass spectrometry
MBP	maltose binding protein
NDP	nucleotide diphosphate
Ni-NTA	Ni-nitrilotriacetic acid
NP-HPLC	normal phase-high performance liquid chromatography
OT	the eukaryotic oligosaccharyltransferase
OTase	general oligosaccharyltransferase
P-diNAcBac	Phosphorylated diNAcBac
Pgl(*Cj*)	Protein glycosylation locus in *C. jejuni*
Pgl(*Ng*)	Pilin glycosylation locus in *N. gonorrhoeae*
PglA-LBT	C-terminal PglA(*Cj*)-LBT
PglA(*Cj*)	Second glycosyltransferase in *C. jejuni*, transfers a GalNAc

PglA(*Ng*)	Second glycosyltransferase in *N. gonorrhoeae*, transfers a Gal
PglB-ATD(*Ng*)	Acetyltransferase domain of PglB(*Ng*)
PglB-PGTD(*Ng*)	phospho-glycosyltransferase domain of PglB(Ng)
PglB(*Cj*)	*C. jejuni* oligosaccharyl transferase
PglB(*Ng*)	Bifunctional enzyme in *N. gonorrhoeae* that sequentially transfers acetyl group to form UDP-diNAcBac and then transfers P-diNAcBac to form Und-PP-diNAcBac.
PglC(*Cj*)	First glycosyltransferase in *C. jejuni*, transfers P-Bac to form Und-PP-Bac
PglC(*Ng*)	*N. gonorrhoeae* aminotransferase, forms UDP-4-amino
PglD(*Cj*)	*C. jejuni* acetyltransferase, forms UDP-diNAcBac
PglD(*Ng*)	*N. gonorrhoeae* dehydratase, forms UDP-4-keto
PglE(*Cj*)	*C. jejuni* aminotransferase, forms UDP-4-amino
PglE(*Ng*)	Third glycosyltransferase in *N. gonorrhoeae*, transfers Gal
PglF(*Cj*)	*C. jejuni* dehydratase, forms UDP-4-keto
PglH(*Cj*)	Fourth glycosyltransferase in *C. jejuni*, transfers three GalNAcs
PglI(*Cj*)	Branching glycosyltransferase in *C. jejuni*, transfers glucose
PglJ(*Cj*)	Third glycosyltransferase in *C. jejuni*, transfers GalNAc
PglO(*Ng*)	Putative oligosaccharyl transferase in *N. gonorrhoeae*
PIRS	polyisoprenol recongnition sequence
sLBT	single Lanthanide Binding Tag
TMHMM	Tied Mixture Hidden Markov Model (a transmembrane prediction program)
UDP-4-amino	Uridine diphosphate 4-amino-2-acetamido-2,4,6-trideoxy-α-D-glucose
UDP-4-keto	Uridine diphosphate 4-keto-2-acetamido-2,4,6-trideoxy-α-D-glucose
UDP-DATDH	Uridine diphosphate 2,4-diacetamido-2,4,6-trideoxy-α-D-hexose
UDP-diNAcBac	Uridine diphosphate 2,4-diacetamido-2,4,6-trideoxy-α-D-glucose
UDP-Gal	Uridine diphosphate galactose
UDP-GalNAc	Uridine diphosphate N-acetylgalactosamine
UDP-Glc	Uridine diphosphate glucose
UDP-GlcNAc	Uridine diphosphate N-acetylglucosamine
Und-OH	Undecaprenol
Und-P	Undecaprenyl phosphate
Und-PP	Undecaprenyl pyrophosphate
Und-PP-Bac	Product of both PglC(*Cj*) and PglB(*Ng*)
Und-PP-Bac-$(Gal)_2$	PglE(*Ng*) product, Gal-β(1,4)-Gal-α(1,3)-Bac-α-1-PP-Und
Und-PP-Bac-$(GalNac)_2$	PglJ(*Ng*) product, GalNAc-α(1,3)-GalNAc-α(1,3)-Bac-α-1-PP-Und
Und-PP-Bac-Gal	PglA(*Ng*) product, Gal-α(1,3)-Bac-α-1-PP-Und
Und-PP-Bac-GalNAc	PglA(*Cj*) product, GalNAc-α(1,3)-Bac- α-1-PP-Und
Und-PP-heptasaccharide	complete undecaprenyl-linked product in *C. jejuni* pathway

Chapter 1: Examining the role of polyisoprenols and polyisoprenyl-phosphates in biology

Introduction

The linear polyisoprenols are a unique class of secondary metabolites within the isoprenoid natural product family (Figure 1-1). These polyisoprenols comprise a small percentage (0.1%) of lipid bilayers in both prokaryotes and eukaryotes.[1-3] Polyprenyl-phosphate derivatives act as oligosaccharide carriers during glycan biosynthesis in many conserved, essential cellular processes including N-linked glycosylation, C- and O-protein mannosylation, and bacterial cell wall biosynthesis. Despite extensive research on the enzymes that utilize polyprenols in glycan assembly pathways, relatively little is known about why nature has chosen polyprenols to act as glycan carriers. This chapter will begin by describing polyprenol structure and polyprenyl-dependent processes in biology, which will be followed by an overview of polyprenol distribution across different species. With this information as essential background, we will explore three different hypotheses that move beyond polyprenol functions to address the more interesting question of "why polyprenols in the first place?"

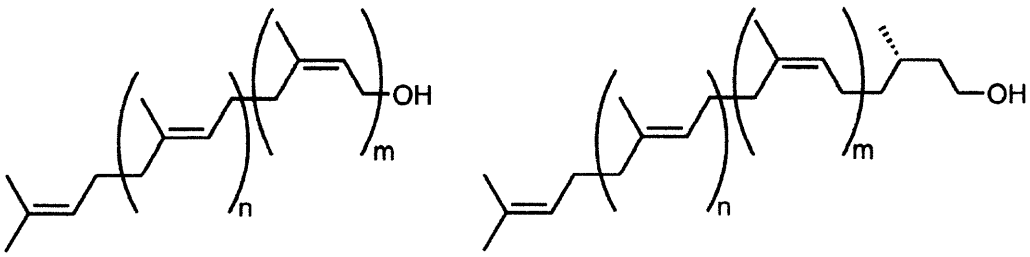


Figure 1-1. Structures of fully unsaturated polyprenols and dolichols.
The length of the polyprenol and the number of isoprene units (n and m) is dependent on the species, where m is *cis* and n is *trans*. *Left*, fully unsaturated polyprenols. *Right*, dolichols with a saturated α-isoprene unit.

The linear polyprenols (or polyisoprenols) contain 7 to 24 isoprene units in either the *trans* or *cis* configuration and can broadly be separated into two subclasses. The first class includes undecaprenol and related homologs, which contain only unsaturated isoprene units,

whereas the second class are designated as dolichols and possess a single saturated α-isoprene unit. In both classes, α refers to the unit closest to the hydroxyl moiety and ω is used to designate the terminal isoprene in the linear structure (Figure 1-1). The polyprenols in individual organisms are usually polydispersed ranging in number of isoprene units with one or two lengths comprising the majority of the polyprenol population. In general, polyprenol length increases with the complexity of organisms (Table 1-1). Bacteria typically contain undecaprenol, a C55 unsaturated polyprenol with two *trans* isoprene units at the ω end and nine *cis* isoprene units. Archaeal organisms typically contain C55-C66 dolichols,[4] *Saccharomyces cerevisiae* (yeast) contain C80 dolichols and mammalian tissues contain C90-C100 dolichols. Plants exhibit the widest range of polyprenol diversity, as both unsaturated polyprenols and dolichols have been identified.[5, 6] The lipid bilayers of most plants contain C55 polyprenols, although these polyprenols contain three ω-*trans* isoprene units instead of two. Plants in the gymnosperm family contain C80-C100 polyprenols, while plants in the genus *Potentilla* contain mixtures of polyprenols with the longest possessing up to 160 carbons.[7] It is curious that such a wide diversity of polyprenol architectures has evolved amongst the three kingdoms of life, as the known functions of these molecules are very similar in all organisms.

Table 1-1. Structures of polyprenols in different representative species.

Species	n	m	Dolichol?	Extended length	Ref
Bacteria	2	8	No	50 Å	8
Mycobacterium smegmatis	3*	3	No	34 Å	9, 10
	1	8	No	46 Å	
Halobacterium halobium (archaea)	2	9	Yes	60 Å	4
Saccharomyces cerevisiae	2	12-13	Yes	73-77 Å	11
Rattus norvegicus (rat)	2	13-16	Yes	76-89 Å	12
Homo sapiens	2	14-17	Yes	81-94 Å	12
Ficus elastica (plant)	3	6-8	No	47-56 Å	13
Aspergillus fumigatus (plant)	2	14-20	Yes	86-107 Å	5
Potentilla crinita (plant)	3	15-17	No	86-94 Å	6
	3	25-27		129-137 Å	
	3	35-37		172-180 Å	

* The four isoprene units at the ω-terminal end are fully saturated in this short polyprenol from *Mycobaterium smegmatis*. Two types of polyprenols are found in this bacterium.

Many studies have focused on the delineation of polyprenol biosynthesis, which involves the enzyme-catalyzed condensation of dimethylallylpyrophosphate (DMAPP) with multiple units of isoprenylpyrophosphate (IPP) (Figure 1-2). Interestingly, there are two major pathways for the synthesis of the precursor DMAPP and IPP molecules. The mevalonate-dependent pathway is found mostly in eukaryotic systems, while the more recently discovered 1-deoxy-D-xyulose-5 phosphate (DOXP) pathway is found in many plants and prokaryotes. The biosynthesis of isoprenoid alcohols has been extensively reviewed elsewhere.[14, 15]

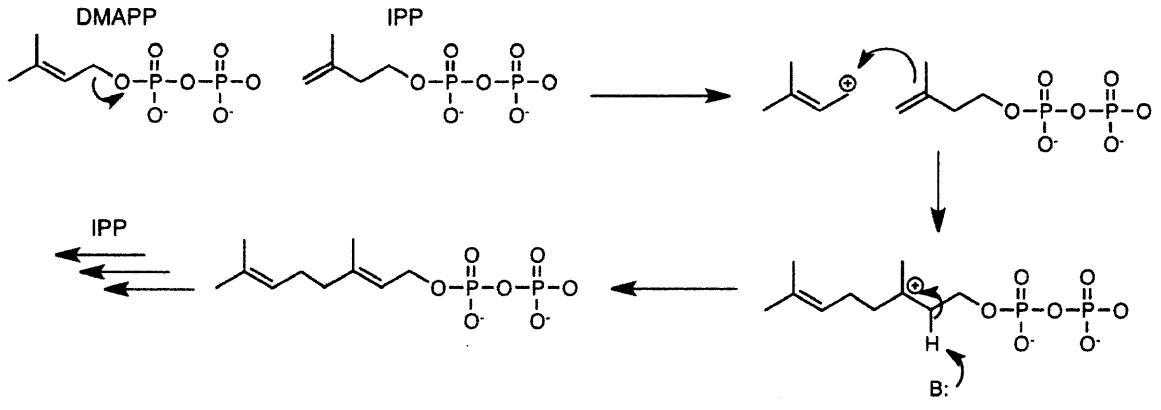


Figure 1-2. Mechanism of isoprene biosynthesis.
Prenyltransferases catalyze the condensation of DMAPP with IPP to form long polyprenyl-diphosphate molecules and direct which stereoisomer (*cis* or *trans*) is produced in each reaction.

Polyprenyl-phosphate dependent processes

Polyprenyl-phosphates are essential substrates for critical cellular functions in both eukaryotes and prokaryotes. N-linked glycosylation may be the most well known polyprenol-dependent process. It involves the synthesis of an oligosaccharide on polyprenyl-diphosphate or polyprenyl-phosphate and occurs in eukaryotes, archaea, and several bacterial species (Figures 1-3 and 1-4).[14, 16] In eukaryotes, a conserved dolichyl-diphosphate (Dol-PP) heptasaccharide is assembled on the cytoplasmic face of the endoplasmic reticulum (ER) membrane.[17-19] The Dol-PP-glycan intermediate is then translocated across the lipid bilayer to the ER lumen, where it is further glycosylated to form a tetradecasaccharide that is transferred to the nitrogen in the primary amide side chain of asparagine residues (Figure 1-3). To date, definitive characterization of N-linked glycosylation in archaea has been limited, but these pathways are known to involve assembly of small, highly modified glycans on dolichyl-phosphate and dolichyl-diphosphate carriers (Figure 1-3).[20, 21] Within the last decade, N-linked glycosylation was identified in the Gram-negative pathogen *Campylobacter jejuni*,[22, 23] however, genomic analysis suggests that homologous pathways may be found in at least 84 different prokaryotic species.[24] The *C. jejuni* pathway represents the most well-characterized bacterial glycosylation system, [16, 25-31] and involves the biosynthesis of an undecaprenyl-diphosphate (Und-PP) heptasaccharide prior to glycan transfer to asparagine in folded proteins in the periplasmic space (Figure 1-4). Interestingly, an analogous O-linked glycosylation pathway has recently been described in *Neisseria* species, in which the glycan is assembled on undecaprenyl-diphosphate prior to glycosylation of serine and threonine residues in periplasmic proteins (Figure 1-4).[32-34] Biochemical characterization of this pathway is the focus of both Chapters 3 and 4.

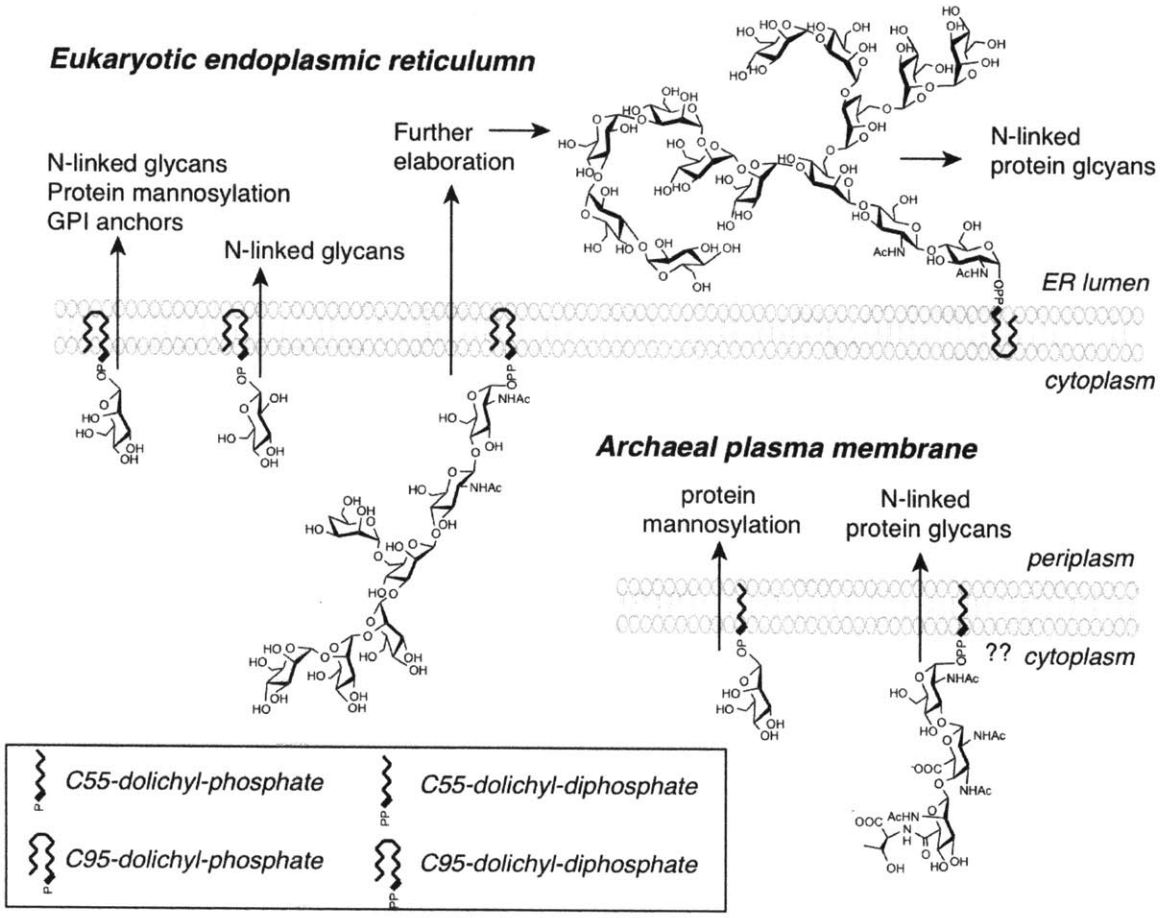


Figure 1-3. Dolichyl-linked glycan substrates in eukaryotes and archaea.
The glycans are depicted in the cellular region where they are biosynthesized and the arrow indicates that the glycans undergo enzyme-catalyzed flipping across the membrane prior to subsequent reactions. *Top*, From left to right, the eukaryotic ER membrane contains multiple species, including dolichyl-phosphate mannose, dolichyl-phosphate glucose, dolichyl-diphosphate heptasaccharide and dolichyl-diphosphate tetradecasaccharide. *Lower right*, The archaeal membrane contains dolichyl-phosphate mannose and dolichyl-linked glycan precursors to N-linked glycans. The glycan produced in *Methanococcus voltae* is shown and the question marks indicate that it is unknown whether this intermediate contains a diphosphate or phosphate linkage, since both have been observed in archaeal species.

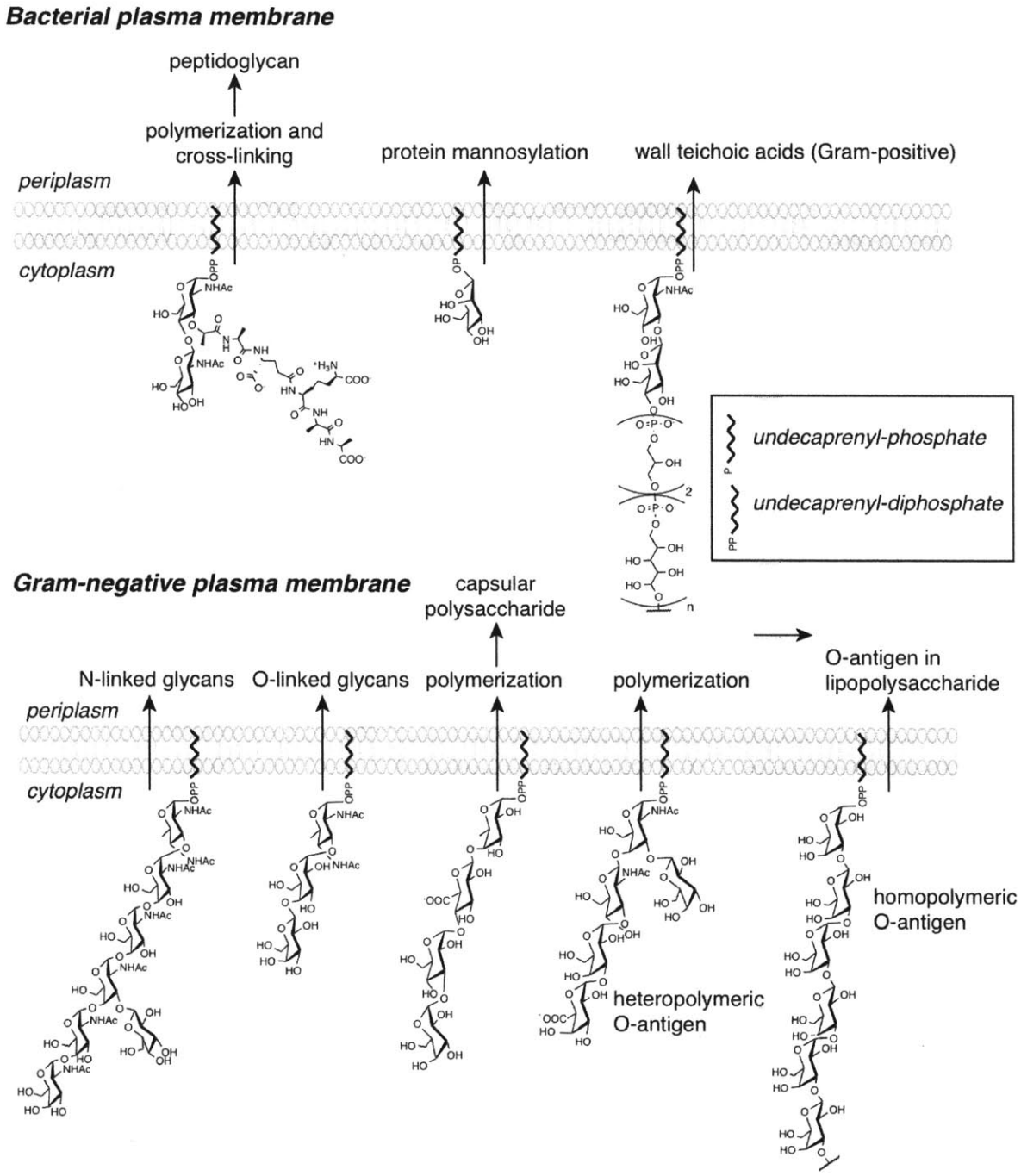


Figure 1-4. Undecaprenyl-linked glycans in bacterial plasma membranes.
The glycans are depicted in the cellular region where they are biosynthesized and the arrow indicates that the glycans undergo enzyme-catalyzed flipping across the membrane prior to subsequent reactions. *Top*, From left to right, both Gram-negative and Gram-positive bacteria contain Lipid II and undecaprenyl-phosphate-mannose, while teichoic acids are present in Gram-positive bacteria like *Staphylococcus aureus*. *Bottom*, From left to right, examples of undecaprenyl-diphosphate-linked glycans found in Gram-negative bacteria are shown including N-linked glycans from *C. jejuni*, O-linked glycans from *N. gonorrhoeae*, capsular polysaccharide

from *E. coli* strain K27, heteropolymeric O-antigen from *E. coli* strain O113, and homopolymeric O-antigens from *Klebsiella pneumoniae*.

In addition to its role in protein glycosylation, undecaprenol acts as membrane anchor for the biosynthesis of many extracellular, oligosaccharide-based structures in bacteria (Figure 1-4). Importantly, peptidoglycan intermediates are synthesized on undecaprenol; these intermediates comprise Und-PP-MurNAc(pentapeptide) (Lipid I) and Und-PP-MurNAc(pentapeptide)-GlcNAc (Lipid II). After Lipid II is translocated across the membrane, peptidoglycan is assembled through disaccharide polymerization and peptide cross-linking forming the rigid cell wall barrier that is essential for bacterial viability. Because peptidoglycan biosynthesis has provided excellent targets for antibiotics, much effort has focused on understanding how Lipid II precursors are biosynthesized and incorporated into the cell wall.[15, 35, 36] Several other essential bacterial components are synthesized via undecaprenyl-diphosphate intermediates including O-antigen polymers, [37, 38] teichoic acids,[39] and capsular polysaccharides.[40-42] These extracellular glycans are involved in bacterial defense mechanisms, mediate crucial microbial-host interactions and may also represent antibacterial target pathways.

Finally, polyprenyl-phosphate-linked saccharides play an essential role in glycan biosynthesis by acting as activated glycan donor intermediates for transfer to protein or other biosynthetic intermediates; these molecules are membrane-bound alternatives to nucleotide-diphosphate (NDP) activated donors.[43] Polyprenyl-phosphate saccharides are assembled on the cytoplasmic surface of a cellular membrane from an NDP-sugar and polyprenyl-phosphate; these substrates are then translocated across the ER membrane in eukaryotes or the plasma membrane in prokaryotes, where they serve as saccharide donors in the ER lumen or periplasm, where NDP-sugars are absent (Figures 1-3 and 1-4). In this way, cellular systems exploit polyprenyl

translocation to transport activated sugar donors to regions that lack NDP-sugars, but still require glycan biosynthesis.

Specifically, polyprenyl-phosphate mannose and polyprenyl-phosphate glucose are the most common polyprenyl-linked glycan donors. Polyprenyl-phosphate mannose is the substrate for protein O-mannosylation, which occurs in eukaryotes and prokaryotes,[43] and C-mannosylation of tryptophan residues in proteins, which is a rare modification that has only been identified in a handful of eukaryotic proteins.[44] In addition to acting as a direct substrate for proteins, Dol-P-mannose is also the mannose donor for glycosylphosphatidylinositol (GPI)-anchor biosynthesis in the ER lumen. GPI anchors are involved in anchoring eukaryotic proteins to the plasma membrane.[45] Importantly, Dol-P-glucose and Dol-P-mannose participate in eukaryotic N-linked glycosylation by serving as monosaccharide donors for the assembly of the Dol-PP-tetradecasaccharide on the ER luminal interface.

Dolichyl-phosphates, as opposed to dolichyl-diphosphates, may also play an essential role in the biology of some archaeal species. Both mono- and di-phosphorylated dolichols have been identified in a variety of species,[4, 46-49] but Dol-P-glycans may be used exclusively by certain archaea such as *Haloferax volcanii*.[47] This finding suggests that Dol-P-monosaccharides may act as donors for the individual glycosyltransferases in archaeal N-linked glycosylation, and importantly, that in some species dolichyl-phosphate may also serve as the glycan carrier instead of dolichyl-diphosphate. The monophosphate linkage is more stable chemically and if this observation holds true for other archaeal species, it could represent a way that these organisms have adapted to extreme conditions.

In contrast to polyprenyl-phosphates, free polyisoprenols to date have no clear function in biological systems, despite the fact that they accumulate at high levels within certain organisms.

As a result, it is believed that these compounds must play a distinct role in cellular physiology, rather than simply providing a substrate pool for polyprenyl-phosphate dependent processes. The remainder of this chapter will focus on the examination of evidence accumulated in recent years concerning the roles of polyprenols and polyprenyl phosphates. In the next section, polyprenol distribution and abundance in membranes and the proposed roles of unmodified polyprenol will be described. We will then evaluate evidence for three hypotheses as to why polyprenyl-phosphates are used as glycan carriers in biology. These hypotheses are: (1) polyprenyl-phosphates exert regulatory roles controlling glycan biosynthesis and cell growth. (2) Glycan translocation across lipid bilayers is a ubiquitous feature of all polyprenyl-dependent pathways, but the flippase-mediated mechanism of this process is poorly understood and polyprenol structures may act to facilitate translocation. (3) Glycan assembly pathways reliant on polyprenyl derivatives often contain many membrane-associated enzymes, which potentially form macromolecular complexes at the lipid bilayer interface dependent on the presence of polyprenyl-linked substrates.

Polyprenol and polyprenyl-phosphate abundance in eukaryotic membranes

In biological membranes, polyprenols can undergo esterification, phosphorylation, and addition of phospho-monosaccharides and diphospho-glycans.[1, 3, 50, 51] Many studies have determined the concentrations of dolichol and dolichyl-phosphate for a variety of eukaryotic cell types including mammalian cell culture lines and tissue extracts from rats and humans (Table 1-2). Typically dolichol and its derivatives are extracted from cellular sources and purified by high performance liquid chromatography (HPLC). The various populations are quantified by comparison to known quantities of standards or by metabolic radiolabeling. Slight variations in

extraction and purification methods probably contribute to some of the inconsistencies observed in Table 1-2. For instance, more recent studies by Keller and coworkers reported a new method for quantification of polyprenyl-phosphates relying on saponification, which resulted in higher dolichyl-phosphate yields.[52, 53]

Table 1-2. Intracellular amounts of dolichol and dolichyl-phosphate in mammalian sources. The amounts were determined by radiolabeled incorporation of mevalonate[a] or acetate,[b] or by HPLC-based methods (given in μg dolichol per g of cells).

Species	Dolichol	Dolichyl-phosphate	Ref
CHO cells	74 cpm[a]	740 cpm[a]	54
	690 cpm[a]	270 cpm[a]	1
Rat liver	270 cpm[a]	59 cpm[a]	55
	43000 cpm[b]	890 cpm[b]	55
	43.0 μg/g	8.75 μg/g	12
	41 μg/g	9.38 μg/g	56
	17.1 μg/g	14.7 μg/g	53
	21.1 - 52.2 μg/g[c]	22.0 - 16.7 μg/g[c]	52
Rat testes	4.3 μg/g	16.0 μg/g	53
	9.36 - 27.9 μg/g[c]	15.3 - 19.8 μg/g[c]	52
Rat spleen	39.3 μg/g	18.7 μg/g	53
	58.9 - 157 μg/g[c]	11.6 - 24.8 μg/g[c]	52
Rat brain	11.3 μg/g	16.2 μg/g	53
	9.36 - 27.9 μg/g[c]	15.3 - 19.8 μg/g[c]	52
Rat kidney	7.8 μg/g	20.0 μg/g	53
	12.3 - 24.1 μg/g[c]	22.0 - 20.0 μg/g[c]	52
Human liver	465 μg/g	10.8 μg/g	12
	486 μg/g	11.9 μg/g	56
Human testis	1493 μg/g	169 μg/g	12

[c] In this study, the dolichol and dolichyl-phosphate amounts were determined at 4, 6, 8, 10 and 14 weeks; the values for 4 and 14 weeks are shown.

More interestingly, other discrepancies in Table 1-2 could stem from additional factors that affect the amount of cellular dolichol, such as aging. In both mammalian tissues[52] as well as plants,[6, 57] polyprenol tends to accumulate over the lifetime of the organism. Interestingly, the relative amounts of dolichol and dolichyl-phosphate quantified from several rat organs suggested that the amount of unmodified dolichol increased with age, whereas dolichyl-phosphate levels

remained constant.[52] This general finding has been verified both in plants[6] and in human brain tissue, which shows 3 to 20-fold accumulation of dolichols over a lifespan, but relatively consistent levels of dolichyl-phosphate.[58-60] The cause of this increase is unknown, but it is suggestive that the accumulation of dolichol could be a factor in the side effects of aging. In addition to aging, altered dolichol levels have been measured in a variety of disease states suggesting that an increased dolichol concentration could serve as a biomarker for various diseases including liver cancer and cirrhosis, cataract disease, lysosomal storage disorders, and neurodegenerative diseases including Alzheimer's and neuronal ceroid lipofuscinosis.[12, 51, 61-66] However, several studies in plants and mammals have proposed that polyisoprenols may have antioxidant properties and participate in scavenging of free radical oxygen species.[67-69] If this result is further validated, then it could provide some explanation of the varied dolichols levels in disease states.

In addition to quantification-based profiling, studies on the cellular localization of dolichols have provided a few hints concerning alternative functions for native and esterified dolichols. The distribution of dolichol, dolichyl ester and dolichyl-phosphate within the intracellular organelles of eukaryotes has been quantified.[50] As expected, it was found that the endoplasmic reticulum, which is the site of N-linked glycosylation and other dolichyl-phosphate dependent processes in eukaryotic cells, contains a significant amount of dolichol.[50] More surprisingly, several studies identified lysosomes as intracellular organelles with high levels of dolichyl esterified with a variety of fatty acids.[50, 70] This observation is interesting in light of the finding that altered dolichol levels were observed in lysosomal storage disorders. It has been suggested that dolichols could play a role in the intracellular trafficking of fatty acids, or alternatively, that fatty acid derivatives of dolichols act as transport vehicle for dolichols.[50] More

recent studies have implicated a cis-prenyltransferase, a key enzyme in dolichol biosynthesis, in protein trafficking and retention within the dynamic ER and Golgi membrane network, suggesting that dolichol and its derivatives may a play a role in protein localization.[71, 72]

Another way to evaluate the potential role of dolichols in the membrane is via biophysical studies on the effects of polyisoprenols and polyisoprenyl derivatives in model membrane vesicles. A variety of techniques have been used to address this question including nuclear magnetic resonance spectroscopy, differential scanning calorimetry, small angle X-ray diffraction, electrophysiology, and freeze-fracture electron microscopy.[73-80] From these studies, it is clear that polyisoprenols increase membrane fluidity, ion permeability and the propensity of membranes to adopt a non-bilayer, hexagonal II conformation (Figure 1-5).[73-77] In addition, dolichol and dolichyl esters were found to be oriented parallel to the membrane surface and form aggregates near the center of the bilayer, whereas dolichyl-phosphates are monodispersed and oriented perpendicular to membrane with the charged headgroup at the membrane surface (Figure 1-5).[75, 78] In addition, small-angle X-ray scattering and NMR studies[76, 81-83] have shown that polyprenols form three separate domains with a coiled center region such that these extended linear molecules (Table 1-1) can reside within the width of the lipid bilayer.

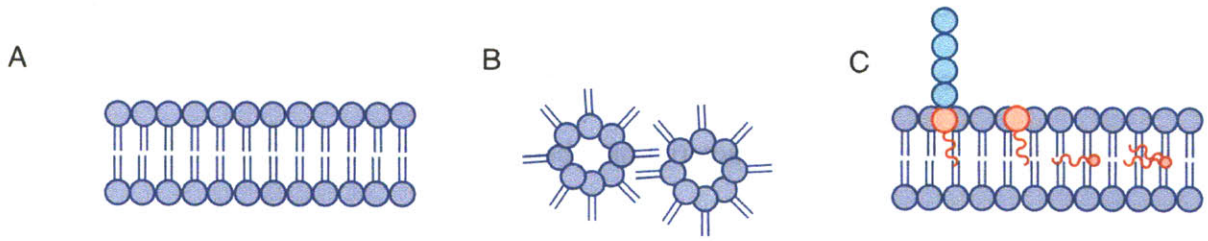


Figure 1-5. Phospholipid structures in presence of polyprenols.
A, Lipids form a typical lipid bilayer in the absence of polyprenol. *B,* Some evidence suggests that the hexagonal II conformation (shown here) is induced in the presence of polyprenols.[74, 75, 82] *C,* Polyprenol derivatives have different orientations in the membrane. From left to right, a lipid bilayer is shown containing a polyprenyl-phosphate glycan and polyprenyl-phosphate perpendicular to the membrane, while polyprenol and esterified polyprenol parallel to the membrane surface.

One caveat with many of the reported studies is that most of the biophysical experiments on polyisoprenols were performed in model membranes containing pure phospholipids. These model membranes are poor substitutes for native membrane bilayers that contain a diverse array of phospholipids and proteins. In a recent set of studies, the inclusion of a small hydrophobic peptide into dolichol-containing membrane vesicles reversed the bilayer destabilization caused by dolichols.[82, 83] Further studies are needed to understand the effects of polyisoprenols in the context of physiological membranes, but it is apparent from these biophysical studies that dolichol and dolichyl-phosphate can influence membrane properties. It will be interesting to learn whether the biophysical effects of dolichol accumulation in the phospholipid bilayer could contribute to aging or other cellular processes.

Undecaprenol and undecaprenyl-phosphate abundance in microbial membranes

Most studies on polyprenols have focused on mammalian tissues and plants, as it is more facile to extract significant amounts of material from macroscopic sources. However,

undecaprenol is a key molecule in the biosynthesis of many essential bacterial structures including peptidoglycan, but only a few studies have examined the relative populations of various undecaprenyl derivatives in bacterial membranes (Table 1-3). Interestingly, it appears as though *Escherichia coli*, unlike *Staphylococcus aureus*, does not possess reserve pools of free undecaprenol.[84] Further studies are required to determine if these results are applicable to all Gram-negative and Gram-positive bacteria. This type of general finding would have important biological implications suggesting that the two classes of bacteria may have developed very different ways of regulating undecaprenyl-phosphate production. Regulation of undecaprenyl-phosphate may represent a distinguishing feature between the two classes of bacteria that has been previously unexplored.

Although the undecaprenol levels have not been examined in other Gram-negative bacteria, several studies indicate that *S. aureus* and other Gram-positive bacteria like *Streptococcus faecalis* contain free undecaprenol (Table 1-3).[85] In addition, the lack of undecaprenol in Gram-negative bacteria is supported by the fact that no undecaprenol kinases have been identified in Gram-negative bacteria, whereas several have been identified in Gram-positive bacteria.[63-67] As mentioned above, free polyprenols are known fluidize model membranes, and thus it is possible that the lack of undecaprenol in *E. coli* stabilizes the plasma membrane. The low undecaprenol levels may be important for maintaining the unique dual membrane architecture of Gram-negative bacteria.

Table 1-3. The percentages of undecaprenol, undecaprenyl-phosphate (Und-P), and undecaprenyl-diphosphate (Und-PP) in various bacterial species.

Species	Undecaprenol	Und-P	Und-PP	Reference
E. coli	nd[a]	22% (75 nmol/g)[b]	78% (260 nmol/g)[b]	84
S. aureus	31% (70 nmol/g)	19% (43 nmol/g)[b]	50% (113 nmol/g)[b]	84
	90%	<10%	[c]	8
S. faecalis	82%	18%	[c]	85

[a] nd, no significant amounts were detected

[b] nmoles polyprenol per gram cell pellet

[c] For the latter two studies, no attempt was made to quantify the amount of Und-PP.

In addition to the presence or absence of undecaprenol, another key difference between Gram-negative and Gram-positive bacteria is the amount of Lipid II. Arguably the most essential derivative of bacterial undecaprenol is Lipid II, which is the intermediate required for peptidoglycan biosynthesis in both Gram-negative and Gram-positive bacteria, a process that is essential to bacterial viability (Figure 1-4). Metabolic labeling has been used to quantify the pools of Lipid I/II relative to other substrates required for peptidoglycan biosynthesis in *E. coli* resulting in estimation of ~3000 copies per cell.[86] Based on the work described in Table 1-3,[84] this would represent <2% of the undecaprenyl-phosphate supply in *E. coli*.

In contrast, the amount of Lipid I/II has been estimated in a variety of Gram-positive bacteria and in every case the amount is much higher than in *E. coli*. Metabolic labeling was used to determine that each cell of *Bacillus megaterium* contained approximately 34,000 Lipid I/II molecules.[87] Other studies have quantified the number of precursors by determining the number of cellular binding sites for various Lipid II targeting antibiotics such as ramoplanin and mersacidin.[88] These reports suggest that there are 50,000 Lipid I/II precursors in *S. aureus,*[89] 70,000 in *Micrococcus luteus,*[90] and 200,000 in *Staphylococcus simulans.*[90] In *S. aureus*, this

amount of Lipid I/II would represent over 40% of the total pool of undecaprenyl derivatives, a considerably larger percentage than in *E. coli*. It is important to note that the amount of Lipid I/II in *E. coli* has not been measured directly and some error would be expected with estimation method previously used.[86]

Although more accurate quantitative studies are needed, it appears that *E. coli* and *S. aureus* contain significantly different levels of undecaprenol and Lipid I/II. Since Gram-positive bacteria have much thicker peptidoglycan layers (20-80 nm) than Gram-negative (7-8 nm), it is not surprising that Gram-positive bacteria have larger pools of Lipid I/II. In addition, the differences may imply that Gram-negative and Gram-positive bacteria rely on different mechanisms for the regulation of undecaprenyl derivatives in cell wall biosynthesis and other undecaprenyl-dependent processes. These distinctions may be specific to *E. coli* and *S. aureus*, but further studies are needed to determine if they represent a general physiological distinction between Gram-negative and Gram-positive bacteria.

Why polyprenols? Regulation of polyprenyl-dependent pathways

Polyprenyl-phosphate is a rate limiting substrate

The first hypothesis as to why polyprenyl-phosphates are ubiquitous glycan carriers involves their role as regulators of essential cellular processes. The amount of dolichyl-phosphate has been shown to modulate glycoprotein biosynthesis in eukaryotes.[2, 91-94] In one study, estrogen-induced differentiation of hen oviducts resulted in increased production of Dol-P-[^{14}C]mannose (Man) when incubated with GDP-[^{14}C]Man. The elevated amounts of Dol-P-[^{14}C]Man were attributed to higher levels of dolichyl-phosphate and not to increased enzymatic activity, thus implying that increased production of dolichyl-phosphate is an important regulator

of cell differentiation.[92] Similarly, when sea urchin embryos were incubated with an inhibitor of polyisoprenol synthesis, abnormal gastrulation was observed, which was correlated with the inability of the cell to produce glycoproteins. Addition of exogenous dolichol allowed for normal gastrulation, suggesting that dolichol-phosphate is a limiting reagent for N-linked glycoprotein biosynthesis and subsequent cellular transformations. The rate of dolichyl-phosphate and glycoprotein synthesis has also been linked to the growth rate of CHO cells and cell division.[2] Several other studies have confirmed that dolichyl-phosphate is a rate limiting substrate in N-linked glycosylation and is thereby a key factor in cellular development.[1, 91] It is clear from these studies that dolichyl-phosphate levels regulate the amount of glycoprotein biosynthesis, thus affecting important transitions during cellular development. Because dolichol biosynthesis operates independently from phospholipid biosynthesis, it can be differentially regulated such that production of glycoproteins is unrelated to other membrane-bound processes.

If the amount of dolichyl-phosphate is rate determining, then this implies that the dolichyl-phosphate biosynthesis must be tightly regulated. Dolichyl-phosphate can be produced by phosphorylation of dolichol and by hydrolysis of dolichyl-diphosphate. For the first route, a CTP-dependent dolichol kinase has been identified, which has been shown to be important during cellular development.[95] However, *de novo* phosphorylation of dolichol is not considered to be the primary source of dolichyl-phosphate (Figure 1-6). Alternatively, dolichyl-phosphate is regenerated through the action of a pyrophosphatase that acts on dolichyl-diphosphate, which is the product resulting from polyprenol biosynthesis and also a byproduct after glycan transfer by the oligosaccharyltransferase. A recent study has established that this second path is the principal source of dolichyl-phosphate in ER membranes (Figure 1-6).[96, 97] Another crucial enzyme for dolichol synthesis is the polyprenyl-diphosphate reductase, which produces the

saturated α-isoprene unit.[54] Importantly, the aforementioned enzymes modulate the amounts of dolichyl-phosphate available for glycan biosynthesis and thus, cellular defects in these enzymes can cause congenital glycosylation disorders (CDGs). CDGs are a class of diseases characterized by aberrant glycosylation, which can arise from mutations in the biosynthesis of both dolichols[98] and glycoproteins.[99] Importantly, CDGs have been identified that are related to mutations in dolichol kinase and polyprenyl-diphosphate reductase, which both regulate dolichyl-phosphate production.[99, 100] It is clear that regulation of dolichyl-phosphate plays an essential role in mediating glycoprotein biosynthesis, which is crucial to eukaryotic cellular function.

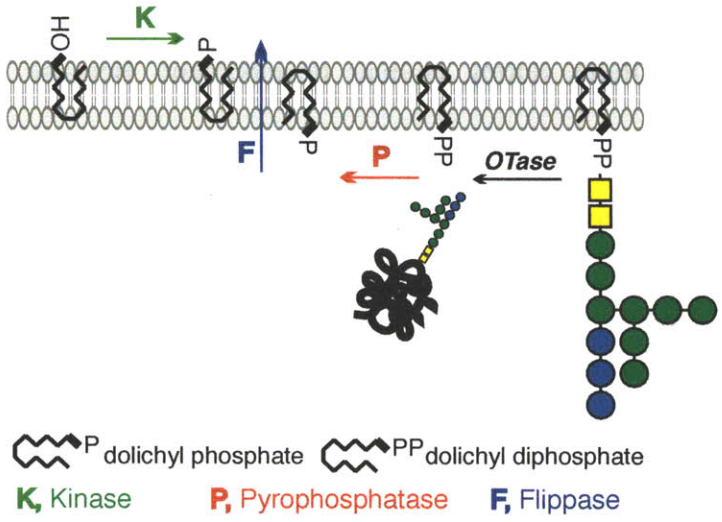


Figure 1-6. Cellular sources of dolichyl-phosphate.
Dolichol is phosphorylated *de novo* by a kinase (K) to generated dolichyl-phosphate or a salvage pathway occurs in which dolichyl-diphosphate is hydrolyzed by a pyrophosphatase (P) and then translocated across the membrane by a flippase (F) to begin another round of glycan biosynthesis. Dolichyl-diphosphate is the intermediate released after glycan transfer to protein.

The situation is less clear in bacterial systems, since it is not known if the amount of undecaprenyl-phosphate or Lipid II regulates cell growth via peptidoglycan biosynthesis. It seems unlikely in the case of *Escherichia coli*, which contains only undecaprenyl-phosphate and

undecaprenyl-diphosphate.[84] In addition, the cellular routes to undecaprenyl-phosphate are poorly defined. Undecaprenol kinases have been identified in Gram-positive bacteria,[101-105] but as mentioned above, not in Gram-negative bacteria, consistent with the lack of significant levels of undecaprenol in *Escherichia coli*. Interestingly, the undecaprenol kinase from *Streptococcus mutans* was identified as a virulence factor and may play a role in bacterial adaptation to different environments.[102] In addition, undecaprenyl-diphosphate pyrophosphatases have also been identified suggesting that undecaprenyl-phosphate is regenerated in a manner similar to dolichyl-phosphate.[106]

Polyprenyl-phosphates are shared amongst competing enzymes and pathways

An interesting facet of polyisoprenol regulation is that in most cases these molecules are substrates for several competing pathways or enzymes. In eukaryotic systems, three glycosyltransferases utilize dolichyl-phosphate to produce the three dolichyl substrates required for N-linked glycosylation: Dol-P-mannose, Dol-P-glucose and Dol-PP-N-acetylglucosamine.[1] It has been shown that the production levels of the three dolichyl transformations are linked *in vivo* such that if one enzyme is inhibited, the activity of the other two increase with the availability of dolichyl-phosphate substrate.[1] These results imply that the enzymes share a common pool of substrate and further confirm that dolichyl-phosphate is a rate limiting substrate.

The situation is less clear in bacterial systems, which contain multiple systems that utilize undecaprenyl-phosphate as a glycan carrier. In *Bacillus licheniformis*, which contains both peptidoglycan and teichoic acids built on undecaprenyl-phosphate, it was proposed that these pathways are not limited by a common pool of undecaprenyl carrier.[107] Inhibition of peptidoglycan biosynthesis by omission of glycan substrate or by treatment with the antibiotic

bacitracin did not lead to significant decreases in the production of teichoic acids. This led the authors to posit the presence of multi-enzyme complexes capable of internally recycling undecaprenyl-phosphate after each biosynthetic cycle, rather than releasing it into the substrate pool after each round of synthesis.[107] Although this is only one study, the regulatory needs of prokaryotic and eukaryotic systems are different in that the three reactions requiring dolichyl-phosphate in eukaryotes are all involved in N-linked glycosylation, but in bacteria, enzymes require undecaprenyl-phosphate for separate biosynthetic ends. Further investigations are required to understand how bacteria manage multiple biosynthetic systems that require undecaprenyl-phosphate.

Enzyme specificity for polyprenols

Finally, a number of studies have shown that eukaryotic enzymes are specific for dolichol,[108-110] while bacterial enzymes are specific for undecaprenol, and many of these are summarized in the following reference.[111] Enzymes that utilize dolichol typically prefer S(-) dolichol to either R(+) dolichol or unsaturated polyprenol of a similar length indicating specificity for the configuration of the α-isoprene subunit.[112] In general, less specificity is found for polyisoprenol length. Fewer studies have examined the isoprene specificity of bacterial enzymes, but these have generally shown preferences for unsaturated undecaprenol over dolichols of similar length.[111] A recent comprehensive study examined three glycosyltransferases in the bacterial N-linked glycosylation pathway in *C. jejuni.* It was found that all three enzymes were specific for both the unsaturation and the *cis/trans* configurations of the isoprene units.[27] Interestingly, this was true even in the case of a soluble glycosyltransferase that is not predicted to contain any transmembrane domains. Enzyme specificity for polyisoprenols in prokaryotes

and eukaryotes is an important component of polyisoprenol regulation because it allows these molecules to be manipulated independently in the milieu of the phospholipid bilayer.

Why polyprenols? Importance in glycan "flipping"

Polyprenol-dependent pathways in bacteria, archaea, and eukaryotes rely universally on translocation or "flipping" of polyprenyl derivatives across lipid bilayers.[97, 110, 113-118] However very little is known about the mechanism of translocation or the role of enzyme catalysts and polyprenols in the process. Studies using a variety of biophysical and enzymatic-based methods have established that polyprenyl-phosphates and glycan modified polyprenols have very slow trans-bilayer diffusion rates with half-lives on the order of hours in model membranes.[79, 80, 119] As described above, *in vivo* studies have implied that polyprenyl-phosphate is typically generated through a recycling mechanism in which the polyprenyl-diphosphate released is hydrolyzed and then returned to the cytoplasmic face of the membrane (Figure 1-6). Biosynthetic rates of N-linked glycan or peptidoglycan production are such that relatively fast rates of polyprenyl-phosphate translocation are required with half-lives of less than one second.[86, 120] The difference between the observed and required flipping rates implies that catalysis, most likely enzyme-based, is required for the translocation of polyprenyl-phosphate derivatives.

Eukaryotic flippases

It has been long established that a variety of polyprenyl-phosphate derivatives are flipped across the ER membrane to facilitate N-linked glycosylation and related processes,[120] however, until the last decade, relatively little was known about the enzymes involved in this process. In the eukaryotic pathway, Dol-PP-$GlcNAc_2$-Man_5 is assembled on the cytoplasmic surface of the

ER membrane before it is transferred across the membrane to the luminal surface, where biosynthesis of the glycan is completed (Figure 1-3). The protein Rft1 was originally identified in the eukaryotic N-linked glycosylation pathway as the flippase responsible for the translocation of Dol-PP-$GlcNAc_2$-Man_5 to the ER lumen.[121] This study showed that deletion strains of Rft1 accumulated significant amounts of Dol-PP-$GlcNAc_2$-Man_5. However, more recent studies have demonstrated that flipping occurs in the absence of Rft1. Fractionated ER membrane samples were reconstituted into proteoliposomes and assayed for flippase activity. The translocation of Dol-PP-$GlcNAc_2$-Man_5 was not associated with the proteoliposomes that contained Rft1 or the glycerophospholipid flippase.[117] Furthermore, sealed microsomal vesicles from Rft1-depleted yeast showed translocation and elongation of the substrate at faster rates than wild type vesicles, providing further evidence that an enzyme other than Rft1 is involved in Dol-PP-$GlcNAc_2$-Man_5 translocation.[122]

In addition to Dol-PP-$GlcNAc_2$-Man_5, other dolichyl derivatives including dolichyl-phosphate, Dol-P-mannose, and Dol-P-glucose are translocated by eukaryotic cells (Figure 1-3). Very recently, Dol-P-Man flippase activity was identified in yeast. This activity could be separated from the Dol-PP-$GlcNAc_2$-Man_5 flippase activity, suggesting that multiple flippases specifically catalyze required translocation events for each of the required substrates.[115] This study also showed that citronellyl-phosphate, a short C10 dolichyl-phosphate analog, inhibited the flippase activity and the authors suggested that the enzyme may be also involved in translocation of dolichyl-phosphate.[115] In this model, the flippase acts as a pore through which the populations of Dol-P-mannose and dolichyl-phosphate equilibrate, with Dol-P-mannose translocating to the ER lumen to act as a glycan donor and dolichyl-phosphate returning to the cytoplasmic leaflet to begin a new cycle of glycan synthesis. This model would predict that the

cytoplasmic loops of the flippase enzyme might preferentially recognize the Dol-P-mannose, whereas the periplasmic loops may prefer dolichyl-phosphate, which would help facilitate the desired flow of molecules, but this model is highly speculative and remains to be tested.

Bacterial flippases

Similarly, in bacteria, it is believed that different flippases catalyze the translocation of the undecaprenyl-linked substrates in each pathway, although distinct flippases have not been characterized for all substrates. The genes encoding prokaryotic pathways are frequently found in clustered operons, which has allowed for facile identification of flippases for most undecaprenol-dependent pathways. Interestingly, it has been shown that two distinct enzyme families perform this function in prokaryotes.[113] The first class of flippases is the ABC-type transporters, which have ATPase activity and are associated with N- and O-linked protein glycosylation as well as the biosynthesis of homopolymeric O-antigen, which is polymerized in the cytoplasm before it is flipped to the periplasm (Figure 1-4). The second class of flippases, wzx-type translocases, shares homology with eukaryotic flippases and shows no dependence on ATP. Wzx-flippases are responsible for translocation of the individual subunits of peptidoglycan, heteropolymeric O-antigen, capsular polysaccharides and teichoic acids across the plasma membrane, which is followed by polymerization in the periplasm. It has been proposed that polymerization may be a driving force for uni-directional translocation, and thus ATP is not required in the latter type of flipping.[113] However, it is not clear how polymerization is different chemically from the reactions in the homopolymeric O-antigen, and N- and O-linked glycosylation pathways, since in all cases, the glycan substrates are transferred from the undecaprenyl carrier after translocation. One possibility is that periplasmic polymerization of

wzx-dependent undecaprenyl substrates requires higher concentrations of the precursors in the membrane and thus an equilibrative-based mechanism is efficient for facile transport of the substrates to the periplasm. In the case of homopolymeric O-antigen, and N- and O-liked glycosylation, the concentration of the undecaprenyl-linked substrates may be too low for this type of mechanism to be effective. Further work is required to understand the biological necessity of the two distinct mechanistic classes of flippases.

In general, the identification of most bacterial flippases has been simplified due to the presence of operons within bacterial genomes, although detailed biochemical characterization is still required in most cases. By comparison, discovering the Lipid II flippase in the peptidoglycan biosynthetic pathway has been a significant challenge, despite the large amount of research focused on this pathway during the last forty years. Recent studies suggested that either MviN or FtsW play a role in Lipid II translocation. *MviN* was identified via a bioinformatics approach, which sought to identify genes ubiquitous in peptidoglycan-containing bacteria, and *mviN* was shown to be essential for *E. coli* growth, although its function was not known.[123] However, *mviN* homologs were not essential for growth in Gram-positive *B. subtilis* implying that this protein is not directly responsible for translocation, since flipping of Lipid II is required for bacterial viability.[123] Alternatively, much debate has focused on the role of FtsW in translocation,[124, 125] but recently definitive biochemical studies demonstrated that the presence of purified FtsW in model membranes induced movement of Lipid II across the bilayer suggesting a role for FtsW in translocation.[114] It is still not known why MviN is essential for growth in Gram-negative bacteria.

Mechanism of translocation

Limited information is available about the mechanism of flippase-mediated translocation, because until recently, no biochemical methods for assaying translocation had been reported. Developing a biochemical assay for monitoring flippase activity has been quite challenging as it requires a model membrane system. In addition, flippase enzymes typically have many transmembrane domains and are relatively intractable to biochemical purification. However, in recent years, several important studies were published, which exploited lectin-binding, chemical reporter and FRET-based methods for quantifying the localization of polyprenyl-phosphate derivatives within membrane vesicles.[114, 115, 117, 122] With the exception of the fluorescence-based method, which requires a non-native glycan substrate, these methods are limited by the timescale of visualization since lipid-flipping occurs faster than the time resolution of the assay, which complicates mechanistic studies. In addition, no X-ray crystal structures of polyprenyl-diphosphate-glycan flippases have been reported, although the structure of a lipid flippase with homology to the ATP-binding cassette flippases has been recently published.[126]

A recent study utilized an unbiased truncation strategy to map the membrane topology of the Wzx flippase involved in O-antigen biosynthesis in *Psuedomonas aeruginosa.*[116] This study established that the flippase contained 12 transmembrane domains, similar to flippases from eukaryotic pathways, and contained several sizeable cytoplasmic loops, which may be involved in recognition of the O-antigen glycan portion of the substrate. In addition, the transmembrane helices included more charged residues than is typical for transmembrane domains. This observation prompted the authors to propose that these mostly cationic residues could line a channel formed by the transmembrane domains to facilitate movement of the O-linked glycan

across the membrane.[116] This study provides the basis for further mutational and structural studies focused on understanding how flippases mediate translocation.

It is important here to consider a second hypothesis for protein-mediated translocation that focuses on polyisoprenol recognition sequences (PIRS), which have been shown to selectively bind polyprenyl-linked substrates in model membranes.[82, 83] PIRS were first identified as dolichol recognition sequences (DRS) in the predicted transmembrane domains of three essential glycosyltransferases in the yeast N-linked glycosylation pathway and were postulated to interact with dolichol.[127] Subsequent studies identified PIRS in other essential eukaryotic glycosyltransferases and *E. coli* enzymes involved in capsular polysaccharide biosynthesis.[128, 129] The importance of these sequences is unclear as the PIRS was found to be essential for GlcNAc-1-phosphate transferase activity, but not for dolichyl-phosphate-mannose synthase activity.[130, 131]

Recent studies have utilized NMR and molecular modeling to study the binding of short PIRS peptides to C95-dolichol and undecaprenol and the related phosphorylated derivatives.[82, 83] These studies have formed the basis for the proposal that polyprenols and polyprenyl-phosphates adopt a coiled structure composed of three separate domains such that the head to tail length is shorter than the width of the membrane (32-33 Å for C95-dolichol and 22 Å for undecaprenol),[83] which is consistent with a prior study on the conformation of dolichol in a membrane bilayer (Figure 1-7).[81] Furthermore, conformational analysis of how polyprenols may interact with PIRS-based peptides suggested that multiple (4-5) peptides might be able to bind each polyprenol molecule (Figure 1-7). The authors propose that polyprenyl-phosphate may promote formation of a hydrophilic channel through recruitment of PIRS in nearby transmembrane domains.[82]



Figure 1-7. Structures of dolichol, dolichyl-phosphate and peptide-bound dolichol.
A, Dolichol structure was determined by SAXS analysis.[81] *B*, Dolichol (left) and dolichyl-phosphate (right) structures were determined by NMR and molecular modeling. The red group represents the phosphate.[83] *C*, Dolichol is shown bound to five PIRS-containing peptides as determined by computational modeling studies and NMR structural data of the peptide and dolichols structures.[82] All three images were adapted from previous structural studies in the following references.[81-83]

The flippase and PIRS hypotheses for glycan translocation are not mutually exclusive. The PIR-sequence is not well defined, but a key element of the peptide is the presence of a Pro residue midway through the sequence.[82, 83] Interestingly, Pro residues are rarely found in transmembrane domains, but are frequently found in the transmembrane domains of transporter-related enzymes.[132] Indeed, many of the flippase enzymes discussed above contain at least one transmembrane domain with a Pro residue. This could suggest the proposed PIRS-polyisoprenol interaction may be important for flippase activity, but further biochemical studies are necessary to elucidate the potential roles of these domains.

Why polyprenol? Models of macromolecular complexes

Enzyme complexes that utilize polyprenols

Several important multi-enzyme complexes that rely on translocated polyprenyl-linked substrates have been extensively characterized, such as the eukaryotic oligosaccharyltransferase (OT) and peptidoglycan biosynthetic machinery in bacteria. Although the role of polyprenols remains unclear, they are interesting complexes to consider in the context of polyprenols. The eukaryotic oligosaccharyltransferase (OT), which contains up to nine membrane-bound subunits,[18, 19, 133, 134] and is involved in the transfer of the tetradecasaccharide from dolichyl-diphosphate to protein. From the standpoint of polyisoprenols, very little is known about the role of dolichols in the interactions with the OT complex, although the Wbp1 subunit has been suggested to bind a dolichyl derivative.[135] Although the presence of dolichyl-based substrates has not been reported in purification of the OT complex, no high-resolution structures are available and dolichyl-phosphate derivatives could be removed during detergent solubilization. *In vitro* reconstitution of the OT complex within a membrane-mimetic system would be one way to unequivocally determine if dolichyl derivatives are essential for complex formation and stability; however, this is an extremely challenging experiment due to the complexity of the membrane-bound subunits.

In many ways, the eukaryotic OT complex is unique, because it is composed of multiple protein subunits acting in tandem to perform one reaction. Most other polyprenyl-dependent pathways involve multiple biosynthetic steps carried out by separate polypeptides. For example, 15-20 different proteins have been identified in the bacterial enzymatic machinery involved in peptidoglycan biosynthesis. Recent studies have identified essential protein-protein interactions that suggest the formation of both an elongase complex, responsible for lateral growth of rod-like

bacteria, and a distinct divisome complex, which mediates peptidoglycan synthesis at the site of cell division.[35] Interestingly, both of these complexes are associated with a cytoskeletal protein, which polymerizes at the cytoplasmic face of the cell surface and appears to recruit the required protein components. In cell elongation, MreB forms a helical pattern on the cell surface, whereas in cell division, FtsZ forms a ring at the center of rod-like cells to signal the site of division. Many studies have identified important protein binding partners for MreB and FtsZ. The elongase and divisome complexes require Lipid II to build peptidoglycan, but it is not known if Lipid II itself can act to help organize and recruit the elongase and divisome molecular machines. As discussed above, the relatively low pools of Lipid I and II found in bacterial cells, particularly Gram-negative bacteria, suggests that cells have found a way to sequester these substrates, but little is known about how this may occur.

Although many studies have elucidated models of the eukaryotic OT complex and peptidoglycan biosynthesis, essentially nothing is known about enzyme complex formation in other polyprenyl-dependent pathways. Typically separate enzymes carry out each step in the biosynthesis of the glycan substrate, translocation of the glycan substrate across the membrane and transfer of the glycan to its final substrate. Macromolecular complex formation would allow for efficient biosynthesis through sequential enzyme action. However, identification of the biosynthetic machinery is complicated since the protein-protein interactions are probably weaker than in the OT complex and are not associated with a cytoskeletal protein as in peptidoglycan biosynthesis.

Polyprenyl-dependent enzyme complexes

Polyprenyl-phosphate derivatives are key components of essential multi-enzyme biosynthetic pathways present in all living species. Macromolecular enzyme association is an attractive model for how pathways may perform efficient and sequential biosynthesis of polyprenyl-linked glycans (Figure 1-8). Furthermore, polyprenyl-linked substrates may act as more than passive glycan carriers by altering membrane fluidity and specifically promoting enzyme complex formation and substrate flux through the pathway. The hypothesis of macromolecular complexes is not new, but previous evidence is scant. As described above, examination of the interplay between teichoic acid and peptidoglycan biosynthesis in *Bacillus licheniformis* led the authors to suggest that the undecaprenyl-phosphate was sequestered from solution by separate biosynthetic complexes.[107] Further experimental evaluation of this hypothesis is challenging and may require *in vitro* reconstitution of a polyprenyl-based pathway in a membrane-mimetic system, and to our knowledge, no such experiments have been performed to date.

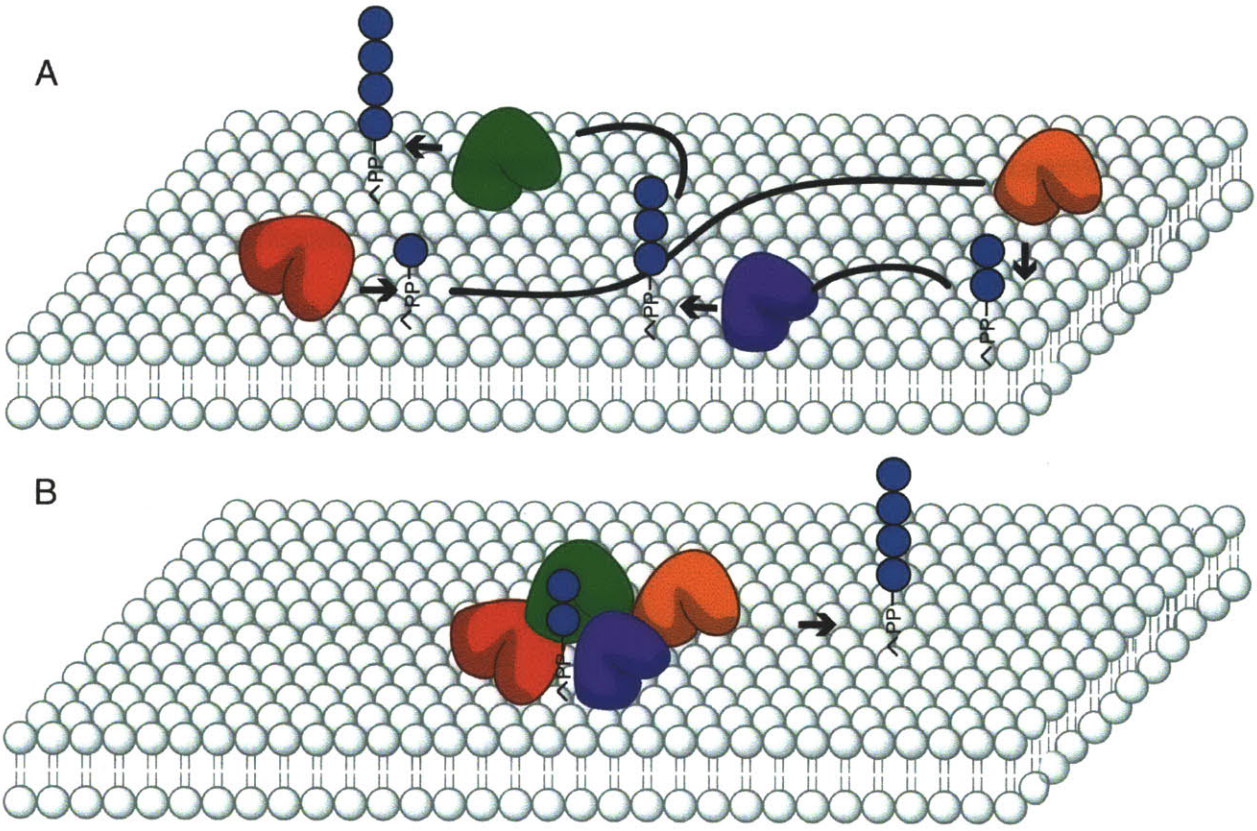


Figure 1-8. Independent and sequential models of polyprenol-dependent glycan biosynthesis. *A*, In the independent model, glycan intermediates are released after every glycosyltransferase reaction. This may result in long breaks before the intermediate encounters the next enzyme in the pathway. *B*, In the sequential model, the enzymes cluster together to produce the glycan efficiently without release of the intermediates to the lipid bilayer.

As described above, the identification of specific interactions between PIRS peptides and polyisoprenyl molecules was initially linked to polyprenyl-phosphate transolcation.[82, 83] A second interpretation, relevant here, is that the proposed ability of polyprenols to bind multiple PIRS domains is actually involved in facilitating protein-protein interactions within a biosynthetic pathway. This model suggests that polyprenols play an important role in recruitment of biosynthetic enzymes into a macromolecular complex, and could also have important implications for how the polyprenyl-linked glycan intermediates are shuttled through the pathway.

Recently, the complete biochemical characterization of enzymes participating in *Campylobacter jejuni* N-linked glycosylation has provided an exciting candidate for exploring enzyme-polyprenol complex formation. Pathway reconstitution in Nanodiscs, an *in vitro* membrane-mimetic system, would allow the role of undecaprenyl-phosphate in a membrane-associated glycan biosynthetic pathway to be further investigated. Chapters 5, 6 and 7 of this thesis describe our initial efforts to establish a model membrane system with *C. jejuni* N-linked glycosylation pathway to address these types of questions.

Conclusions

Polyprenols and polyprenyl-phosphates are closely related molecules, but knowledge about the function of these molecules in biological systems is surprisingly sparse. Much less is known about unmodified polyprenols, which act as precursors to polyprenyl-phosphates in developing cells, but have no other clearly established function in biological systems and may be unnecessary in Gram-negative bacteria. Furthermore, polyprenols can accumulate at relatively high levels within many eukaryotic organisms as well as Gram-positive bacteria; biosynthesis of these molecules is a complex energy-dependent process and it seems unlikely that cells would produce polyprenols needlessly. Further work is necessary to establish the physical effects that polyprenols may exert on complex biological membranes and how polyprenol accumulation affects aging, disease, and intracellular transport processes.

In contrast, the role of polyprenyl-phosphates as facilitators of glycan assembly in essential bioprocesses such as N-linked protein glycosylation and peptidoglycan biosynthesis has been thoroughly characterized. In addition, within these biosynthetic pathways, dolichyl-

phosphates may influence enzyme and pathway regulation, flippase-mediated glycan translocation across the lipid bilayer, and macromolecular enzyme complex formation. From these three hypotheses, it is interesting to consider which better explains the evolutionary conservation of polyprenols and dolichols. The first reason involving regulation of dolichyl-phosphates is well established in eukaryotes, and it is certainly important that dolichyl-phosphates have independent biosynthetic and recycling pathways from other membrane components. However, it is not clear that the long, complex structure of polyisoprenols is necessary for regulatory function, suggesting that this may not be a satisfying conclusion. In contrast, the ubiquitous use of polyprenols for translocation of glycans and their potential role in facilitating enzyme complex formation offer much more tantalizing implications. Both of these processes involve localization and movement within the lipid bilayer, processes that may be facilitated by the known biophysical effects exerted by polyisoprenols and by the specific interactions that may occur between polyisoprenols and certain peptide sequences. With the biochemical characterization of polyprenyl-dependent pathways and exciting new membrane mimetic technologies in place, the tools are available to elucidate the roles of polyprenyl-phosphates in glycan translocation and biosynthetic complex assembly.

Acknowledgements

I would like to offer my unending gratitude to Dr. Angelyn Larkin, who provided exceedingly helpful insight into this chapter.

References

1. Rosenwald, A. G.; Stoll, J.; Krag, S. S., Regulation of glycosylation. Three enzymes compete for a common pool of dolichyl phosphate in vivo. *J Biol Chem* **1990,** 265, (24), 14544-53.
2. Kabakoff, B. D.; Doyle, J. W.; Kandutsch, A. A., Relationships among dolichyl phosphate, glycoprotein synthesis, and cell culture growth. *Arch Biochem Biophys* **1990,** 276, (2), 382-9.
3. Krag, S. S., The importance of being dolichol. *Biochem Biophys Res Commun* **1998,** 243, (1), 1-5.
4. Lechner, J.; Wieland, F.; Sumper, M., Biosynthesis of sulfated saccharides N-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from *halobacteria. J. Biol. Chem.* **1985,** 260, (2), 860-6.
5. Stone, K. J.; Butterworth, P. H.; Hemming, F. W., Characterization of the hexahydropolyprenols of Aspergillus fumigatus Fresenius. *Biochem J* **1967,** 102, (2), 443-55.
6. Swiezewska, E.; Sasak, W.; Mankowski, T.; Jankowski, W.; Vogtman, T.; Krajewska, I.; Hertel, J.; Skoczylas, E.; Chojnacki, T., The search for plant polyprenols. *Acta Biochim. Pol.* **1994,** 41, (3), 221-60.
7. Swiezewska, E.; Chojnacki, T., The occurrence of unique, long-chain polyprenols in the leaves of *Potentilla* species. *Acta Biochim. Pol.* **1989,** 36, (2), 143-58.
8. Higashi, Y.; Strominger, J. L.; Sweeley, C. C., Biosynthesis of the peptidoglycan of bacterial cell walls. XXI. Isolation of free C55-isoprenoid alcohol and of lipid intermediates in peptidoglycan synthesis from Staphylococcus aureus. *J. Biol. Chem.* **1970,** 245, (14), 3697-702.
9. Wolucka, B. A.; de Hoffmann, E., Isolation and characterization of the major form of polyprenyl-phospho-mannose from Mycobacterium smegmatis. *Glycobiology* **1998,** 8, (10), 955-62.
10. Wolucka, B. A.; McNeil, M. R.; de Hoffmann, E.; Chojnacki, T.; Brennan, P. J., Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. *J Biol Chem* **1994,** 269, (37), 23328-35.
11. Adair, W. L., Jr.; Cafmeyer, N., Characterization of the Saccharomyces cerevisiae cis-prenyltransferase required for dolichyl phosphate biosynthesis. *Arch Biochem Biophys* **1987,** 259, (2), 589-96.
12. Eggens, I.; Chojnacki, T.; Kenne, L.; Dallner, G., Separation, quantitation and distribution of dolichol and dolichyl phosphate in rat and human tissues. *Biochim. Biophys. Acta* **1983,** 751, (3), 355-68.
13. Stone, K. J.; Wellburn, A. R.; Hemming, F. W.; Pennock, J. F., The characterization of ficaprenol-10, -11 and 12 from the leaves of Ficus elastica (decorative rubber plant). *Biochem J* **1967,** 102, (1), 325-30.
14. Jones, M. B.; Rosenberg, J. N.; Betenbaugh, M. J.; Krag, S. S., Structure and synthesis of polyisoprenoids used in N-glycosylation across the three domains of life. *Biochim. Biophys. Acta* **2009,** 1790, (6), 485-94.

15. Bouhss, A.; Trunkfield, A. E.; Bugg, T. D.; Mengin-Lecreulx, D., The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol. Rev.* **2008,** 32, (2), 208-33.
16. Weerapana, E.; Imperiali, B., Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **2006,** 16, (6), 91R-101R.
17. Burda, P.; Aebi, M., The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta* **1999,** 1426, (2), 239-57.
18. Dempski, R. E., Jr.; Imperiali, B., Oligosaccharyl transferase: gatekeeper to the secretory pathway. *Curr. Opin. Chem. Biol.* **2002,** 6, (6), 844-50.
19. Kelleher, D. J.; Gilmore, R., An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **2006,** 16, (4), 47R-62R.
20. Calo, D.; Kaminski, L.; Eichler, J., Protein glycosylation in Archaea: sweet and extreme. *Glycobiology* **2010,** 20, (9), 1065-76.
21. Jarrell, K. F.; Jones, G. M.; Kandiba, L.; Nair, D. B.; Eichler, J., S-layer glycoproteins and flagellins: reporters of archaeal posttranslational modifications. *Archaea* **2010,** 2010.
22. Szymanski, C. M.; Yao, R.; Ewing, C. P.; Trust, T. J.; Guerry, P., Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **1999,** 32, (5), 1022-30.
23. Wacker, M.; Linton, D.; Hitchen, P. G.; Nita-Lazar, M.; Haslam, S. M.; North, S. J.; Panico, M.; Morris, H. R.; Dell, A.; Wren, B. W.; Aebi, M., N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **2002,** 298, (5599), 1790-3.
24. Kumar, M.; Balaji, P. V., Comparative genomics analysis of completely sequenced microbial genomes reveals the ubiquity of N-linked glycosylation in prokaryotes. *Mol. Biosyst.* **2011**.
25. Troutman, J. M.; Imperiali, B., *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **2009,** 48, (12), 2807-16.
26. Chen, M. M.; Glover, K. J.; Imperiali, B., From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* **2007,** 46, (18), 5579-85.
27. Chen, M. M.; Weerapana, E.; Ciepichal, E.; Stupak, J.; Reid, C. W.; Swiezewska, E.; Imperiali, B., Polyisoprenol specificity in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2007,** 46, (50), 14342-8.
28. Glover, K. J.; Weerapana, E.; Chen, M. M.; Imperiali, B., Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2006,** 45, (16), 5343-50.
29. Olivier, N. B.; Chen, M. M.; Behr, J. R.; Imperiali, B., *In vitro* biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* **2006,** 45, (45), 13659-69.
30. Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. U.S.A.* **2005,** 102, (40), 14255-9.
31. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* **2005,** 12, (12), 1311-5.

32. Vik, A.; Aas, F. E.; Anonsen, J. H.; Bilsborough, S.; Schneider, A.; Egge-Jacobsen, W.; Koomey, M., Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. U S A* **2009,** 106, (11), 4447-52.
33. Aas, F. E.; Vik, A.; Vedde, J.; Koomey, M.; Egge-Jacobsen, W., *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol. Microbiol.* **2007,** 65, (3), 607-24.
34. Hegge, F. T.; Hitchen, P. G.; Aas, F. E.; Kristiansen, H.; Lovold, C.; Egge-Jacobsen, W.; Panico, M.; Leong, W. Y.; Bull, V.; Virji, M.; Morris, H. R.; Dell, A.; Koomey, M., Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of *Neisseria gonorrhoeae* type IV pili. *Proc. Natl. Acad. Sci. U S A* **2004,** 101, (29), 10798-803.
35. den Blaauwen, T.; de Pedro, M. A.; Nguyen-Disteche, M.; Ayala, J. A., Morphogenesis of rod-shaped sacculi. *FEMS Microbiol. Rev.* **2008,** 32, (2), 321-44.
36. de Kruijff, B.; van Dam, V.; Breukink, E., Lipid II: a central component in bacterial cell wall synthesis and a target for antibiotics. *Prostaglandins Leukot. Essent. Fatty Acids* **2008,** 79, (3-5), 117-21.
37. Samuel, G.; Reeves, P., Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr. Res.* **2003,** 338, (23), 2503-19.
38. Wright, A.; Dankert, M.; Fennessey, P.; Robbins, P. W., Characterization of a polyisoprenoid compound functional in O-antigen biosynthesis. *Proc. Natl. Acad. Sci. U S A* **1967,** 57, (6), 1798-803.
39. Mancuso, D. J.; Chiu, T. H., Biosynthesis of glucosyl monophosphoryl undecaprenol and its role in lipoteichoic acid biosynthesis. *J. Bacteriol.* **1982,** 152, (2), 616-25.
40. Whitfield, C., Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annu Rev Biochem* **2006,** 75, 39-68.
41. Troy, F. A.; Vijay, I. K.; Tesche, N., Role of undecaprenyl phosphate in synthesis of polymers containing sialic acid in *Escherichia coli*. *J. Biol. Chem.* **1975,** 250, (1), 156-63.
42. Masson, L.; Holbein, B. E., Role of lipid intermediate(s) in the synthesis of serogroup B *Neisseria meningitidis* capsular polysaccharide. *J. Bacteriol.* **1985,** 161, (3), 861-7.
43. Lommel, M.; Strahl, S., Protein O-mannosylation: conserved from bacteria to humans. *Glycobiology* **2009,** 19, (8), 816-28.
44. Furmanek, A.; Hofsteenge, J., Protein C-mannosylation: facts and questions. *Acta Biochim Pol* **2000,** 47, (3), 781-9.
45. Fujita, M.; Kinoshita, T., Structural remodeling of GPI anchors during biosynthesis and after attachment to proteins. *FEBS Lett* **2010,** 584, (9), 1670-7.
46. Mescher, M. F.; Hansen, U.; Strominger, J. L., Formation of lipid-linked sugar compounds in *Halobacterium salinarium*. Presumed intermediates in glycoprotein synthesis. *J. Biol. Chem.* **1976,** 251, (23), 7289-94.
47. Kuntz, C.; Sonnenbichler, J.; Sonnenbichler, I.; Sumper, M.; Zeitler, R., Isolation and characterization of dolichol-linked oligosaccharides from *Haloferax volcanii*. *Glycobiology* **1997,** 7, (7), 897-904.
48. Zhu, B. C.; Laine, R. A., Dolichyl-phosphomannose synthase from the archae Thermoplasma acidophilum. *Glycobiology* **1996,** 6, (8), 811-6.
49. Wieland, F.; Dompert, W.; Bernhardt, G.; Sumper, M., Halobacterial glycoprotein saccharides contain covalently linked sulphate. *FEBS Lett* **1980,** 120, (1), 110-4.

50. Tollbom, O.; Valtersson, C.; Chojnacki, T.; Dallner, G., Esterification of dolichol in rat liver. *J. Biol. Chem.* **1988,** 263, (3), 1347-52.
51. Chojnacki, T.; Dallner, G., The biological role of dolichol. *Biochem J* **1988,** 251, (1), 1-9.
52. Keller, R. K.; Nellis, S. W., Quantitation of dolichyl phosphate and dolichol in major organs of the rat as a function of age. *Lipids* **1986,** 21, (5), 353-5.
53. Keller, R. K.; Fuller, M. S.; Rottler, G. D.; Connelly, L. W., Extraction of dolichyl phosphate and its quantitation by straight-phase high-performance liquid chromatography. *Anal. Biochem.* **1985,** 147, (1), 166-73.
54. Stoll, J.; Rosenwald, A. G.; Krag, S. S., A Chinese hamster ovary cell mutant F2A8 utilizes polyprenol rather than dolichol for its lipid-dependent asparagine-linked glycosylation reactions. *J. Biol. Chem.* **1988,** 263, (22), 10774-82.
55. Tavares, I. A.; Coolbear, T.; Hemming, F. W., Increased hepatic dolichol and dolichyl phosphate-mediated glycosylation in rats fed cholesterol. *Arch. Biochem. Biophys.* **1981,** 207, (2), 427-36.
56. Ekstrom, T. J.; Chojnacki, T.; Dallner, G., Metabolic labeling of dolichol and dolichyl phosphate in isolate hepatocytes. *J. Biol. Chem.* **1984,** 259, (16), 10460-8.
57. Chojnacki, T.; Vogtman, T., The occurrence and seasonal distribution of C50-C60-polyprenols and of C100-and similar long-chain polyprenols in leaves of plants. *Acta Biochim. Pol.* **1984,** 31, (1), 115-26.
58. Pullarkat, R. K.; Reha, H., Accumulation of dolichols in brains of elderly. *J Biol Chem* **1982,** 257, (11), 5991-3.
59. Andersson, M.; Appelkvist, E. L.; Kristensson, K.; Dallner, G., Distribution of dolichol and dolichyl phosphate in human brain. *J Neurochem* **1987,** 49, (3), 685-91.
60. Andersson, M.; Appelkvist, E. L.; Kristensson, K.; Dallner, G., Age-dependent changes in the levels of dolichol and dolichyl phosphates in human brain. *Acta Chem Scand B* **1987,** 41, (2), 144-6.
61. Olsson, J. M.; Schedin, S.; Teclebrhan, H.; Eriksson, L. C.; Dallner, G., Enzymes of the mevalonate pathway in rat liver nodules induced by 2-acetylaminofluorene treatment. *Carcinogenesis* **1995,** 16, (3), 599-605.
62. Olsson, J. M.; Eriksson, L. C.; Dallner, G., Lipid compositions of intracellular membranes isolated from rat liver nodules in Wistar rats. *Cancer Res* **1991,** 51, (14), 3774-80.
63. Eggens, I.; Ericsson, J.; Tollbom, O., Cytidine 5'-triphosphate-dependent dolichol kinase and dolichol phosphatase activities and levels of dolichyl phosphate in microsomal fractions from highly differentiated human hepatomas. *Cancer Res* **1988,** 48, (12), 3418-24.
64. Gajjar, D.; Jozwiak, A.; Swiezewska, E.; Alapure, B.; Parmar, T.; Johar, K.; Vasavada, A. R., Quantification of dolichol in the human lens with different types of cataracts. *Mol Vis* **2009,** 15, 1573-9.
65. Ng Ying Kin, N. M.; Palo, J.; Haltia, M.; Wolfe, L. S., High levels of brain dolichols in neuronal ceroid-lipofuscinosis and senescence. *J Neurochem* **1983,** 40, (5), 1465-73.
66. Keller, R. K.; Armstrong, D.; Crum, F. C.; Koppang, N., Dolichol and dolichyl phosphate levels in brain tissue from English setters with ceroid lipofuscinosis. *J Neurochem* **1984,** 42, (4), 1040-7.
67. Penuelas, J.; Munne-Bosch, S., Isoprenoids: an evolutionary pool for photoprotection. *Trends Plant Sci* **2005,** 10, (4), 166-9.

68. Loreto, F.; Pinelli, P.; Manes, F.; Kollist, H., Impact of ozone on monoterpene emissions and evidence for an isoprene-like antioxidant action of monoterpenes emitted by Quercus ilex leaves. *Tree Physiol* **2004,** 24, (4), 361-7.
69. Bergamini, E.; Bizzarri, R.; Cavallini, G.; Cerbai, B.; Chiellini, E.; Donati, A.; Gori, Z.; Manfrini, A.; Parentini, I.; Signori, F.; Tamburini, I., Ageing and oxidative stress: a role for dolichol in the antioxidant machinery of cell membranes? *J Alzheimers Dis* **2004,** 6, (2), 129-35.
70. Wong, T. K.; Decker, G. L.; Lennarz, W. J., Localization of dolichol in the lysosomal fraction of rat liver. *J. Biol. Chem.* **1982,** 257, (11), 6614-8.
71. Sato, M.; Sato, K.; Nishikawa, S.; Hirata, A.; Kato, J.; Nakano, A., The yeast RER2 gene, identified by endoplasmic reticulum protein localization mutations, encodes cis-prenyltransferase, a key enzyme in dolichol synthesis. *Mol Cell Biol* **1999,** 19, (1), 471-83.
72. Belgareh-Touze, N.; Corral-Debrinski, M.; Launhardt, H.; Galan, J. M.; Munder, T.; Le Panse, S.; Haguenauer-Tsapis, R., Yeast functional analysis: identification of two essential genes involved in ER to Golgi trafficking. *Traffic* **2003,** 4, (9), 607-17.
73. Wang, X.; Mansourian, A. R.; Quinn, P. J., The effect of dolichol on the structure and phase behaviour of phospholipid model membranes. *Mol Membr Biol* **2008,** 25, (6-7), 547-56.
74. van Duijn, G.; Valtersson, C.; Chojnacki, T.; Verkleij, A. J.; Dallner, G.; de Kruijff, B., Dolichyl phosphate induces non-bilayer structures, vesicle fusion and transbilayer movement of lipids: a model membrane study. *Biochim Biophys Acta* **1986,** 861, (2), 211-23.
75. Valtersson, C.; van Duyn, G.; Verkleij, A. J.; Chojnacki, T.; de Kruijff, B.; Dallner, G., The influence of dolichol, dolichol esters, and dolichyl phosphate on phospholipid polymorphism and fluidity in model membranes. *J Biol Chem* **1985,** 260, (5), 2742-51.
76. Knudsen, M. J.; Troy, F. A., Nuclear magnetic resonance studies of polyisoprenols in model membranes. *Chem Phys Lipids* **1989,** 51, (3-4), 205-12.
77. Janas, T.; Chojnacki, T.; Swiezewska, E., The effect of undecaprenol on bilayer lipid membranes. *Acta Biochim Pol* **1994,** 41, (3), 351-8.
78. de Ropp, J. S.; Troy, F. A., 2H NMR investigation of the organization and dynamics of polyisoprenols in membranes. *J Biol Chem* **1985,** 260, (29), 15669-74.
79. McCloskey, M. A.; Troy, F. A., Paramagnetic isoprenoid carrier lipids. 2. Dispersion and dynamics in lipid membranes. *Biochemistry* **1980,** 19, (10), 2061-6.
80. McCloskey, M. A.; Troy, F. A., Paramagnetic isoprenoid carrier lipids. 1. Chemical synthesis and incorporation into model membranes. *Biochemistry* **1980,** 19, (10), 2056-60.
81. Murgolo, N. J.; Patel, A.; Stivala, S. S.; Wong, T. K., The conformation of dolichol. *Biochemistry* **1989,** 28, (1), 253-60.
82. Zhou, G. P.; Troy, F. A., 2nd, NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure. *Glycobiology* **2005,** 15, (4), 347-59.
83. Zhou, G. P.; Troy, F. A., 2nd, Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions. *Glycobiology* **2003,** 13, (2), 51-71.
84. Barreteau, H.; Magnet, S.; El Ghachi, M.; Touze, T.; Arthur, M.; Mengin-Lecreulx, D.; Blanot, D., Quantitative high-performance liquid chromatography analysis of the pool levels of undecaprenyl phosphate and its derivatives in bacterial membranes. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2009,** 877, (3), 213-20.
85. Umbreit, J. N.; Stone, K. J.; Strominger, J. L., Isolation of polyisoprenyl alcohols from *Streptococcus faecalis*. *J. Bacteriol.* **1972,** 112, (3), 1302-5.

86. van Heijenoort, Y.; Gomez, M.; Derrien, M.; Ayala, J.; van Heijenoort, J., Membrane intermediates in the peptidoglycan metabolism of *Escherichia coli*: possible roles of PBP 1b and PBP 3. *J. Bacteriol.* **1992,** 174, (11), 3549-57.
87. Fuchs-Cleveland, E.; Gilvarg, C., Oligomeric intermediate in peptidoglycan biosynthesis in *Bacillus megaterium*. *Proc. Natl. Acad. Sci. U. S. A.* **1976,** 73, (11), 4200-4.
88. Helm, J. S.; Chen, L.; Walker, S., Rethinking ramoplanin: the role of substrate binding in inhibition of peptidoglycan biosynthesis. *J Am Chem Soc* **2002,** 124, (47), 13970-1.
89. Somner, E. A.; Reynolds, P. E., Inhibition of peptidoglycan biosynthesis by ramoplanin. *Antimicrob. Agents Chemother.* **1990,** 34, (3), 413-9.
90. Brotz, H.; Bierbaum, G.; Leopold, K.; Reynolds, P. E.; Sahl, H. G., The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II. *Antimicrob. Agents Chemother.* **1998,** 42, (1), 154-60.
91. Spiro, M. J.; Spiro, R. G., Control of N-linked carbohydrate unit synthesis in thyroid endoplasmic reticulum by membrane organization and dolichyl phosphate availability. *J Biol Chem* **1986,** 261, (31), 14725-32.
92. Lucas, J. J.; Levin, E., Increase in the lipid intermediate pathway of protein glycosylation during hen oviduct differentiation. *J Biol Chem* **1977,** 252, (12), 4330-6.
93. Carson, D. D.; Lennarz, W. J., Relationship of dolichol synthesis to glycoprotein synthesis during embryonic development. *J Biol Chem* **1981,** 256, (9), 4679-86.
94. Carson, D. D.; Lennarz, W. J., Inhibition of polyisoprenoid and glycoprotein biosynthesis causes abnormal embryonic development. *Proc Natl Acad Sci U S A* **1979,** 76, (11), 5709-13.
95. Rossignol, D. P.; Scher, M.; Waechter, C. J.; Lennarz, W. J., Metabolic interconversion of dolichol and dolichyl phosphate during development of the sea urchin embryo. *J Biol Chem* **1983,** 258, (15), 9122-7.
96. Rossignol, D. P.; Lennarz, W. J.; Waechter, C. J., Induction of phosphorylation of dolichol during embryonic development of the sea urchin. *J Biol Chem* **1981,** 256, (20), 10538-42.
97. Rush, J. S.; Gao, N.; Lehrman, M. A.; Waechter, C. J., Recycling of dolichyl monophosphate to the cytoplasmic leaflet of the endoplasmic reticulum after the cleavage of dolichyl pyrophosphate on the lumenal monolayer. *J Biol Chem* **2008,** 283, (7), 4087-93.
98. Haeuptle, M. A.; Hennet, T., Congenital disorders of glycosylation: an update on defects affecting the biosynthesis of dolichol-linked oligosaccharides. *Hum Mutat* **2009,** 30, (12), 1628-41.
99. Kranz, C.; Jungeblut, C.; Denecke, J.; Erlekotte, A.; Sohlbach, C.; Debus, V.; Kehl, H. G.; Harms, E.; Reith, A.; Reichel, S.; Grobe, H.; Hammersen, G.; Schwarzer, U.; Marquardt, T., A defect in dolichol phosphate biosynthesis causes a new inherited disorder with death in early infancy. *Am J Hum Genet* **2007,** 80, (3), 433-40.
100. Cantagrel, V.; Lefeber, D. J.; Ng, B. G.; Guan, Z.; Silhavy, J. L.; Bielas, S. L.; Lehle, L.; Hombauer, H.; Adamowicz, M.; Swiezewska, E.; De Brouwer, A. P.; Blumel, P.; Sykut-Cegielska, J.; Houliston, S.; Swistun, D.; Ali, B. R.; Dobyns, W. B.; Babovic-Vuksanovic, D.; van Bokhoven, H.; Wevers, R. A.; Raetz, C. R.; Freeze, H. H.; Morava, E.; Al-Gazali, L.; Gleeson, J. G., SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* **2010,** 142, (2), 203-17.
101. Jerga, A.; Lu, Y. J.; Schujman, G. E.; de Mendoza, D.; Rock, C. O., Identification of a soluble diacylglycerol kinase required for lipoteichoic acid production in Bacillus subtilis. *J. Biol. Chem.* **2007,** 282, (30), 21738-45.

102. Lis, M.; Kuramitsu, H. K., The stress-responsive *dgk* gene from *Streptococcus mutans* encodes a putative undecaprenol kinase activity. *Infect. Immun.* **2003,** 71, (4), 1938-43.
103. Kalin, J. R.; Allen, C. M., Lipid activation of undecaprenol kinase from *Lactobacillus plantarum*. *Biochim. Biophys. Acta* **1980,** 619, (1), 76-89.
104. Kalin, J. R.; Allen, C. M., Jr., Characterization of undecaprenol kinase from *Lactobacillus plantarum*. *Biochim. Biophys. Acta.* **1979,** 574, (1), 112-22.
105. Higashi, Y.; Siewert, G.; Strominger, J. L., Biosynthesis of the peptidoglycan of bacterial cell walls. XIX. Isoprenoid alcohol phosphokinase. *J. Biol. Chem.* **1970,** 245, (14), 3683-90.
106. El Ghachi, M.; Bouhss, A.; Blanot, D.; Mengin-Lecreulx, D., The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. *J. Biol. Chem.* **2004,** 279, (29), 30106-13.
107. Anderson, R. G.; Hussey, H.; Baddiley, J., The mechanism of wall synthesis in bacteria. The organization of enzymes and isoprenoid phosphates in the membrane. *Biochem. J.* **1972,** 127, (1), 11-25.
108. Low, P.; Peterson, E.; Mizuno, M.; Takigawa, T.; Chojnacki, T.; Dallner, G., Reaction of optically active S- and R-forms of dolichyl phosphates with activated sugars. *Biochem Biophys Res Commun* **1985,** 130, (1), 460-6.
109. Palamarczyk, G.; Lehle, L.; Mankowski, T.; Chojnacki, T.; Tanner, W., Specificity of solubilized yeast glycosyl transferases for polyprenyl derivatives. *Eur J Biochem* **1980,** 105, (3), 517-23.
110. Rush, J. S.; van Leyen, K.; Ouerfelli, O.; Wolucka, B.; Waechter, C. J., Transbilayer movement of Glc-P-dolichol and its function as a glucosyl donor: protein-mediated transport of a water-soluble analog into sealed ER vesicles from pig brain. *Glycobiology* **1998,** 8, (12), 1195-205.
111. Rush, J. S.; Rick, P. D.; Waechter, C. J., Polyisoprenyl phosphate specificity of UDP-GlcNAc:undecaprenyl phosphate N-acetylglucosaminyl 1-P transferase from E.coli. *Glycobiology* **1997,** 7, (2), 315-22.
112. Chamot-Rooke, J.; Mikaty, G.; Malosse, C.; Soyer, M.; Dumont, A.; Gault, J.; Imhaus, A. F.; Martin, P.; Trellet, M.; Clary, G.; Chafey, P.; Camoin, L.; Nilges, M.; Nassif, X.; Dumenil, G., Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination. *Science* **2011,** 331, (6018), 778-82.
113. Alaimo, C.; Catrein, I.; Morf, L.; Marolda, C. L.; Callewaert, N.; Valvano, M. A.; Feldman, M. F.; Aebi, M., Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J.* **2006,** 25, (5), 967-76.
114. Mohammadi, T.; van Dam, V.; Sijbrandi, R.; Vernet, T.; Zapun, A.; Bouhss, A.; Diepeveen-de Bruin, M.; Nguyen-Disteche, M.; de Kruijff, B.; Breukink, E., Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J.* **2011.**
115. Sanyal, S.; Menon, A. K., Stereoselective transbilayer translocation of mannosyl phosphoryl dolichol by an endoplasmic reticulum flippase. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** 107, (25), 11289-94.
116. Islam, S. T.; Taylor, V. L.; Qi, M.; Lam, J. S., Membrane Topology Mapping of the O-Antigen Flippase (Wzx), Polymerase (Wzy), and Ligase (WaaL) from Pseudomonas aeruginosa PAO1 Reveals Novel Domain Architectures. *MBio* **2010,** 1, (3).
117. Sanyal, S.; Frank, C. G.; Menon, A. K., Distinct flippases translocate glycerophospholipids and oligosaccharide diphosphate dolichols across the endoplasmic reticulum. *Biochemistry* **2008,** 47, (30), 7937-46.

118. Power, P. M.; Seib, K. L.; Jennings, M. P., Pilin glycosylation in *Neisseria meningitidis* occurs by a similar pathway to wzy-dependent O-antigen biosynthesis in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **2006,** 347, (4), 904-8.
119. Hanover, J. A.; Lennarz, W. J., The topological orientation of N,N'-diacetylchitobiosylpyrophosphoryldolichol in artificial and natural membranes. *J Biol Chem* **1979,** 254, (18), 9237-46.
120. Lennarz, W. J., Protein glycosylation in the endoplasmic reticulum: current topological issues. *Biochemistry* **1987,** 26, (23), 7205-10.
121. Helenius, J.; Ng, D. T.; Marolda, C. L.; Walter, P.; Valvano, M. A.; Aebi, M., Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* **2002,** 415, (6870), 447-50.
122. Rush, J. S.; Gao, N.; Lehrman, M. A.; Matveev, S.; Waechter, C. J., Suppression of Rft1 expression does not impair the transbilayer movement of Man5GlcNAc2-P-P-dolichol in sealed microsomes from yeast. *J. Biol. Chem.* **2009,** 284, (30), 19835-42.
123. Ruiz, N., Bioinformatics identification of MurJ (MviN) as the peptidoglycan lipid II flippase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2008,** 105, (40), 15553-7.
124. Ehlert, K.; Holtje, J. V., Role of precursor translocation in coordination of murein and phospholipid synthesis in *Escherichia coli*. *J. Bacteriol.* **1996,** 178, (23), 6766-71.
125. Lara, B.; Mengin-Lecreulx, D.; Ayala, J. A.; van Heijenoort, J., Peptidoglycan precursor pools associated with MraY and FtsW deficiencies or antibiotic treatments. *FEMS Microbiol. Lett.* **2005,** 250, (2), 195-200.
126. Ward, A.; Reyes, C. L.; Yu, J.; Roth, C. B.; Chang, G., Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc Natl Acad Sci U S A* **2007,** 104, (48), 19005-10.
127. Albright, C. F.; Orlean, P.; Robbins, P. W., A 13-amino acid peptide in three yeast glycosyltransferases may be involved in dolichol recognition. *Proc Natl Acad Sci U S A* **1989,** 86, (19), 7366-9.
128. Pavelka, M. S., Jr.; Wright, L. F.; Silver, R. P., Identification of two genes, kpsM and kpsT, in region 3 of the polysialic acid gene cluster of Escherichia coli K1. *J Bacteriol* **1991,** 173, (15), 4603-10.
129. Zhu, X. Y.; Lehrman, M. A., Cloning, sequence, and expression of a cDNA encoding hamster UDP-GlcNAc:dolichol phosphate N-acetylglucosamine-1-phosphate transferase. *J Biol Chem* **1990,** 265, (24), 14250-5.
130. Datta, A. K.; Lehrman, M. A., Both potential dolichol recognition sequences of hamster GlcNAc-1-phosphate transferase are necessary for normal enzyme function. *J Biol Chem* **1993,** 268, (17), 12663-8.
131. Zimmerman, J. W.; Robbins, P. W., The hydrophobic domain of dolichyl-phosphate-mannose synthase is not essential for enzyme activity or growth in Saccharomyces cerevisiae. *J Biol Chem* **1993,** 268, (22), 16746-53.
132. Brandl, C. J.; Deber, C. M., Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc Natl Acad Sci U S A* **1986,** 83, (4), 917-21.
133. Lennarz, W. J., Studies on oligosaccharyl transferase in yeast. *Acta Biochim Pol* **2007,** 54, (4), 673-7.
134. Knauer, R.; Lehle, L., The oligosaccharyltransferase complex from yeast. *Biochim Biophys Acta* **1999,** 1426, (2), 259-73.
135. Pathak, R.; Hendrickson, T. L.; Imperiali, B., Sulfhydryl modification of the yeast Wbp1p inhibits oligosaccharyl transferase activity. *Biochemistry* **1995,** 34, (13), 4179-85.

Chapter 2: Chemoenzymatic synthesis of polyprenyl-phosphates and polyprenyl-phosphate derivatives

Most of the work described in this chapter was previously published in the following reference:

Hartley, M. D.; Larkin, A.; Imperiali, B., Chemoenzymatic synthesis of polyprenyl phosphates. *Bioorg. Med. Chem.* **2008,** 16, (9), 5149-5156.

Dr. Angelyn Larkin performed the PglB assay described in this chapter and collaborated on several aspects of optimizing the synthesis of undecaprenyl-diphosphate disaccharide.

Introduction

Polyprenyl-phosphates, present in the membranes of all organisms, are critical components in several prokaryotic and eukaryotic processes. Eukaryotic organisms utilize dolichyl-phosphate (Dol-P, Figure 2-1, **4**), which contains 18-20 isoprene subunits, as an oligosaccharide carrier in N-linked protein glycosylation.[1] A shorter dolichyl-phosphate with 11-12 isoprene units is suggested to play a similar role in the corresponding glycosylation pathways in certain archaea.[2] Furthermore, in recent years, undecaprenyl-phosphate (Und-P, Figure 2-1, **2**) has been implicated in bacterial protein glycosylation pathways, such as the N-linked protein glycosylation (Pgl) pathway in *Campylobacter jejuni*[3-6] and the O-linked protein glycosylation (Pgl) pathway in *Neisseria gonorrhoeae* (Figure 2-2).[7, 8] Undecaprenyl-phosphate comprises 11 isoprene units and unlike dolichol, the α-subunit is unsaturated. Importantly, undecaprenyl-phosphate is also required for peptidoglycan biosynthesis, which is essential to bacterial viability.[9] It is also integral to the biosynthesis of capsular polysaccharides and lipopolysaccharide O-antigens in Gram-negative bacteria and wall teichoic acids in Gram-positive bacteria.[10]

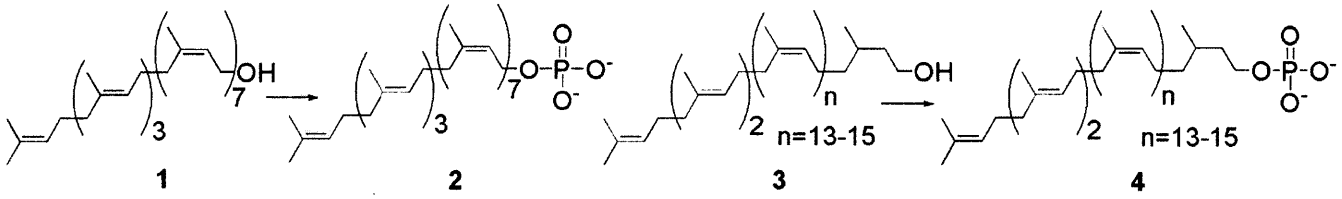


Figure 2-1. Polyprenol phosphorylation reactions.
Undecaprenol (**1**) and dolichol (**3**) can be phosphorylated to form undecaprenyl-phosphate (**2**) and dolichyl-phosphate (**4**), respectively. This reaction can be done synthetically or chemoenzymatically by the action of a kinase, which can donate the γ-phosphate of ATP.

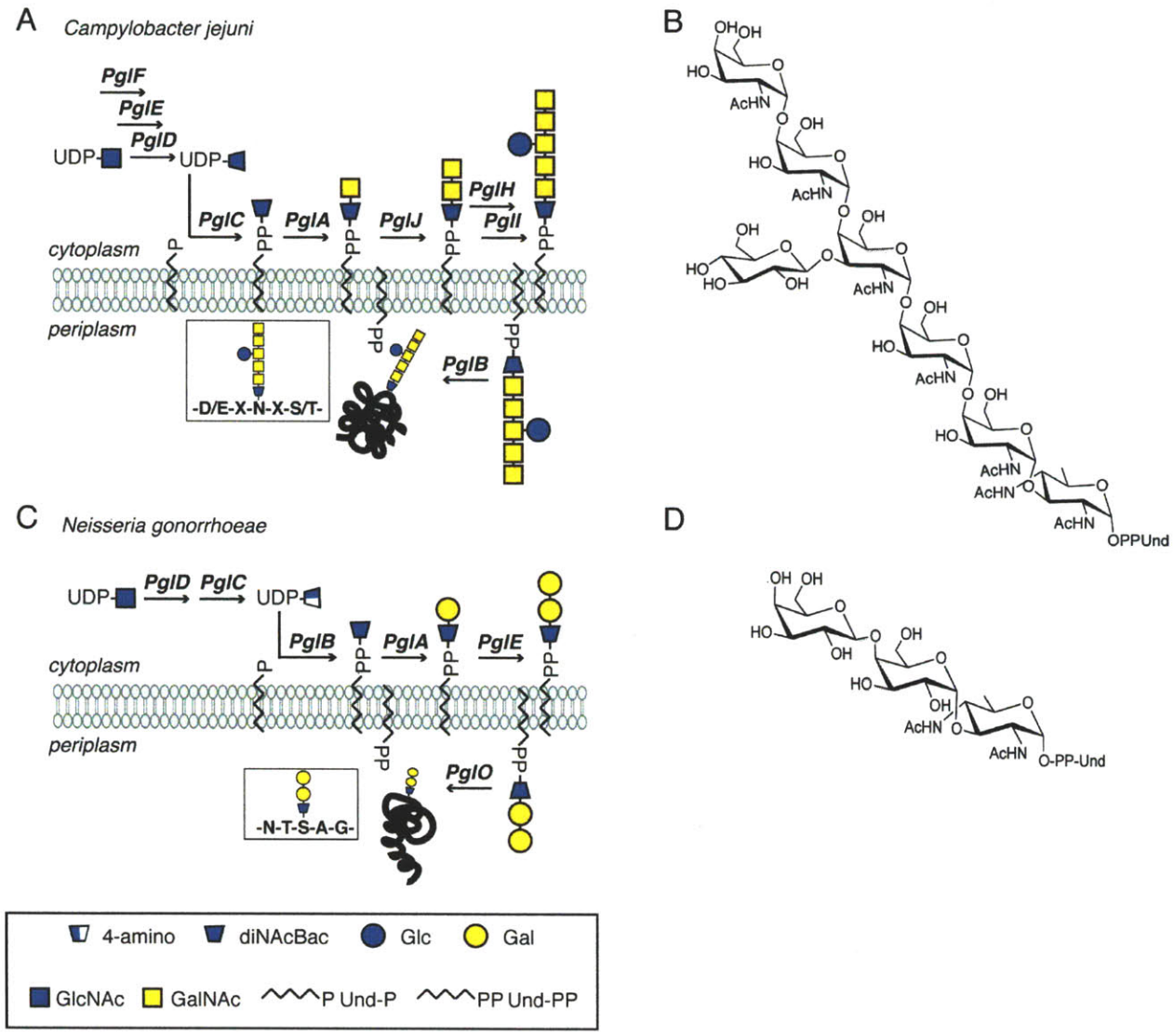


Figure 2-2. Undecaprenyl-dependent protein glycosylation pathways.
A, *C. jejuni* N-linked protein glycosylation pathway. *B*, *C. jejuni* Und-PP-heptasaccharide. *C*, *N. gonorrhoeae* O-linked protein glycosylation pathway. *D*, *N. gonorrhoeae* Und-PP-trisaccharide.

Interestingly, although the Pgl pathways from *N. gonorrhoeae* and *C. jejuni* are among the best-characterized bacterial glycosylation pathways, recent analysis of 380 bacterial pathogens with sequenced genomes suggests that 71% contain homologous Pgl enzymes that may be involved in protein glycosylation and most likely require undecaprenyl-phosphate as a

substrate.[11] In both Pgl pathways (Figure 2-2), undecaprenyl-phosphate is the membrane-bound precursor for the glycosyltransferase-mediated construction of the undecaprenyl-diphosphate (Und-PP) glycan, the substrate for the protein oligosaccharyltransferase (OTase). Thus, undecaprenyl-phosphate is a critical precursor for *in vitro* studies of the glycosyltransferases[5, 12, 13] and oligosaccharyltransferases (OTases)[6, 14] in *C. jejuni* and *N. gonorrhoeae*.

Because polyprenyl-phosphates and their derivatives are essential chemical tools for the biochemical characterization of several fundamental cellular processes, a convenient and effective method for their preparation is desirable. In general, chemical methods have been relied upon to obtain the desired substrates, since it is difficult to isolate large amounts of the phosphorylated polyprenols from natural sources. Undecaprenyl-phosphate (**2**) and dolichyl-phosphate (**4**) have been chemically synthesized using various approaches (Figure 2-1).[13, 15-19] However, chemical transformations are often impractical on small scales and it can be difficult to quantify the small amounts of polyprenyl-phosphates produced in these reactions. At the outset of our studies, we deemed an enzymatic route to polyprenyl-phosphates to be an excellent alternative, because it would offer the possibility of a simple, one-step reaction that could be performed on a variety of scales using ATP as the phosphoryl donor also enabling use of [γ-^{32}P] or [γ-^{32}P]ATP for sensitive quantification (Figure 2-1).

Cellular undecaprenol often exists in prokaryotic membranes as undecaprenyl-phosphate or undecaprenyl-diphosphate. Dolichol kinase has been shown to be important for cellular development in eukaryotic organisms,[20] and it is known that eukaryotic organisms store dolichol in the unmodified form.[21] However, in bacteria, it is not known whether *de novo* generation of undecaprenyl-phosphate (Figure 2-1, **2**) is an important source of this molecule (Figure 2-3), which would occur by direct phosphorylation of undecaprenol (Figure 2-1, **1**), although several

undecaprenol kinases in Gram-positive bacteria have been identified.[22, 23] In addition to undecaprenol phosphorylation, pyrophosphatase-mediated hydrolysis of undecaprenyl-diphosphate is an alternative source of undecaprenyl-phosphate (Figure 2-3).[24] The phosphatase salvage pathway is probably a major source of undecaprenyl-phosphate, since undecaprenyl-diphosphate is released after polyprenol biosynthesis and glycan transfer in most assembly pathways, including peptidoglycan biosynthesis and the Pgl pathways in *C. jejuni* and *N. gonorrhoeae*.

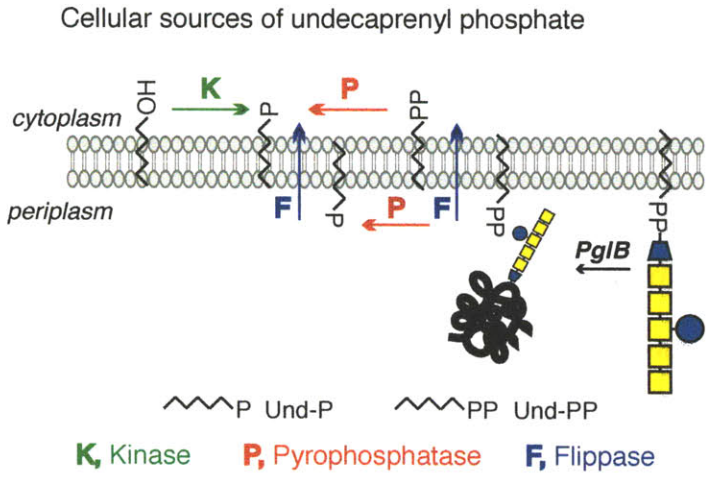


Figure 2-3. Two cellular sources of Und-P.
The first is a *de novo* pathway, in which a kinase (green) phosphorylates undecaprenol. The second is a salvage pathway, in which Und-PP is released after biosynthesis or oligosaccharide transfer (the PglB reaction from *C. jejuni* is shown). Und-PP is then hydrolyzed by a pyrophosphatase (red) and translocated across the membrane by a putative flippase (blue), although the order of these two steps is unknown.

We first focused on identifying a kinase that could be used to phosphorylate undecaprenol and then examined the utility of the enzyme in the phosphorylation of the dolichol family of polyprenols. Three enzymes were selected as potential kinases that could be used to

phosphorylate undecaprenol. In the literature, undecaprenol kinases from *Staphylococcus aureus*[22] and *Lactobacillus plantarum*[23] have been biochemically characterized, but since these identifications occurred before genome sequencing was common, the protein and related gene sequences are still unknown. A more recent study revealed undecaprenol kinase activity in a *Streptococcus mutans* homolog of diacylglycerol (DGK) kinase.[25] Using *S. mutans* DGK as a lead, a bioinformatics-based search revealed a *C. jejuni* enzyme with 36% homology to the *S. mutans* DGK. Thus, the *S. mutans* and *C. jejuni* DGKs were selected as two potential candidates for expression and activity screening.

In addition, a further search of the genomic database of *C. jejuni* was carried out to identify genes with homology to *E. coli bacA*, since strains resistant to bacitracin were found to have an overexpressed *bacA* gene.[26] Bacitracin inhibits bacterial growth by sequestering undecaprenyl-diphosphate[27] and it was believed that *bacA* might encode an undecaprenol kinase that compensated for a reduction in the pool of undecaprenyl-diphosphate by increasing the *de novo* production of undecaprenyl-phosphate from undecaprenol. However, subsequent *in vitro* functional assays in *E. coli* revealed that *E. coli* BacA is responsible for the hydrolysis of undecaprenyl-diphosphate regenerating undecaprenyl-phosphate (Figure 2-3).[24] Despite this finding in the *E. coli* system, the *C. jejuni* BacA homolog was selected as a third potential target as the enzyme had not been annotated and minimally represented a protein that bound an undecaprenyl-phosphate derivative.

All three enzymes were predicted to have multiple bilayer-spanning domains as determined the transmembrane prediction program, TMHMM.[28] Purification of membrane bound proteins is difficult and it is convenient to work with such proteins in native membranes as cell envelope fractions to maintain function of the enzymes. However, it is difficult to quantify

the amount of a specific enzyme within a crude cell envelope preparation. Quantitative Western blot analysis is one possible approach, however, it has many drawbacks. Most importantly, it is difficult to generate consistent standards for comparison with the protein of interest, because the integral membrane proteins found in cell envelope fractions are often intractable to purification and can behave anomalously with respect to mobility and staining in gel electrophoresis experiments.

To avoid these problems, an alternative method of enzyme quantification was used, in which the three target enzymes were co-expressed with lanthanide-binding-tags (LBT). LBTs are useful tools that have been applied in techniques such as X-ray crystallography and NMR.[29, 30] Here, the expression of the LBT-protein fusion allows for reliable and reproducible quantification of the amount of LBT-labeled protein present in a crude extract. The unique and tight binding of the LBT to a lanthanide metal ion, such as Tb^{3+}, results in an increased luminescence at 544 nm. The luminescence can be measured and used to estimate the amount of LBT-labeled protein present in the crude cell envelope fraction. The double-LBT (dLBT) developed recently[30] binds two Tb^{3+} ions and was utilized here to determine the relative concentrations of the kinases in the cell envelope fractions.

Results and Discussion

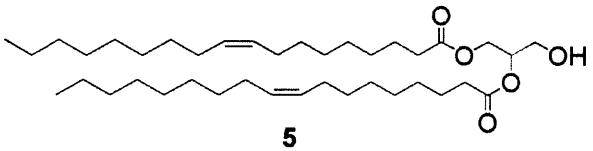
Substrate specificity of *C. jejuni* BacA, *S. mutans* DGK, and *C. jejuni* DGK

The open reading frames of *S. mutans* DGK, the *C. jejuni* DGK homolog, and the *C. jejuni* BacA homolog were amplified by PCR from the genomic DNA of *S. mutans* and *C. jejuni,* respectively. The potential kinases were expressed heterologously in *E. coli* BL21 cells and

were isolated as cell envelope fractions.

A radioactivity-based assay was performed to evaluate the three potential kinase candidates; the transfer of [^{32}P]-γ-phosphate from the aqueous soluble [γ-^{32}P]-ATP to the organic soluble undecaprenol was monitored. Initial assays of BacA suggested that it had no measurable phosphorylation activity. BacA cell envelope fractions had turnover levels (30% ± 4.4%, mean ± s.e., n=4) comparable to background controls in which cell envelope fractions containing an empty pET-24a(+) vector were assayed with undecaprenol (28% ± 1.4%, mean ± s.e., n=4) and in which BacA cell envelope fractions were assayed without undecaprenol (27% ± 1.7%, mean ± s.e., n=4).

Initial assays of *S. mutans* DGK and *C. jejuni* DGK showed potential kinase activity, but the background activity in the assay was high due to the presence of the cell envelope fraction. Thus, both enzymes were purified and assayed immediately with undecaprenol, dolichol (Figure 2-2, **1** and **3**), or diacylglycerol (Figure 2-4, **5**). The results of these assays (Table 2-1) suggested that *S. mutans* DGK phosphorylated both undecaprenol and dolichol (Figure 2-2, **1** and **3**) with high conversion and diacylglycerol with a lower yield (Figure 2-4, **5**). In contrast, *C. jejuni* DGK showed no activity with undecaprenol or dolichol, but appeared to phosphorylate diacylglycerol to some degree. Diacylglycerol phosphorylation yields may be improved by optimizing the reaction with the addition of essential phospholipids,[31] but this was not attempted.



5

Figure 2-4. The structure of diacylglycerol.

Table 2-1. Specificity of *S. mutans* DGK and *C. jejuni* DGK for undecaprenol, diacylglycerol, and dolichol.

Substrate	% turnover with *S. mutans* DGK[a]	% turnover with *C. jejuni* DGK[a]
None	17 ± 1.8	10 ± 2.3
Undecaprenol	56 ± 3.4	10 ± 1.7
Diacylglycerol	32 ± 4.5	24 ± 2.6
Dolichol	55 ± 2.9	12 ± 2.9

[a] Values reported reflect mean ± s.e., n=4. In addition, the reactions were performed on a 0.5 nmole scale in 100 μL.

The identities of the phosphorylated products generated by *S. mutans* DGK were confirmed by isolation of the [^{32}P]-radiolabeled polyprenyl-phosphates using normal-phase HPLC chromatography. Comparison by TLC confirmed that the products isolated from the kinase reaction were identical to chemically synthesized undecaprenyl-phosphate and dolichyl-phosphate.

Development of a one-pot enzymatic route to undecaprenyl-diphosphate disaccharide

Previously, the Und-PP-linked disaccharide in the *C. jejuni* pathway (Figure 2-5, **7**) was prepared by incubating synthetically prepared undecaprenyl-phosphate with the first two glycosyltransferases in the *C. jejuni* pathway, PglC and PglA (Figure 2-5). PglC is responsible for the transfer of 1-phospho-N,N'-diacetylbacillosamine (diNAcBac) from UDP-diNAcBac to give **6**.[12] The next enzyme in the pathway, PglA, catalyzes the transfer of N-acetylgalactosamine (GalNAc) from UDP-GalNAc to form Und-PP-diNAcBac-[^{3}H]-GalNAc (Figure 2-5, **7**),[5] which was detected by monitoring transfer of the radiolabeled [^{3}H]-GalNAc from the aqueous phase to the organic phase.

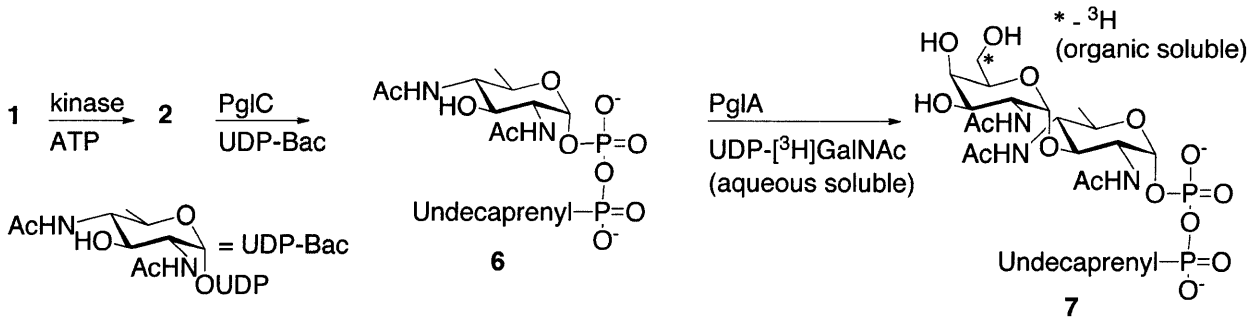


Figure 2-5. Biosynthesis of the *C. jejuni* undecaprenyl-diphosphate-linked disaccharide. The reaction is coupled to the phosphorylation of undecaprenol by *S. mutans* DGK kinase.

As an alternative to this method, a coupled reaction was developed in which Und-PP-diNAcBac-[^{3}H]-GalNAc (**7**) was generated enzymatically by the action of *S. mutans* or *C. jejuni* DGK, PglC and PglA in a one-step reaction from undecaprenol, ATP, and UDP-sugar substrates. Under initial conditions, the reaction with *S. mutans* DGK showed promising activity (~41% yield). Interestingly, the *C. jejuni* DGK reaction afforded a small amount of tritiated product (7%) at an amount higher than the background turnover (3%) present in the empty cell envelope fractions (Figure 2-6). This is in contrast to the [^{32}P]-ATP experiments in which the *C. jejuni* DGK had no detectable phosphorylation activity. BacA, however, showed levels comparable to the control reaction confirming that it has no discernable phosphorylation activity under these assay conditions.

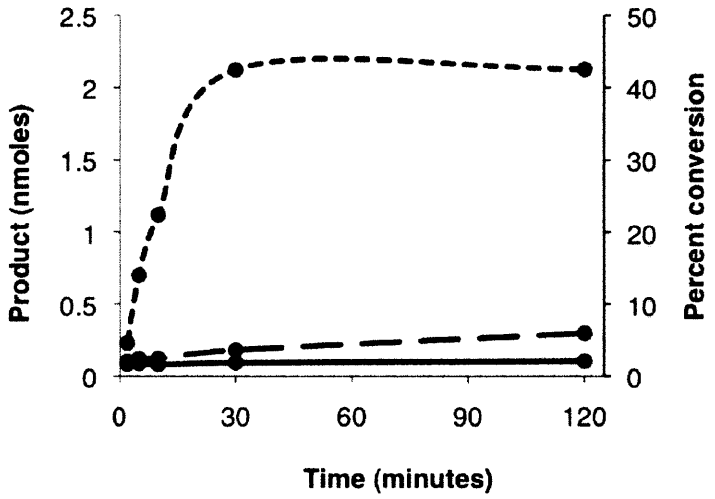
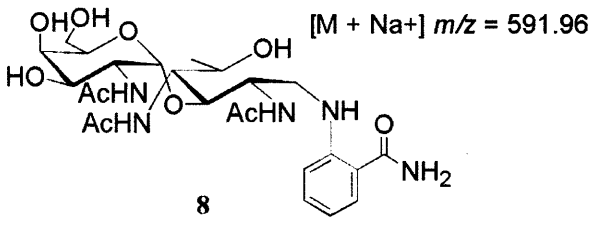


Figure 2-6. Time course assay in which the kinase candidates are coupled to PglC and PglA. Dotted line, *S. mutans* DGK; dashed line, *C. jejuni* DGK; solid line, blank cell envelope fraction.

The identity of the radiolabeled disaccharide product was confirmed as previously described using fluorescence-based HPLC and MALDI MS.[5] Specifically, the glycan was cleaved from the isoprenyl-diphosphate under acidic conditions and labeled with 2-aminobenzamide (2-AB) via reductive amination. The 2-AB-labeled disaccharide (Figure 2-7, **8**) was separated by normal phase analytical HPLC and the fractions were collected for MALDI MS, which confirmed the identity of the product.



8

Figure 2-7. Structure of the 2-AB-labeled diNAcBac-GalNAc.

Quantification of kinase enzymes by luminescence

C. jejuni DGK appeared to phosphorylate undecaprenol when coupled to PglC and PglA, although not with the robust activity shown by the *S. mutans* DGK (Figure 2-6). To establish whether differing amounts of expression affected turnover, it was necessary to quantify the amount of kinase in the cell envelope fractions. To do this, the two DGKs were expressed with a double-Lanthanide-Binding-Tag (dLBTs) as an N-terminal fusion of each protein. The dLBT-kinases were expressed and isolated as cell envelope fractions. Coupled reactions with PglC and PglA showed that the amount of turnover was not affected by the presence of the dLBT on the kinase. The turnover seen after two hours was reproducibly similar for the tagged and untagged versions of the enzymes (Table 2-2). The dLBT luminescence revealed that the *C. jejuni* DGK (10 μM) was present in a five-fold excess over the *S. mutans* DGK (2 μM) suggesting that enzyme concentration was not a factor in the low activity of *C. jejuni* DGK with undecaprenol.

Table 2-2. Results from studies of dLBT-kinase activity and luminescence.

	% turnover in disaccharide reaction	Protein present in reaction (μg)	Cell envelope fraction concentration (μM)
C. jejuni DGK	6%	-	-
S. mutans DGK	61%	-	-
dLBT-*C. jejuni* DGK	17%	3.3	10
dLBT-*S. mutans* DGK	59%	0.7	2

While *S. mutans* DGK phosphorylated both dolichol and undecaprenol in high yields, the homologous DGK in *C. jejuni* showed only small amounts of turnover in the presence of other enzymes. It is likely that the consumption of undecaprenyl-phosphate by the subsequent glycosyltransferases acted as a driving force to increase the reaction of *C. jejuni* DGK with

undecaprenol, even though undecaprenol was not the preferred substrate. This could explain the observation that *C. jejuni* DGK showed enhanced activity towards undecaprenol in the presence of coupling enzymes.

The results of the dLBT quantification of *S. mutans* kinase and *C. jejuni* DGK also provided insight into the enzymatic functions of the two enzymes. The higher concentration of *C. jejuni* DGK in the cell envelope fraction (10 μM versus 2 μM for *S. mutans* DGK) suggests that the higher levels of undecaprenol phosphorylation by *S. mutans* DGK is intrinsic to the enzyme under these assay conditions and is not caused by higher expression levels of the enzyme. The native substrate of *C. jejuni* DGK is more likely to be diacylglycerol, and not undecaprenol, as suggested by the substrate specificity assay (Table 2-1) and protein homology with the *E. coli* DGK. In addition, these results are consistent with recent work that characterizes a DGK homolog in *Bacillus subtilis* that also exhibits undecaprenol kinase activity.[32] It was noted that all bacterial species identified with undecaprenol kinases, like *S. mutans*,[25] *B. subtilis*,[32] *S. aureus*,[22] and *L. plantarum*,[23] are Gram-positive bacteria, whereas the DGK in the Gram negative bacteria *E. coli*[33] and *C. jejuni* do not significantly phosphorylate undecaprenol *in vitro*. Sequence analysis by ClustalW[34] also reveals significant differences in the N-terminus of DGKs between Gram-negative and Gram-positive bacteria (Figure 2-8). Also, comparing the TMHMM predictions[28] for the different DGKs reveals that Gram-negative DGKs are predicted to contain three distinct transmembrane helices, whereas Gram-positive DGKs are predicted to include only two transmembrane helices (Figure 2-9). Recent analysis of *E. coli* membranes determined that only undecaprenyl-phosphate and undecaprenyl-diphosphate were present, suggesting that this bacterium generates undecaprenyl-phosphate solely through the hydrolysis of undecaprenyl-diphosphate and would not require an undecaprenol kinase.[35] As *C. jejuni* is also a

Gram-negative bacterium, the lack of a DGK-like undecaprenol kinase is consistent with the hypothesis that Gram-negative and positive bacteria use different mechanisms of undecaprenol recycling.

```
CLUSTAL 2.1 multiple sequence alignment


C. jejuni    ------------MKPKYH----FLNNARYALEGLFALFKNEMAFRIELCIIIPAIVFSF 43
E. coli      ------------MANNTTGFTRIIKAAGYSWKGLRAAWINEAAFRQEGVAVLLAVVIAC 47
S. mutans    ---MPMDLRDNKQSQKKWKNR-TLTSSLEFALTGIFTAFKEERNMKKHAVSALLAVIAGL 56
B. subtilis MFTKQKEFWTLMDSKDHRNELNRFFKSFVHAGRGIWETARTERNFQFHAAAACAVLICGF 60
                              :      : .   .:  *:      *  :: .      .:: .

C. jejuni    FLKISFLEHLLLISVLILILIVEALNSAIEACVDLITNEWHE-KAKIAKDCASAAVFFSV 102
E. coli      WLDVDAITRVLLISSVMLVMIVEILNSAIEAVVDRIGSEYHE-LSGRAKDMGSAAVLIAI 106
S. mutans    VFKVSVIEWLFLLLSIFLVITFEIVNSAIENVVDLASDYHFSMLAKNAKDMAAGAVLVIS 116
B. subtilis LVELSIIEWMIIFLLIGGMFSLELLNTAIEHTVDLITDKHHP-LAKAAKDAAAGAVCVFA 119
             ..:. :  :::: :  :: .* :*:***  **   .  .   :  *** .:.** .

C. jejuni    LLALFVWGFILYNIYL----- 118
E. coli      IVAVITWCILLWSHFG----- 122
S. mutans    GFAALTGLIIFVPKIWFLLFH 137
B. subtilis VISSIIGLLIFLPKL------ 134
             .: :   :::
```

Figure 2-8. ClustalW sequence alignment of various DGKs.
The DGKs from Gram-negative *C. jejuni* and *E. coli* are compared to the DGKs from Gram-positive *S. mutans* and *B. subtilis*.[34]

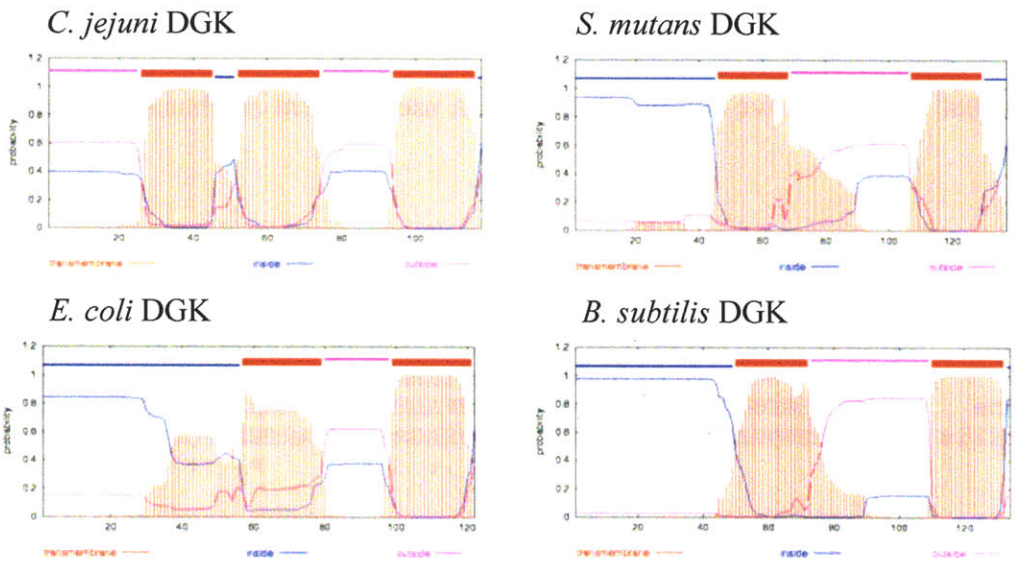


Figure 2-9. TMHMM prediction analysis of various DGKs.
The membrane topologies as predicted by TMHMM[28] are shown for two Gram-negative bacteria (*C. jejuni* and *E. coli*) and two Gram-positive bacteria (*S. mutans* and *B. subtilis*).

Optimization of Und-PP-diNAcBac-[^{3}H]-GalNAc synthesis for assays of PglB

S. mutans DGK was the optimal undecaprenol kinase in our screen and thus the reaction conditions for the kinase-dependent preparation of Und-PP-diNAcBac-[^{3}H]-GalNAc were optimized. To identify the lowest specific activity (i.e. highest substrate concentration) that would maintain high turnover (~70%) of the radiolabeled UDP-GalNAc, we performed a screen in which the amount of radioactivity in each reaction (0.15 μCi) remained constant, but the specific activity of the UDP-[^{3}H]GalNAc was varied. As seen in Table 2-3, high turnover of radiolabeled UDP-GalNAc was seen at specific activities of 300 mCi/mmol or more for *S. mutans* DGK. As a final comparison, the efficiency of *C. jejuni* DGK dropped substantially as the concentration of sugar substrate increased; the only reaction that proceeded with high turnover was at the highest specific activity (15,000 mCi/mmol). These results established that the *C. jejuni* DGK is less useful than *S. mutans* DGK as an undecaprenol kinase. Attempts to increase the turnover of the *S. mutans* DGK reaction by increasing reaction times or by adding

more kinase were unsuccessful. The reaction was scaled-up to produce 3.8 nmoles (300 mCi/mmol) of radiolabeled substrate, which is enough substrate to perform ~50 PglB reactions, in which the radiolabeled glycan is transferred to a peptide consensus sequence (Figure 2-10).

Table 2-3. Percent turnover of *C. jejuni* DGK and *S. mutans* DGK.
The DGKs were coupled to PglC and PglA to synthesize undecaprenyl-diphosphate [^{3}H]disaccharide with UDP-GalNAc at varying specific activities.

Specific activity of UDP-[6-^{3}H]GalNAc in mCi/mmol	*C. jejuni* DGK	*S. mutans* DGK
15,000	62	67
3000	20	70
300	5	67
30	3	24
3	4	23

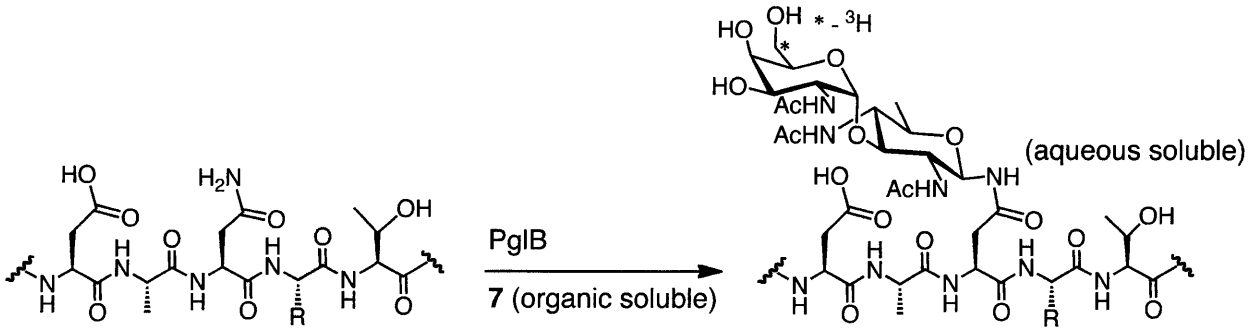


Figure 2-10. Assay of the OTase PglB.
The radiolabeled disaccharide is transferred from the isoprenyl carrier onto a peptide.

To definitively show the efficacy of the radiolabeled undecaprenyl-diphosphate-linked disaccharide (**7**) produced by the coupled enzymatic reaction, the product was tested as a substrate for PglB, the *C. jejuni* oligosaccharyl transferase. A PglB activity assay (Figure 2-10) monitors transfer of the radiolabeled disaccharide from the organic soluble polyprenyl carrier onto a peptide in the aqueous layer containing an optimized version of the required sequon in bacterial N-linked glycosylation, E/DXNXT/S.[14] The disaccharide is of sufficient length for

transfer to peptide substrates by PglB, which *in vivo* transfers a heptasaccharide from the isoprenyl carrier.[6] This type of assay has been used extensively to determine the peptide specificity of PglB[14] as well as to screen for potential inhibitors.[36] An assay of PglB using the radiolabeled undecaprenyl-diphosphate disaccharide produced by *S. mutans* DGK as a substrate is shown in Figure 2-11.

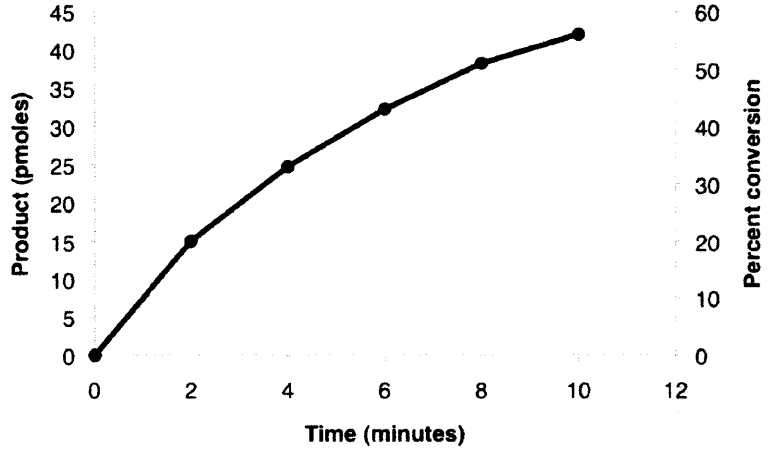


Figure 2-11. A time course of a PglB reaction.
The radiolabeled undecaprenyl-diphosphate disaccharide generated by the kinase-coupled reaction was used as a substrate.

General utilization of *S. mutans* DGK

The *S. mutans* DGK has provided a simple method for generating small amounts of radiolabeled undecaprenyl-phosphate, which is useful in a variety of applications. The multi-step chemical synthesis of the radiolabeled compound would be challenging on a small scale and complicated by the short half-live of radioactive phosphorous. Thus, the enzyme provides a facile, one-step route towards this compound, which can be easily purified by normal phase HPLC with detection by radioactivity. Although [^{32}P]-γ-ATP was used in the initial

characterization of *S. mutans* DGK to examine enzyme function, more recently, [^{33}P]-γ-ATP has proven more useful for biosynthetic purposes, since the half-life of ^{33}P (25.3 days) is twice that of ^{32}P (14.3 days). The enzyme-mediated phosphorylation of undecaprenol reaction typically proceeds with 40-50% conversion and has been prepared with a range of specific activities (40 mCi/mmol - 3000 Ci/mmol) and concentrations (7 nM - 500 μM) emphasizing the overall utility of the enzyme. Importantly, undecaprenyl-phosphate is the substrate for phospho-glycosyltransferases, the first enzymes in the both the *C. jejuni* and *N. gonorrhoeae* pathways. The work described in Chapter 4, 5 and 6 focuses on the utilization of Nanodiscs, a membrane-mimetic system,[37] to study the interaction of the early glycosyltransferases in the *C. jejuni* pathway. As such, generation of radiolabeled [^{33}P]Und-P is crucial for functional studies of the *C. jejuni* enzymes in Nanodiscs and for further biophysical studies on the effects of polyprenol in lipid bilayers.

In addition, after the initial development of the *S. mutans* DGK as a tool to generate Und-PP-diNAcBac-[^{3}H]GalNAc for PglB assays, the coupled kinase assay was extended to a variety of other glycosyltransferases. In an analogous manner, the *S. mutans* DGK was used to generate the *C. jejuni* Und-PP-trisaccharide, Und-PP-diNAcBac-[^{3}H](GalNAc)$_2$, by inclusion of PglJ, the third glycosyltransferase in the pathway (Figure 2-12, **9**).[38] The *S. mutans* DGK has been coupled to the glycosyltransferases from the *N. gonorrhoeae* O-linked glycosylation pathway to generate Und-PP-[^{3}H]diNAcBac and Und-PP-diNAcBac-[^{3}H]Gal (Figure 2-12, **10** and **11**). The Und-PP-linked glycans biosynthesized by the coupled *S. mutans* kinase reactions have been used as substrates for glycosyltransferases in the *C. jejuni* pathway[38] and *N. gonorrhoeae* pathway as described in Chapter 3. Und-PP-linked products were also shown to be substrates for the O-linked OTase from *N. gonorrhoeae* as described in Chapter 4. Furthermore, *S. mutans* DGK was

utilized for in situ generation of undecaprenyl-phosphate in a multi-enzyme screen to identify small molecule inhibitors of the *C. jejuni* UDP-diNAcBac biosynthetic enzymes.[39] One caveat to general use of the undecaprenol kinase is that it requires 0.5-1.0% Triton X-100 or similar amounts of β-D-dodeceylmaltoside for full activity, which has been observed to inhibit the activity of select enzymes. However, at this time, the kinase has been coupled to six different glycosyltransferases demonstrating the general applicability of this enzyme.

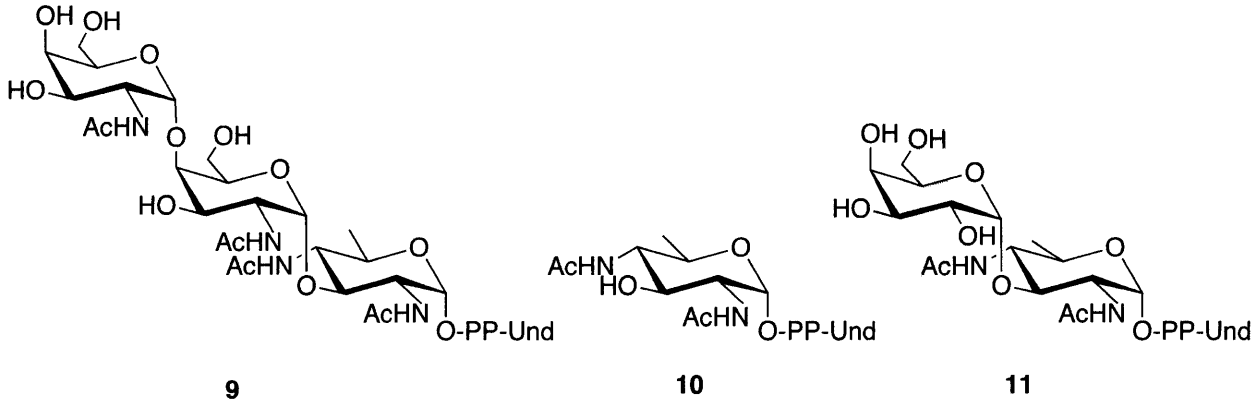


Figure 2-12. Und-PP-linked glycan substrates generated by kinase coupled reaction. Product **9** was biosynthesized by PglC, PglA, and PglJ from *C. jejuni*, **10** was produced by PglB from *N. gonorrhoeae*, and **11** was produced by PglB and PglA from *N. gonorrhoeae*.

Conclusions

In summary, the *S. mutans* DGK has been used to efficiently synthesize polyprenyl-phosphates, including undecaprenyl-phosphate and dolichyl-phosphate, from the unmodified polyprenols using ATP as the phosphoryl donor. In addition, the kinase has been used in conjunction with enzymes from the bacterial N-linked glycosylation pathway in *C. jejuni* and the O-linked glycosylation pathway in *N. gonorrhoeae* to efficiently generate radiolabeled undecaprenyl-diphosphate-linked glycans needed to probe the function of several important

glycosyltransferases and OTases. It is likely that this kinase could potentially be used to generate the polyprenyl-phosphate substrates needed for *in vitro* biochemical studies of diverse cellular pathways including the biosynthesis of peptidoglycan, O-antigens, capsular polysaccharides, and teichoic acids.

In addition, this chapter describes the use of dLBT-kinase fusion proteins to quantify the amount of protein present in partially purified cell envelope fractions. This technique exploits the unique luminescence generated by the interaction of Tb^{3+} with dLBT to specifically quantify the amount of protein present in the crude fraction and it is expected that this method could be extended for quantification of LBT-fusion proteins in other crude mixtures.

Acknowledgements

It was a pleasure to collaborate with Dr. Angelyn Larkin on this project and she performed the PglB assay and aided in the optimization of the Und-PP-diNAcBac-GalNAc biosynthesis. I am grateful to Dr. Nelson Olivier for his sage guidance on this project and his critical reading of the initial manuscript. In addition, I would like to thank Marcie Jaffee for her thorough critique of this thesis chapter. I would also like to acknowledge Dr. Matthieu Sainlos for obtaining the MALDI MS data and Dr. Jerry Troutman for his assistance in the polyprenol HPLC purification. Finally, I am grateful to Dr. Langdon Martin for sharing his dLBT-ubiquitin protein as well as his extensive LBT-related knowledge.

Experimental Procedures

Materials

The genomic DNA of *S. mutans* (25175D) and *C. jejuni* (700819D-5) were acquired from the American Type Culture Collection. Undecaprenol and the radiolabeled substrates were purchased from American Radiolabeled Chemicals, UDP-diNAcBac was prepared as previously described,[4] and the UDP-GalNAc was purchased from Sigma-Aldrich. The pure solvent upper phase (PSUP) solvent used in the enzyme assays was prepared by mixing 235 ml of H_2O containing 1.83 g of KCl with 15 ml $CHCl_3$ and 240 ml MeOH. The fluorophore-labeled disaccharide was separated on a normal-phase GlykoSepN HPLC column from ProZyme, and the isoprenes were separated on a normal-phase Varian Microsorb HPLC column. A Fluoromax 2 instrument from Jobin Yvon Horiba was used to measure the LBT luminescence. A LS6500 Beckman Scintillation Counter was used to determine the radioactivity present in the assay samples.

Cloning of kinase genes into expression vector

C. jejuni bacA and *C. jejuni dgk* were cloned from *C. jejuni* genomic DNA (ATCC 700819) using the polymerase chain reaction. Primers for both proteins inserted a BamH I restriction site at the N-terminus and an Xho I restriction site at the C-terminus (Table 2-4). *S. mutans dgk* was cloned from *S. mutans* genomic DNA (ATCC 25175D) using PCR. Similar primers inserted a BamH I site at the N-terminus. The genes were inserted into a pET-24a(+) expression vector (Novagen), which contains a T7 tag (MASMTGGQQMG) at the N-terminus and a His_6 tag at the C-terminus.

Table 2-4. Primers for the amplification of the candidate kinase genes from genomic DNA.

Primer	Sequence
C. jejuni BacA fwd	CGCGGATCCATGGAAAATTTATATGCTTTAATACTTGG
C. jejuni BacA rev	CGGCTCGAGTAGCTTAAATTCACTTCCAGCATTTAAAATC
C. jejuni DGK fwd	CGCGGATCCATGAAGCCTAAATATCATTTTTTAAATAACG
C. jejuni DGK rev	CGGCTCGAG-ATGAAAAAGCAAAAACCAAATTTTTGG
S. mutans DGK fwd	CGCGGATCCATGCCTATGGACTTAAGAGATAATAAGC
S. mutans DGK rev	CGGCTCGAGATGAAAAAGCAAAAACCAAATTTTTGG

Expression of enzymes in *E. coli* BL21 cells

Test expressions (~5 mL) were performed to identify the optimal expression strains, growth temperatures and timescales for the three enzymes, which are described here. *S. mutans* DGK and *C. jejuni* BacA were expressed in *E. coli* BL21 Gold cells and *C. jejuni* DGK was expressed in *E. coli* BL21 Codon Plus-RIL cells. A liter of Luria-Bertana Broth was inoculated with 5 ml of an overnight starter culture and incubated with shaking at 37°C. When the OD_{600} of *S. mutans* DGK and *C. jejuni* DGK reached ~0.4 AU, the cultures were cooled to 16°C. At 0.6 AU, the cultures were induced with 1 mM isopropyl β-D-thiogalactoside and incubated with shaking for ~20 hours at 16°C. The procedure was similar for *C. jejuni* BacA, with the exception that the cells were not cooled prior to induction and were incubated at 37°C overnight. After growth, the cells were harvested by centrifugation and pellets were stored at -80°C.

Isolation of kinases in cell envelope fractions

The kinases were isolated in cell envelope fractions (CEFs). The frozen cell pellets containing 2-3 g of cell pellet from 1 L of cell culture were thawed into 40 ml of buffer containing 50 mM Tris-Acetate (pH 8.0) and 1 mM EDTA per liter of cells. The cells were lysed on ice by two intervals of sonication for 75 seconds at an amplitude of 60% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes

in between rounds of sonication. A low speed spin (45 minutes, 5,697 x *g*) removed most of the cellular debris; it was followed by high speed spin (65 minutes, 142,414 x *g*) to isolate the cell envelope fraction. The cell envelope fraction was homogenized in 0.5 ml of 50 mM Tris (pH 8.0) per gram of cell pellet and aliquoted into smaller fractions for storage at -80°C.

Assay of kinases with [^{32}P]-ATP

To a tube containing 13 nmoles of dried polyprenol substrate, 3 µl dimethyl sulfoxide (DMSO) and 7 µl 7% Triton X-100 were added. The tube was vortexed and sonicated for several minutes to ensure solubilization of the substrate. To the same tube, 5 µM [^{32}P]-ATP (~4.5 mCi/mmol), 40 µl of purified enzyme or 10 µl of enzyme CEF, 30 mM Tris-Acetate (pH 8.0), 50 mM $MgCl_2$, and dH_2O were added to a total volume of 100 µl. The reaction was initiated with ATP and generally quenched after 1 hour into 1 ml of 2:1 $CHCl_3$:MeOH and extracted three times (400µl) with PSUP. The organic layer was dried under nitrogen gas and prepared for scintillation counting by addition of 200 µl of Solvable and 5 ml of Formula 989 (Perkin-Elmer) with vortexing. After waiting 30 minutes for the decay of chemiluminescence, samples were counted for 1-5 minutes.

Purification of kinases from cell envelope fraction

To purify the kinases from the cell envelope fraction, 50 µl of sample was incubated with 5% Triton X-100 for one hour at 4°C. Then the samples were spun at 16,110 *x g* in a tabletop microcentrifuge. The resultant supernatant was incubated with 100 µl Ni-NTA resin for 1 hour. The resin was placed in a 0.2 µm filter in a microcentrifuge tube for the subsequent wash and elution steps. The resin was washed twice with 250 µl of buffer containing 50 mM Tris-acetate

(pH 8.0) and twice with 250 µl of the same buffer supplemented with 45 mM imidazole; each wash was centrifuged for 45 s at 1306 *x g*. The protein was eluted in 200 µl of 50 mM Tris-acetate, pH 8.0 with 300 mM imidazole. Enzymes were assayed immediately following the elution step.

Purification of polyprenyl-phosphates by normal phase HPLC

The phosphorylated isoprenols were separated on a normal phase analytical HPLC column using 4:1 CH_3Cl: MeOH (solvent A) and 10:10:3 CH_3Cl:MeOH:H_2O, 2M ammonium acetate (solvent B). A gradient of 50% to 30% of solvent A over 20 minutes was used at a flow rate of 1 ml/min. Fractions of 0.5 ml were collected and were subjected to scintillation counting to detect the radiolabeled isoprene derivatives. TLC conditions used for undecaprenyl-phosphate comparison were 65:25:4 $CHCl_3$:MeOH:H_2O.

Synthesis of radiolabeled undecaprenyl-diphosphate disaccharide

To synthesize the desired Und-PP-diNAcBac-[^{3}H]GalNAc, undecaprenol (20 µg, 26 nmol) in hexanes was measured into a microcentrifuge tube and the solvent removed by evaporation. After adding DMSO (6 µL), followed by 1% Triton X-100 to solubilize the isoprene, 50 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.5), *S. mutans* kinase CEF (50 µL), PglA (purified to ~1 mg/ml, 20 uL), PglC cell envelope fraction (20 uL), and dH_2O were combined. In a separate tube, 10 mM ATP, 2 mM UDP-Bac, 50 µM UDP-GalNAc (150 mCi/mmol) were mixed, and the reaction was initiated by the addition of the substrate mixture to the enzyme for a final reaction volume of 200 µL. The reaction was stirred every 30 minutes for 2 hrs, then quenched through the addition of 2:1 $CHCl_3$:MeOH (1.3 mL). The aqueous layer was removed,

and the organic layer extracted with PSUP (3x, 300 μL).

The synthesis and purification of [^{33}P]Und-P, Und-PP-[^{3}H]diNAcBac and Und-PP-diNAcBac-[^{3}H]Gal are described in Chapters 3 and 7. The synthesis and purification of Und-PP-diNAcBac-[^{3}H](GalNAc)$_2$ was carried out by Dr. Jerry Troutman and is detailed in the following reference.[38]

Quantification of the dLBT-kinase in a cell envelope fraction using luminescence

The amount of dLBT-tagged protein present in a sample can be determined by measuring the luminescence of the sample incubated with Tb^{3+} under denaturing conditions. A Fluoromax 2 instrument (Jobin Yvon Horiba) was used to obtain the luminescence emission spectra as previously described.[40] The samples were prepared by diluting 60-70 μL of cell envelope fraction in a total of 3 mL 6 M guanidine hydrochloride. A 40-fold dilution minimum was maintained to ensure complete denaturation of the protein and availability of the dLBT for Tb^{3+} binding. The luminescence was measured by integrating the area under the Tb^{3+} emission peak at 544 nm. Measurements were recorded before and after the addition of 10 μl of 1 mM Tb^{3+} and to determine the background luminescence. The background luminescence was low for the purified standards (0.5-5% depending on the protein concentration), whereas it was much higher for the kinase cell envelope fractions (10-50%) due to the relatively low concentrations of the protein. For all subsequent analysis, the background luminescence of the sample was subtracted from the luminescence values of the sample incubated with Tb^{3+}.

A purified dLBT-ubiquitin construct was used as a standard to determine the amount of luminescence emitted by a known amount of dLBT. The concentration of the dLBT-ubiquitin was determined using UV spectroscopy (ε_{280} = 15590 $M^{-1}cm^{-1}$). A linear standard graph was

prepared for a range of dLBT-ubiquitin concentrations from 50 nM to 2 μM (Figure 2-13). The luminescence of the dLBT-kinases was recorded and compared to the standards to approximate the amount of kinase present.

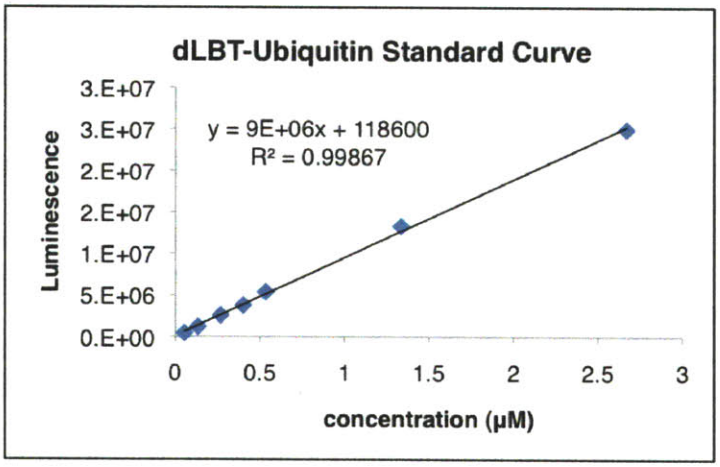


Figure 2-13. A standard curve of dLBT-ubiquitin luminescence.

References

1. Hendrickson, T. L.; Imperiali, B., Metal ion dependence of oligosaccharyl transferase: implications for catalysis. *Biochemistry* **1995,** 34, (29), 9444-50.
2. Lechner, J.; Wieland, F.; Sumper, M., Biosynthesis of sulfated saccharides N-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from *halobacteria. J. Biol. Chem.* **1985,** 260, (2), 860-6.
3. Szymanski, C. M.; Wren, B. W., Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Microbiol.* **2005,** 3, (3), 225-37.
4. Olivier, N. B.; Chen, M. M.; Behr, J. R.; Imperiali, B., *In vitro* biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* **2006,** 45, (45), 13659-69.
5. Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. U.S.A.* **2005,** 102, (40), 14255-9.
6. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* **2005,** 12, (12), 1311-5.
7. Vik, A.; Aas, F. E.; Anonsen, J. H.; Bilsborough, S.; Schneider, A.; Egge-Jacobsen, W.; Koomey, M., Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. U S A* **2009,** 106, (11), 4447-52.
8. Aas, F. E.; Vik, A.; Vedde, J.; Koomey, M.; Egge-Jacobsen, W., *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol. Microbiol.* **2007,** 65, (3), 607-24.
9. Higashi, Y.; Strominger, J. L.; Sweeley, C. C., Biosynthesis of the peptidoglycan of bacterial cell walls. XXI. Isolation of free C55-isoprenoid alcohol and of lipid intermediates in peptidoglycan synthesis from Staphylococcus aureus. *J. Biol. Chem.* **1970,** 245, (14), 3697-702.
10. Troy, F. A.; Vijay, I. K.; Tesche, N., Role of undecaprenyl phosphate in synthesis of polymers containing sialic acid in *Escherichia coli*. *J. Biol. Chem.* **1975,** 250, (1), 156-63.
11. Kumar, M.; Balaji, P. V., Comparative genomics analysis of completely sequenced microbial genomes reveals the ubiquity of N-linked glycosylation in prokaryotes. *Mol. Biosyst.* **2011**.
12. Glover, K. J.; Weerapana, E.; Chen, M. M.; Imperiali, B., Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2006,** 45, (16), 5343-50.
13. Weerapana, E.; Glover, K. J.; Chen, M. M.; Imperiali, B., Investigating bacterial N-linked glycosylation: synthesis and glycosyl acceptor activity of the undecaprenyl pyrophosphate-linked bacillosamine. *J. Am. Chem. Soc.* **2005,** 127, (40), 13766-7.
14. Chen, M. M.; Glover, K. J.; Imperiali, B., From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* **2007,** 46, (18), 5579-85.

15. Ye, X. Y.; Lo, M. C.; Brunner, L.; Walker, D.; Kahne, D.; Walker, S., Better substrates for bacterial transglycosylases. *J Am Chem Soc* **2001,** 123, (13), 3155-6.
16. Branch, C. L.; Burton, G.; Moss, S. F., *Synth. Commun.* **1999,** 29, 2639-2644.
17. Warren, C. D.; Jeanloz, R. W., Chemical synthesis of dolichyl phosphate and dolichyl glycosyl phosphates and pyrophosphates or "dolichol intermediates". *Methods Enzymol.* **1978,** 50, 122-37.
18. Suzuki, S.; Mori, F.; Takigawa, T.; Ibata, K., *Tetrahedron Lett.* **1983,** 24, 5103.
19. Imperiali, B.; Zimmerman, J. W., *Tetrahedron Lett.* **1988,** 29, 5343.
20. Rossignol, D. P.; Lennarz, W. J.; Waechter, C. J., Induction of phosphorylation of dolichol during embryonic development of the sea urchin. *J Biol Chem* **1981,** 256, (20), 10538-42.
21. Chojnacki, T.; Dallner, G., The biological role of dolichol. *Biochem J* **1988,** 251, (1), 1-9.
22. Higashi, Y.; Siewert, G.; Strominger, J. L., Biosynthesis of the peptidoglycan of bacterial cell walls. XIX. Isoprenoid alcohol phosphokinase. *J. Biol. Chem.* **1970,** 245, (14), 3683-90.
23. Kalin, J. R.; Allen, C. M., Jr., Characterization of undecaprenol kinase from *Lactobacillus plantarum*. *Biochim. Biophys. Acta.* **1979,** 574, (1), 112-22.
24. El Ghachi, M.; Bouhss, A.; Blanot, D.; Mengin-Lecreulx, D., The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. *J. Biol. Chem.* **2004,** 279, (29), 30106-13.
25. Lis, M.; Kuramitsu, H. K., The stress-responsive *dgk* gene from *Streptococcus mutans* encodes a putative undecaprenol kinase activity. *Infect. Immun.* **2003,** 71, (4), 1938-43.
26. Cain, B. D.; Norton, P. J.; Eubanks, W.; Nick, H. S.; Allen, C. M., Amplification of the bacA gene confers bacitracin resistance to Escherichia coli. *J. Bacteriol.* **1993,** 175, (12), 3784-9.
27. Siewert, G.; Strominger, J. L., Bacitracin: An Inhibitor of the Dephosphorylation of Lipid Pyrophosphate, an Intermediate in the Biosynthesis of the Peptidoglycan of Bacterial Cell Walls. *Proc. Natl. Acad. U.S.A.* **1967,** 57, (3), 767-773.
28. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
29. Silvaggi, N. R.; Martin, L. J.; Schwalbe, H.; Imperiali, B.; Allen, K. N., Double-lanthanide-binding tags for macromolecular crystallographic structure determination. *J Am Chem Soc* **2007,** 129, (22), 7114-20.
30. Martin, L. J.; Hahnke, M. J.; Nitz, M.; Wohnert, J.; Silvaggi, N. R.; Allen, K. N.; Schwalbe, H.; Imperiali, B., Double-lanthanide-binding tags: design, photophysical properties, and NMR applications. *J. Am. Chem. Soc.* **2007,** 129, (22), 7106-13.
31. Walsh, J. P.; Bell, R. M., sn-1,2-Diacylglycerol kinase of Escherichia coli. Mixed micellar analysis of the phospholipid cofactor requirement and divalent cation dependence. *J Biol Chem* **1986,** 261, (14), 6239-47.
32. Jerga, A.; Lu, Y. J.; Schujman, G. E.; de Mendoza, D.; Rock, C. O., Identification of a soluble diacylglycerol kinase required for lipoteichoic acid production in Bacillus subtilis. *J. Biol. Chem.* **2007,** 282, (30), 21738-45.
33. Bohnenberger, E.; Sandermann, H., Jr., Diglyceride kinase from Escherichia coli. Purification in organic solvent and some properties of the enzyme. *Eur. J. Biochem.* **1979,** 94, (2), 401-7.

34. Thompson, J. D.; Higgins, D. G.; Gibson, T. J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **1994,** 22, (22), 4673-80.
35. Barreteau, H.; Magnet, S.; El Ghachi, M.; Touze, T.; Arthur, M.; Mengin-Lecreulx, D.; Blanot, D., Quantitative high-performance liquid chromatography analysis of the pool levels of undecaprenyl phosphate and its derivatives in bacterial membranes. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2009,** 877, (3), 213-20.
36. Larkin, A. Investigation of Asparagine-Linked Glycosylation in Archaeal and Bacterial Systems. Massachusetts Institute of Technology, Cambridge, MA, 2010.
37. Nath, A.; Atkins, W. M.; Sligar, S. G., Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **2007,** 46, (8), 2059-69.
38. Troutman, J. M.; Imperiali, B., *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **2009,** 48, (12), 2807-16.
39. Morrison, J. P.; Troutman, J. M.; Imperiali, B., Development of a multicomponent kinetic assay of the early enzymes in the Campylobacter jejuni N-linked glycosylation pathway. *Bioorg Med Chem* **2010,** 18, (23), 8167-71.
40. Franz, K. J.; Nitz, M.; Imperiali, B., Lanthanide-binding tags as versatile protein coexpression probes. *Chembiochem* **2003,** 4, (4), 265-71.

Chapter 3: Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of undecaprenyl-diphosphate-linked glycan intermediates

Most of the work in this chapter has been accepted for publication in the following references:

Hartley, M.D.; Morrison, M.J.; Aas, F.E.; Børud, B.; Koomey, M.; and Imperiali, B., Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylbacillosamine. *Biochem.* **2011**, in press.

Børud, B.; Viburiene, R.; Hartley, M.D.; Paulsen, B.S.; Egge-Jacobsen, W.; Imperiali, B.; Koomey. Genetic and molecular analyses reveal an evolutionary trajectory for glycan synthesis in a bacterial protein glycosylation system. *Proc Natl Acad Sci U S A*, **2011**, in press.

Michael Morrison purified and biochemically characterized PglD and PglC and performed the kinetic characterization of PglC and PglB-ATD. Generation and characterization of UDP-diNAcBac was performed in collaboration. Finn Erik Aas and Bente Borud provided the DNA constructs encoding all of the Pgl enzymes.

Introduction

In *Neisseria gonorrhoeae*, individual pilin subunits rapidly assemble and disassemble to form the flagellar-like Type IV pili, which mediate essential interactions with host cells and affect many aspects of pathogenicity including surface motility, bacteria-host communication, cell signaling, bacterial dissemination and biofilm formation.[1, 2] Recently, the gonococcal pilin glycosylation system was shown to be a general O-linked system in which many structurally distinct periplasmic proteins undergo glycosylation.[3] Glycan-modifications on pili, flagella and other extracellular proteins have been implicated in bacterial pathogenicity, which has led to increased interest in bacterial glycosylation pathways as potential antibacterial targets.[4-9] The focus of this study is the protein glycosylation (*pgl*) locus identified in *N. gonorrhoeae*, which is responsible for glycan addition to distinct serine residues.[10, 11]

The protein glycan modifications present in null strains of specific *N. gonorrhoeae pgl* genes have been analyzed by top-down mass spectrometry (MS)[3, 10] and the following model of the protein glycosylation pathway has been developed (Figure 3-1). The core *pgl* locus contains four genes, three of which (*pglD*, *pglC*, and *pglB*) are required for the synthesis of an undecaprenyl-diphosphate 2,4-diacetamido-2,4,6-trideoxy-α-D-hexose (Und-PP-DATDH).[12] The term 'DATDH' is used to indicate that the stereochemistry of this sugar has not been previously determined. PglD and PglC perform NAD^+-dependent dehydratase and pyridoxal-dependent aminotransferase reactions, respectively, to convert a UDP-HexNAc to UDP-2-acetamido-4-amino-2,4,6-trideoxy-α-D-hexose (UDP-4-amino), but it is not known if the initial substrate is UDP-GlcNAc or UDP-GalNAc. PglB is a bifunctional enzyme, which catalyzes the amino acetylation of UDP-4-amino to form UDP-DATDH and the transfer of phospho-DATDH to

undecaprenyl-phosphate (Und-P). The fourth gene, *pglF*, shares homology with ATP-dependent ABC transporter-type flippases and is putatively involved in the translocation of the undecaprenyl-diphosphate-linked glycan across the plasma membrane. Although the function of this gene has not been demonstrated, the *pglF*-null strain does exhibit diminished glycosylation.[10, 11]

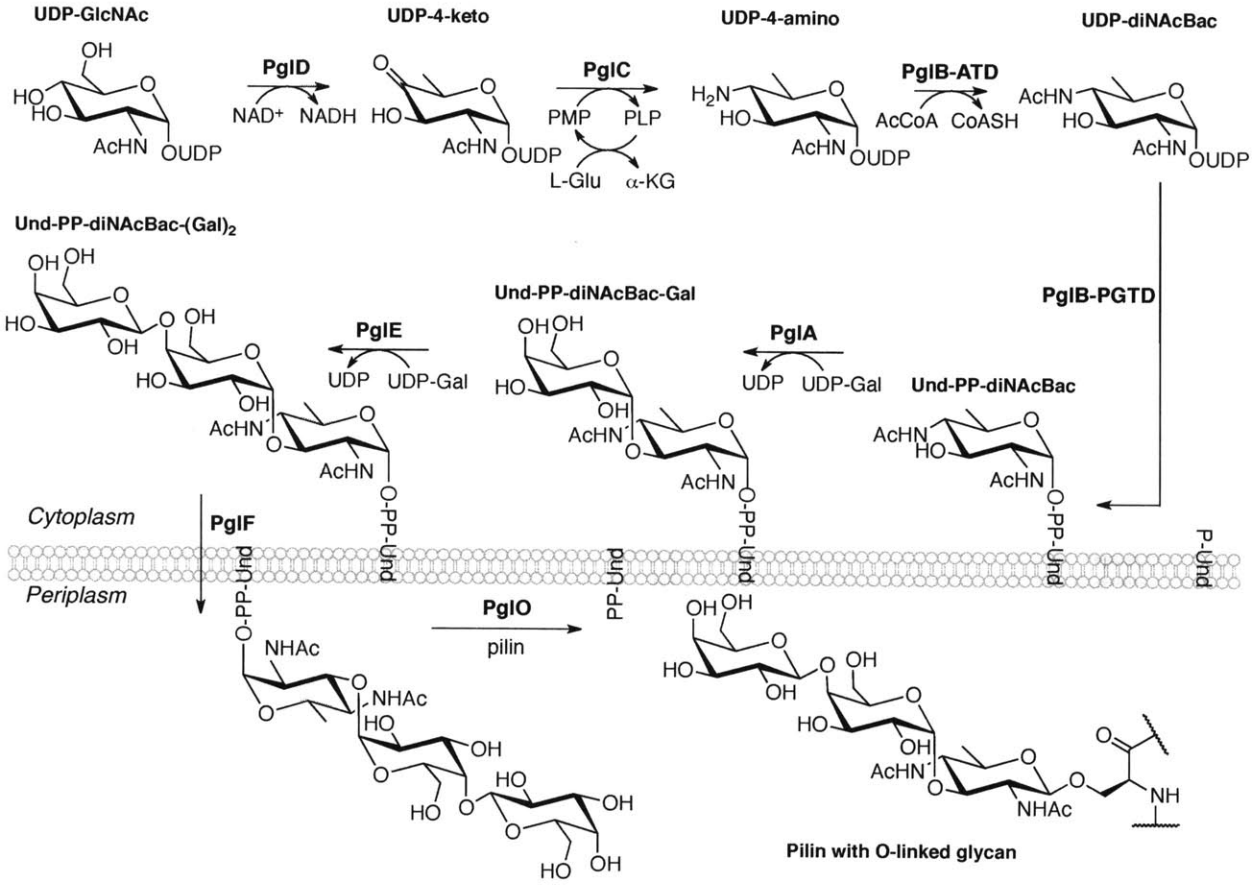


Figure 3-1. O-linked protein glycosylation pathway in *Neisseria gonorrhoeae*.
The correct stereochemistry of the diNAcBac sugar is shown as determined by this study.

Interestingly, the remaining genes involved in pilin glycosylation are not found in the core *pgl* locus. The products of the *pglA* and *pglE* genes further elaborate the polyprenyl-diphosphate-linked DATDH with the transfer of two sequential galactose units (Figure 3-1).[10] The *pglA* and *pglE* genes undergo phase variation in which the genes are alternately turned on

and off. Phase-variant *pglA* alleles have been proposed to be associated with more virulent strains of *N. gonorrhoeae*, although these studies have been disputed.[6, 7, 10] In addition to PglA and PglE, an alternate glycosyltransferase PglH adds a Glc unit instead of Gal to Und-PP-DATDH (Figure 3-2).[13] Finally, a gene has been identified, *pglO*, which shares homology with the O-antigen ligase (WaaL) family and was required for formation of the protein-glycan linkage (Figure 3-1).[3]

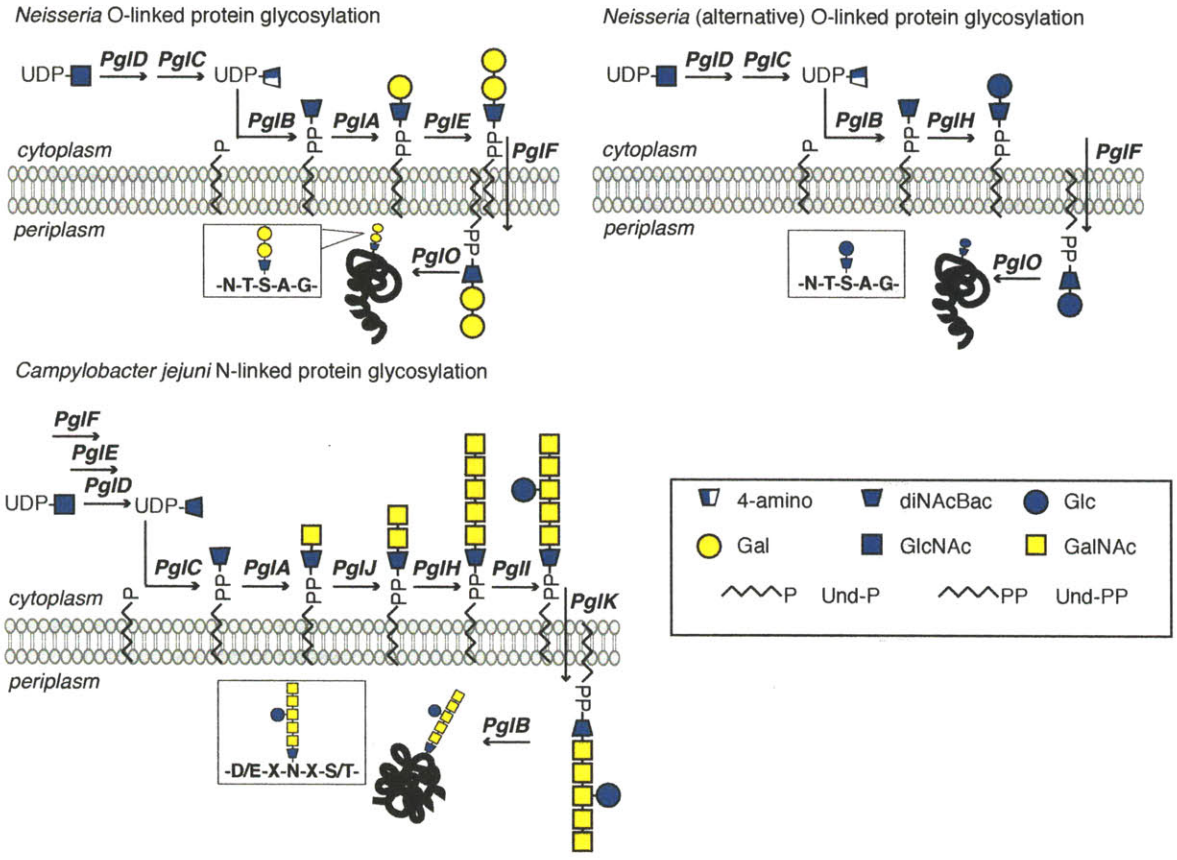


Figure 3-2. Schematic representations of bacterial protein glycosylation pathways: *Upper left*, O-linked pathway in *Neisseria gonorrhoeae* and *Neisseria meningitidis*. *Upper right*, alternative O-linked pathway in *Neisseria* species. *Lower left*, N-linked pathway in *Campylobacter jejuni*.

Interestingly, the Pgl pathway is not unique among bacteria; approximately 71% of sequenced bacterial pathogens contain five or more genes that are homologous to those found in

this pathway.[14] A closely related pathway has been identified in *Neisseria meningitidis* and these two pathways represent the first examples of O-linked protein glycosylation that exploit polyprenyl-linked intermediates; all other identified O-linked protein glycosylation pathways involve sequential transfer of individual saccharide units from nucleotide or polyprenyl-phosphate activated glycan donors to protein substrates. However, many more bacterial pathogens may contain similar types of O-linked protein glycosylation pathways.[14]

Another important parallel pathway is the N-linked protein glycosylation (also designated Pgl) pathway in *Campylobacter jejuni.*[15] The first three enzymes (PglD, C, and B) in the *N. gonorrhoeae* O-linked glycosylation pathway share functional homology with the first four enzymes in the *C. jejuni* pathway, with the exception that the *C. jejuni* locus encodes separate enzymes for the sequential acetyltransferase and phospho-glycosyltransferase reactions (Table 3-1). Both the *N. gonorrhoeae* and *C. jejuni* pathways produce an initial Und-PP-DATDH intermediate, but this intermediate is elaborated in distinct ways (Figure 3-2). The *N. gonorrhoeae* pathway produces a serine-linked mono-, di- or trisaccharide[12] and the *C. jejuni* pathway generates an asparagine-linked heptasaccharide.[16] The *C. jejuni* glycosylation pathway serves as an important model for the *N. gonorrhoeae* system and previous work has resulted in the complete biochemical characterization of the *C. jejuni* Pgl pathway enzymes except for the flippase (PglK).[17-20]

Table 3-1. Percent sequence identity (%) between Pgl proteins.
Comparisons are shown between *N. gonorrhoeae* (*Ng*) and *C. jejuni* (*Cj*) and between *N. gonorrhoeae* (*Ng*) and *N. meningitidis* (*Nm*) proteins.

Function	*N. gonorrhoeae*	*C. jejuni*	% *Ng* and *Cj*	*N. meningitidis*	% *Ng* and *Nm*
dehydratase	PglD	PglF	29.8	PglD	92.5
aminotransferase	PglC	PglE	21.2	PglC	92.8
acetyltransferase	PglB-ATD	PglD	29.7	PglB-ATD	84.9
P-glycosyltransferase	PglB-GTD	PglC	52.3	PglB-GTD	90.3
glycosyltransferase	PglA	PglA	21.3	PglA	95.5
glycosyltransferase	PglE	PglJ	12.0	PglE	93.0

In this chapter, the biochemical functions of the proteins PglD, C, B, A, E, and H from *N. gonorrhoeae* are characterized for the first time through *in vitro* biochemical analysis. Importantly, the previously undefined stereochemical assignment of the UDP-DATDH produced by PglD, C, and B is unequivocally shown to be UDP-diNAcBAc, which is also the identity of the first sugar added in the *C. jejuni* N-linked glycosylation pathway. Additionally, *in vitro* assays demonstrate that the phospho-glycosyltransferase (PglB) and three glycosyltransferases (PglA, PglE, and PglH) build the glycan on an undecaprenyl-diphosphate carrier prior to *en bloc* transfer to protein and that these enzymes display strict specificity for the UDP-saccharide donor.

Results and Discussion

Determination of UDP-DATDH stereochemistry by NMR

The biosynthesis of UDP-DATDH from UDP-GlcNAc was carried out in the presence of purified dehydratase (PglD), aminotransferase (PglC), and acetyltransferase-phospho-glycosyltransferase (PglB). PglC, a soluble protein, was purified to homogeneity (Figure 3-3, lane 2). TMHMM, a transmembrane prediction program,[21] predicts that PglD and PglB contain four and one transmembrane helices, respectively. SDS-PAGE analysis of purified PglD demonstrated that the desired protein product is the dominant component (Figure 3-3, lane 1). PglB was purified for this experiment, but the enzyme was used as a partially purified cell envelope fraction (CEF) in all other assays to avoid problems with protein stability (Figure 3-3, lane 4). The anti-His_4 Western blot analysis revealed that purified PglD and PglB CEFs contained His_6-tagged truncation products that formed during protein expression (Figure 3-3, lanes 1 and 4).

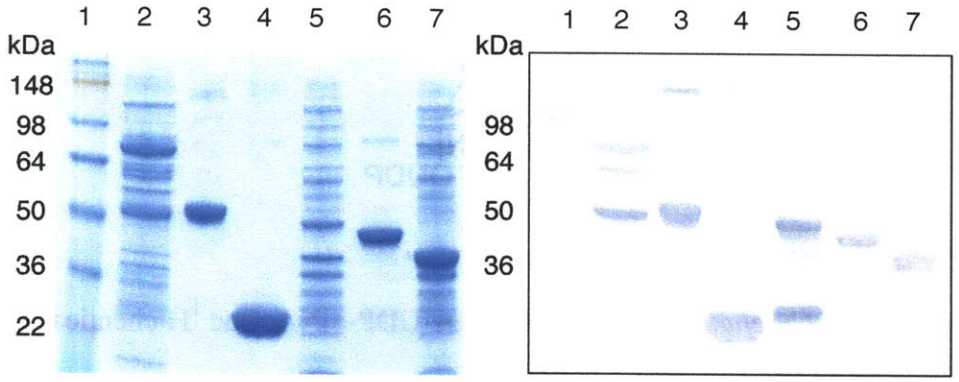


Figure 3-3. SDS-PAGE and Western blot analysis of purified Pgl proteins.
Left, SDS-PAGE stained with Coomasie Blue. *Right*, Western blot analysis using a His_4-specific antibody. Lane 1, molecular weight standards; lane 2, PglD (71 kDa); lane 3, PglC (43 kDa); lane 4, PglB-ATD (21 kDa); lane 5, PglB CEF (44 kDa); lane 6, PglA (43 kDa); lane 7, PglE CEF (38 kDa).

The biosynthesis of UDP-diNAcBac by the action of PglD, PglC, and PglB was followed by capillary electrophoresis to ensure complete turnover of the substrates. Product purification by reverse phase-HPLC (C18) was used to remove unreacted substrates and cofactors leading to a final UDP-DATDH purity of >95%. To determine the final stereochemistry of the sugar, ^{1}H NMR (see Appendix) was employed and the chemical shifts and coupling constants were compared with UDP-diNAcBac from the *C. jejuni* pathway (Table 3-2). The values for the UDP-DATDH sugar from *N. gonorrhoeae* exactly match the values of UDP-diNAcBac from *C. jejuni*.[20] Further confirmation was provided by the ^{31}P, ^{13}C, and ^{1}H-^{1}H COSY NMR spectra (see Appendix). Therefore, the stereochemistry of the DATDH sugar in the *N. gonorrhoeae* pathway is confirmed as diNAcBac (Figure 3-4).

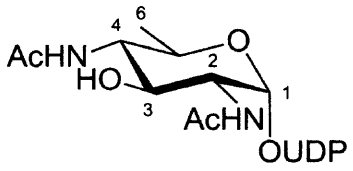


Figure 3-4. Structure of UDP-diNAcBac.

Table 3-2. Comparison of *C. jejuni* and *N. gonorrhoeae* UDP-diNAcBac ^{1}H chemical shifts and coupling constant assignments.

	C. jejuni[20]		*N. gonorrhoeae*	
Moiety	d_H (ppm)	J	d_H (ppm)	J
H1	5.48 (dd)	$J_{1,2}$ = 3.2 Hz $J_{1,P}$ = 6.9 Hz	5.46 (dd)	$J_{1,2}$ = 3.2 Hz $J_{1,P}$ = 6.9 Hz
H2	4.02 (m)		4.03 (m)	
H3	3.79 (at)	$J_{2,3}$ = 10.2 Hz $J_{3,4}$ = 10.2 Hz	3.76 (at)	$J_{2,3}$ = 10.1 Hz $J_{3,4}$ = 10.2 Hz
H4	3.69 (at)	$J_{4,5}$ = 10.2 Hz	3.67 (at)	$J_{4,5}$ = 10.1 Hz
H5	4.05 (m)		4.03 (m)	
H6	1.19 (d)	$J_{5,6}$ = 6.2 Hz	1.17 (d)	$J_{5,6}$ = 6.2 Hz

The extreme diversity of bacterial glycans is highly significant since glycoconjugates typically decorate the bacterial cell surface facilitating interactions with host cells[4, 8] and potentially confounding the immune response.[6, 22-25] Bacillosamine was originally identified as an unusual 2,4-diamino-2,4,6-trideoxy-α-D-glucose in *Bacillus subtilis*,[26] but it also appears in the mono- and di-aminoacetylated forms in a large variety of bacterial glycoconjugates.[27] It has been found frequently in the O-antigen and capsular polysaccharide of Gram negative bacteria, but has also been identified in the S-layer of Gram positive bacteria and as the UDP-activated donor in cellular extracts.[28] DiNAcBac was initially discovered in N-linked glycans in *C. jejuni*,[16] but more recently, a second route to diNAcBac was biochemically characterized in *C. jejuni*, in which GDP-diNAcBac is an intermediate in the CMP-legionaminic acid biosynthetic pathway.[29]

The pilin oligosaccharide in *N. gonorrhoeae* was thought originally to comprise Gal-α-(1,3)-GlcNAc-β-Ser 63.[30] Mass spectrometry and bioinformatic analysis suggested that the linking sugar unit was DATDH instead of GlcNAc.[10] Herein, we confirm the stereochemical assignment of this sugar for the first time showing that the DATDH sugar produced in *N. gonorrhoeae* is diNAcBac (Table 3-2). This adds to the growing number of oligosaccharides identified in bacteria that contain forms of bacillosamine.

PglD, PglC and PglB produce Und-PP-diNAcBac in *N. gonorrhoeae*; these three enzymes have functional homology to PglF(*Cj*), PglE(*Cj*), PglD(*Cj*), and PglC(*Cj*) in *C. jejuni*, which produce the same polyprenyl-linked intermediate (Figure 3-2). Even though the early enzymes in these two pathways carry out identical functions, the sequence identity is relatively low (25-30%), except for the phospho-glycosyltransferase domain of PglB, which has 52% identity with PglC(*Cj*) (Table 3-1). These numbers starkly contrast the sequence identity observed between *N. gonorrhoeae* and *N. meningitidis*, which indicate much closer homologies (>84%, Table 3-1). These numbers imply that *C. jejuni* and *N. gonorrhoeae* pathways are only distantly related from an evolutionary standpoint.

Characterization of PglB-ATD and PglC

The similarity of the *N. gonorrhoeae* protein glycosylation pathway to the pathway in *C. jejuni* suggests that the acetyltransferase domain of PglB acts first on UDP-4-amino to generate UDP-diNAcBac, which is then utilized as a substrate by the phospho-glycosyltransferase domain of PglB (PglB-PGTD, Figure 3-1). The C-terminal acetyltransferase domain of full-length PglB (PglB-ATD, based upon a ClustalW alignment with PglD(*Cj*)) was expressed and purified (Figure 3-3, lane 3). This provided a suitable amount of well-behaved, soluble protein in the

absence of the N-terminal phospho-glycosyltransferase domain, which is predicted by TMHMM to contain a single transmembrane domain.[21] Functional analysis of PglB-ATD described below confirmed that this domain acetylates UDP-4-amino to produce UDP-diNAcBac, which is the substrate for PglB-PGTD mediated transfer of P-diNAcBac to Und-P (Figue 3-1).

Both the aminotransferase (PglC) and acetyltransferase (PglB-ATD) reactions exhibited typical Michaelis-Menten kinetics over a wide range of substrate concentrations. Initial velocity data were used to calculate kinetic parameters of L-glutamate and UDP-4-keto for PglC and AcCoA and UDP-4-amino for PglB-ATD (Table 3-3).

Table 3-3. Steady-state parameters for PglC and PglB-ATD.

Enzyme	Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
PglC	L-glutamate	0.025 ± 0.001	4900 ± 900	5.1
PglC	UDP-4-keto	0.039 ± 0.002	233 ± 35	167
PglB-ATD	AcCoA	0.928 ± 0.032	456 ± 34	2035
PglB-ATD	UDP-4-amino	0.416 ± 0.016	122 ± 17	3410

Work is ongoing to understand the lower level of sequence homology between the *C. jejuni* and *N. gonorrhoeae* diNAcBac biosynthetic enzymes. As a first step, this study describes the kinetic parameters of both the aminotransferase (PglC) and the acetyltransferase (PglB-ATD) (Table 3-3). The apparent K_m of the UDP-4-keto sugar for PglC (233 μM) was comparable to the PglE(*Cj*) homologue (48 μM and 610 μM)[31, 32] and well within typical K_m values for this type of substrate. Likewise, K_m values of the UDP-4-amino substrate for the acetyltransferases PglB-ATD (122 μM) and PglD(*Cj*) (410 μM)[20] lead to a similar conclusion. However, the *N. gonorrhoeae* enzymes presented here are catalytically much less efficient (k_{cat} is 10-100 fold less for PglC and 1000-fold less for PglB-ATD) than their *C. jejuni* counterparts with respect to the UDP-sugar. This observation is reflected in the differences between their specificity constants

(k_{cat}/K_m) (Table 3-3). The high acetyltransferase activity in the *C. jejuni* pathway is used to drive the biosynthesis of the UDP-diNAcBac sugar.[20] A similar phenomenon is observed in the *N. gonorrhoeae* pathway, with a 20-fold enhancement in k_{cat}/K_m of PglB-ATD with respect to aminotransferase activity.

For PglB-ATD, one cannot rule out interplay between the acetyltransferase domain and the missing C-terminal phospho-glycosyltransferase domain. Therefore, care must be taken in interpreting the reduced *N. gonorrhoeae* acetyltransferase efficiency as compared to PglD(*Cj*). Further work will be necessary to clarify how domain interactions affect kinetic parameters. In addition, the low sequence homology between the *C. jejuni* and *N. gonorrhoeae* UDP-diNAcBac pathway enzymes (Table 3-2) could contribute to the observed differences in catalytic efficiency.

Biochemical characterization of the *N. gonorrhoeae* glycosyltransferases

As mentioned above, TMHMM[21] predicts that PglB has a single N-terminal transmembrane helix. In addition, PglE is predicted to contain two C-terminal transmembrane helices. Purification of these proteins by detergent solubilization and extraction resulted in low yields and loss of activity; to avoid these problems, both PglB and PglE were purified as crude CEFs for the glycosyltransferase assays. SDS-PAGE and Western blot analysis showed that PglB and PglE are the predominant bands present in the respective CEFs (Figure 3-3, lanes 4 and 6). In all assays involving PglB and PglE, negative controls with CEFs lacking overexpressed PglB or PglE showed no glycosyltransferase activity (data not shown). PglA is predicted to be soluble and was purified to homogeneity (Figure 3-3, lane 5).

The tritium-labeled product of PglB was prepared from undecaprenyl-phosphate, UDP-4-amino, and [^{3}H]AcCoA, and the tritium-labeled products of PglA and PglE were biosynthesized

from UDP-[^{3}H]Gal and the Und-PP-linked mono- or disaccharide, respectively. The PglB, PglA, and PglE products were analyzed by NP-HPLC. Und-PP-[^{3}H]diNAcBac, Und-PP-diNAcBac-[^{3}H]Gal and Und-PP-diNAcBac-Gal-[^{3}H]Gal were retained on the column with retention times consistent with glycan size (Figure 3-5). Each product was analyzed separately in order to confirm the identity of the peaks (Figure 3-5). In addition, the glycosyltransferase products were characterized by 2-aminobenzamide (2-AB) fluorescence-labeling protocol as previously described.[33] The 2-AB labeled disaccharide and trisaccharide were purified and MALDI MS was used to verify the masses of the products (Figure 3-6). These studies definitively annotate the biochemical functions of PglB, PglA, and PglE as the phospho-glycosyltransferase and the two glycosyltransferases that produce Und-PP-linked mono-, di- and trisaccharides, respectively.

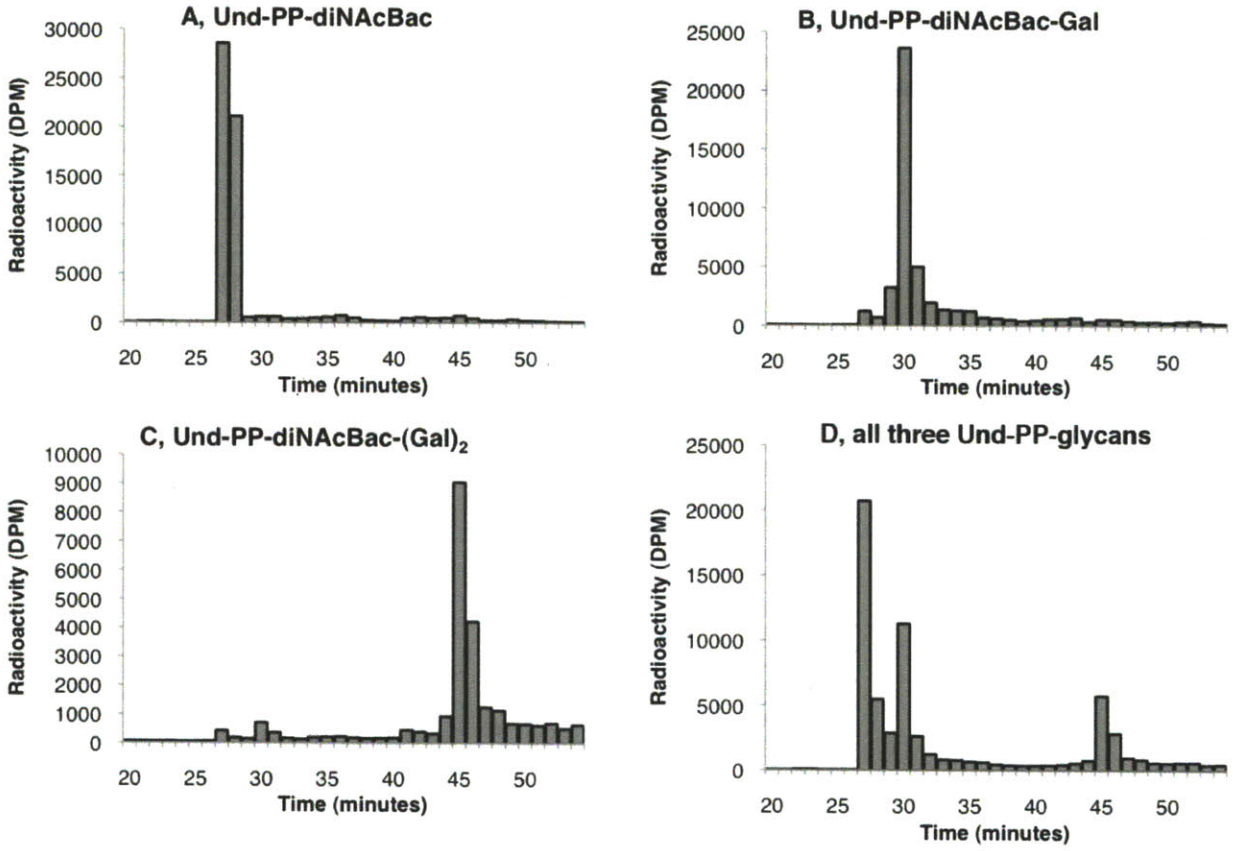


Figure 3-5. Normal phase HPLC separation of radiolabeled glycosyltrasnferase products. The retention times of the products are as follows: *A*, Und-PP-[^{3}H]diNAcBac (27 minutes), *B*, Und-PP-diNAcBac-[^{3}H]Gal (30 minutes), and *C*, Und-PP-diNAcBac-Gal-[^{3}H]Gal (45 minutes). The HPLC trace in *D* shows all three Und-PP-linked glycans together. Fractions were eluted at 1 mL/min and radioactivity (DPM) was determined by scintillation counting.

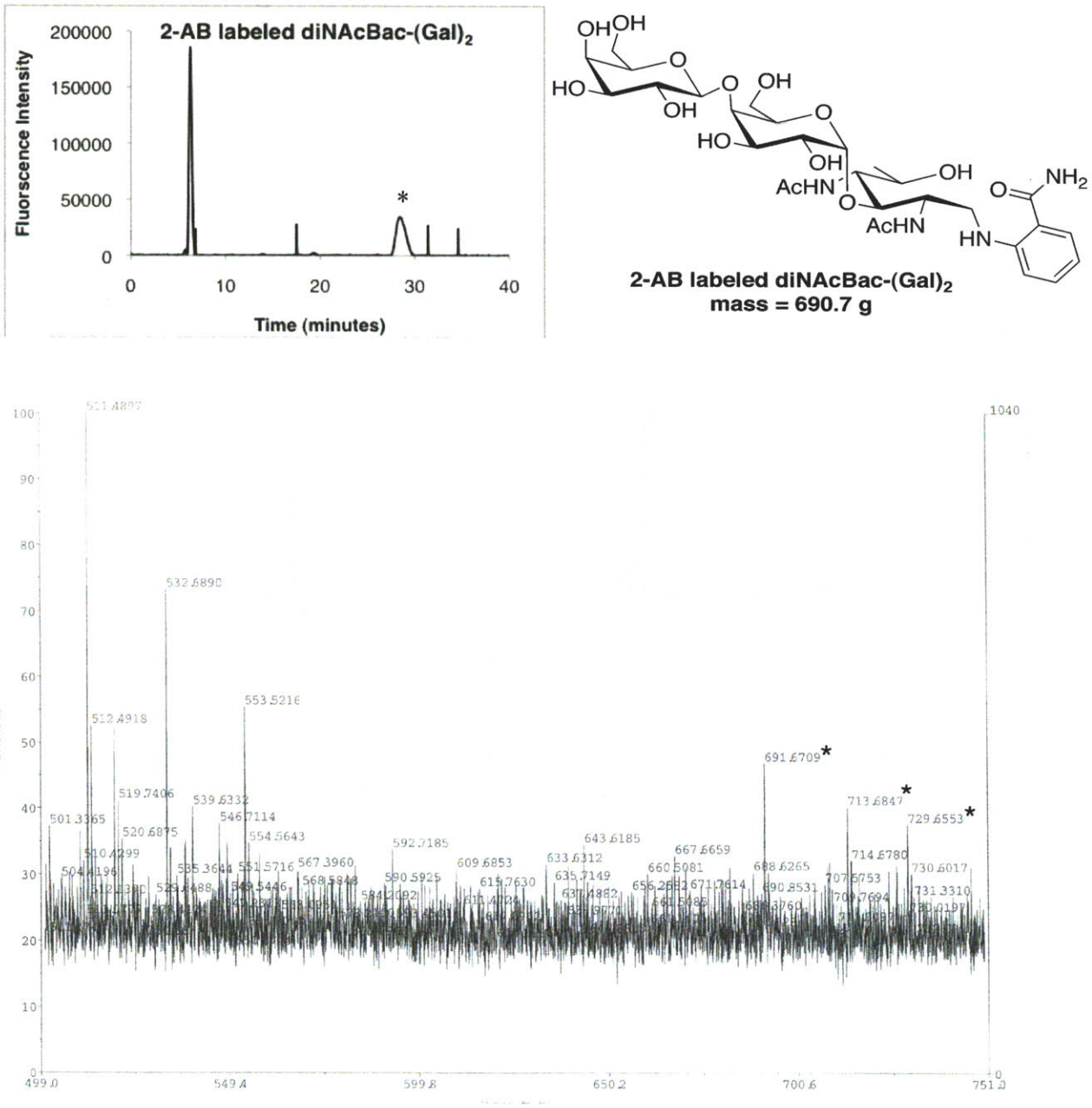


Figure 3-6. Normal phase HPLC with fluorescence detection of 2-AB labeled glycans.
Top, The 2AB-labeled trisaccharide product of PglB, PglA, and PglE elutes at 29 minutes. The product peak is marked with an asterisk. *Bottom*, MALDI MS confirmed the identity of the separated fluorescent product. The masses corresponding to [M + H] (691.7), [M + Na] (713.7) and [M + K] (729.7) are indicated by asterisks. The low masses (<600) are consistent with background from the 2,5-dihydroxybenzoic acid matrix. The same experiment was performed with diNAcBac-Gal (product of PglB and PglA, data not shown).

The substrate specificities of PglB, PglA and PglE were explored through the use of radioactivity-based assays. Organic extraction of the hydrophobic undecaprenyl-linked product allowed for quantification of the amount of radiolabeled sugar transferred to the undecaprenyl substrate.[18, 34] The undecaprenyl-phosphate required for the PglB reaction was generated *in situ* from undecaprenol and ATP with *S. mutans* undecaprenol kinase as previously described.[35] The isoprenyl-linked substrates for the other assays (Und-PP-diNAcBac for PglA and Und-PP-diNAcBac-Gal for PglE) were produced enzymatically and purified by NP-HPLC. These substrates are typically synthesized in 50 nmole reactions with 100 μL total volume. Production of Und-PP-diNAcBac typically proceeds with 20-40% conversion, whereas the Und-PP-diNAcBac-Gal biosynthesis results in turnover of 50-60%.

The activities of the three enzymes were screened with UDP-Glc, UDP-Gal, UDP-GlcNAc, UDP-GalNAc and in the case of PglB, UDP-diNAcBac (Figure 3-7). Enzyme and substrate concentrations were initially varied to identify conditions in which appreciable turnove could be observed in the first 10 minutes of the reaction and these conditions were used in Figure 3-7. In addition, the ability of PglB and PglA to distinguish between UDP-4-amino and UDP-diNAcBac was evaluated through a coupled assay (Figure 3-8). In all cases, the enzymes were highly specific for the corresponding predicted sugar substrate; PglB exclusively transferred phospho-diNAcBac, while PglA and PglE transferred only Gal (Figure 3-7).

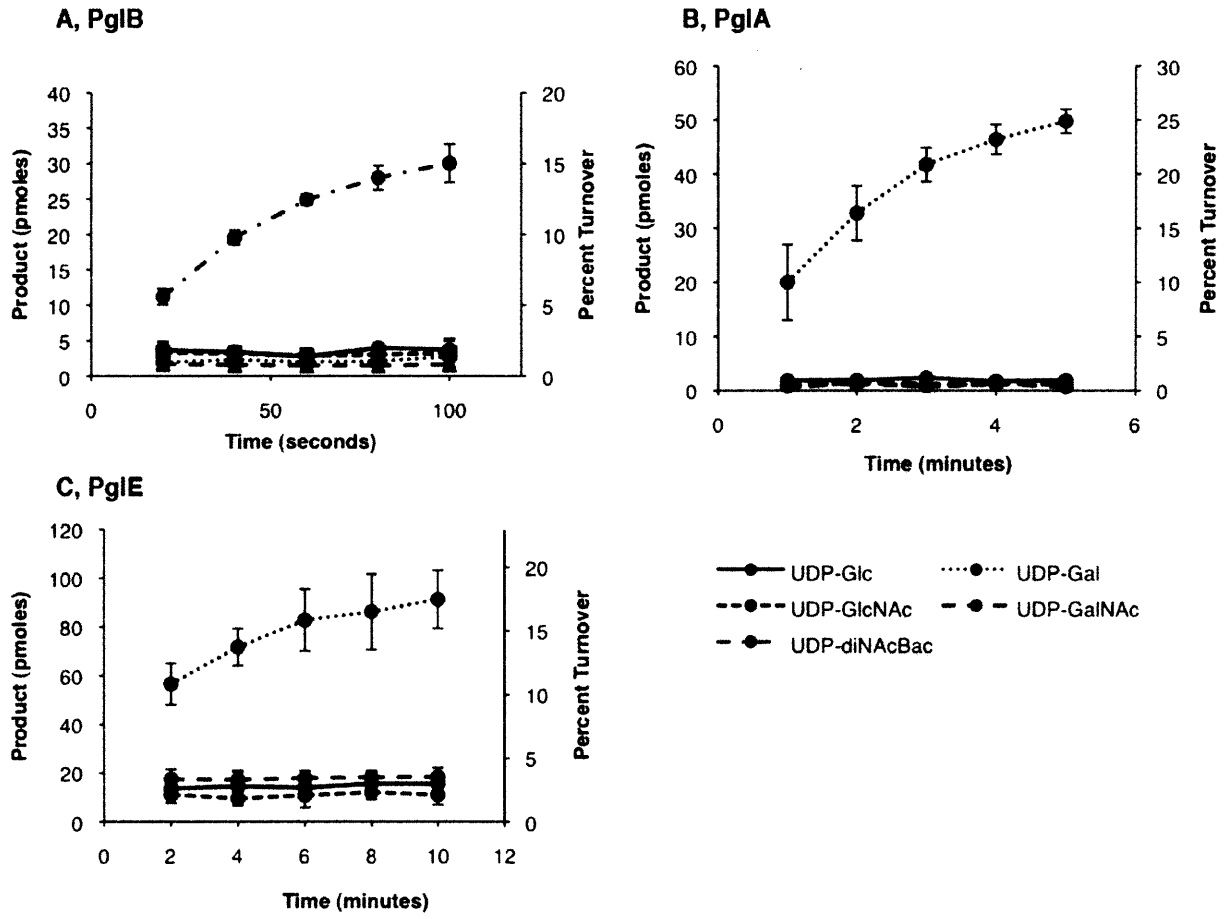


Figure 3-7. The UDP-specificity preferences of PglB, PglA, and PglE.
A, PglB prefers UDP-diNAcBac; *B*, PglA prefers UDP-Gal; *C*, PglE, prefers UDP-Gal. The reactions were carried out in a volume of 100 μL. The assays were performed in triplicate and the error bars indicate standard deviation.

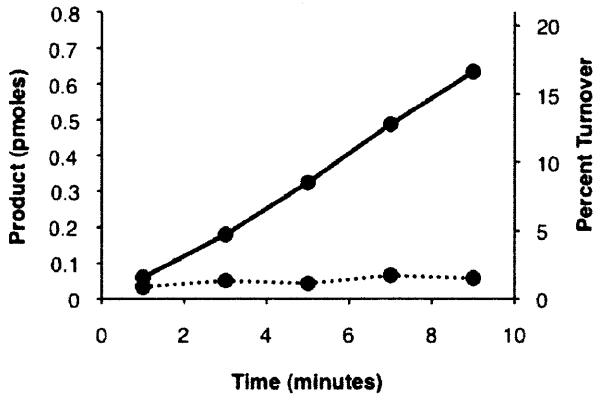


Figure 3-8. PglB and PglA specificity assay with UDP-4-amino.
To test the ability of PglB-GTD and PglA to distinguish between UDP-diNAcBac (solid line), and UDP-4-amino (dotted line), the activities of PglB and PglA were assessed simultaneously via a coupled assay based on radioactivity. As observed in the graph, the pathway is specific for the diNAcBac substrate.

Facile enzymatic synthesis of Und-PP-diNAcBac-Gal and Und-PP-diNAcBac-GalNAc produced by the *Neisseria* and *C. jejuni* pathways, respectively, allowed for examination of the substrate specificity of PglE for the acceptor oligosaccharide. Somewhat surprisingly, PglE is able to add a Gal residue to both the native substrate, Und-PP-diNAcBac-Gal, and to the *C. jejuni* substrate Und-PP-diNAcBac-GalNAc (Figure 3-9). This confirms *in vivo* studies in which the *C. jejuni* PglA was expressed in *N. gonorrhoeae* and the resultant trisaccharide (diNAcBac-GalNAc-Gal) was observed as a covalent pilin modification.[10]

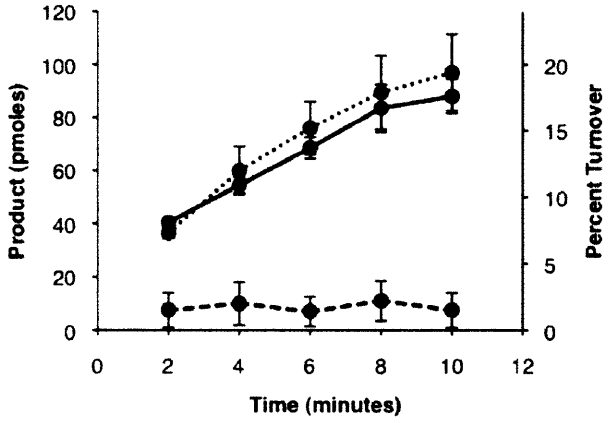


Figure 3-9. The Und-PP-glycan specificity of PglE.
PglE will transfer Gal onto Und-PP-diNAcBac-Gal (dotted line) and Und-PP-diNAcBac-GalNAc (solid line), but not Und-PP-diNAcBac-Glc (dashed line).

Characterization of PglH, an alternative glycosyltransferase

Recently an alternative glycosyltransferase in *N. gonorrhoeae*, PglH, was identified and shown to transfer a hexose unit to Und-PP-diNAcBac (Figure 3-2).[13] *In vivo* characterization of the O-linked protein glycans in *N. gonorrhoeae* failed to establish the stereochemistry of this alternative saccharide addition, although the presence of Und-PP-linked disaccharide in a strain lacking GalE epimerase suggested that this hexose was either Glc or Man. The *in vitro* activity assay developed above for PglB, PglA and PglE was extremely useful for examining PglH sugar specificity.

The PglH allele from *N. meningitidis* strain Z2491 was expressed and purified as an amino-terminal maltose binding protein (MBP)-fusion to assist with protein solubility and purification (Figure 3-10). Purified MBP-PglH was incubated with Und-PP-diNAcBac and radiolabeled versions of five different NDP-sugars: UDP-Glc, UDP-Gal, UDP-GlcNAc, UDP-GalNAc, and GDP-Man. All five of these activated sugars are found endogenously in neisserial

species and could serve as the native substrate for PglH. As shown in Figure 3-11, PglH is specific for transfer of glucose, showing that UDP-Glc is the preferred substrate of this enzyme.

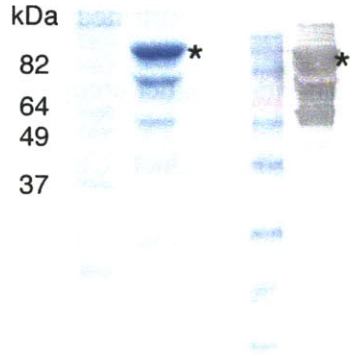


Figure 3-10. SDS-PAGE and anti-MBP Western blot analysis of purified MBP-PglH.
MBP-PglH is the top band indicated by an asterisk. Expression of MBP-PglH (marked by asterisk) results in the formation of two truncation products that co-purify with the full-length protein.

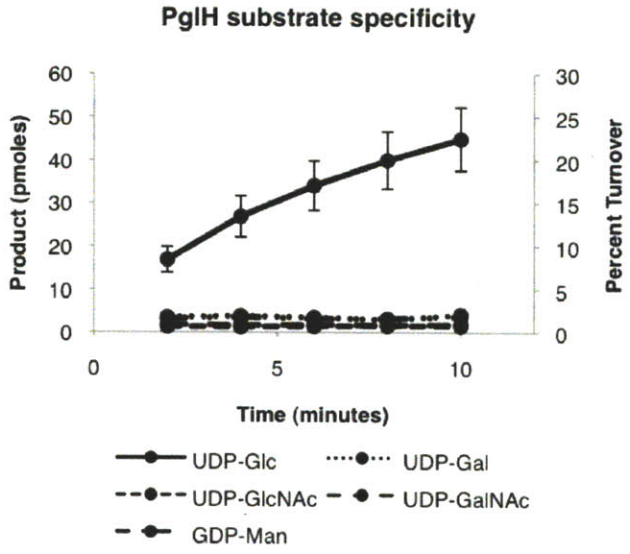


Figure 3-11. UDP-sugar specificity of PglH.
PglH is active in the presence of UDP-Glc (solid line), but not UDP-Gal, UDP-GlcNAc, UDP-GalNAc, or GDP-Man.

To further establish the identity of the radiolabeled product observed in the assay, the PglH product, Und-PP-diNAcBac-Glc, was generated and treated under acidic conditions to

hydrolyze the glycosyl diphosphate linkage and the resultant disaccharide was labeled with 2-aminobenzamide. The fluorescently-labeled sugar was purified using normal phase HPLC and the identity of the disaccharide peak was confirmed by MALDI MS providing definitive evidence that PglH produces a disaccharide product comprising diNAcBac-Glc (Figure 3-12). In addition, analysis of the radiolabeled PglH product, Und-PP-diNAcBac-[^{3}H]Glc, by NP-HPLC revealed that the retention time (30 minutes) was very similar to the Und-PP-diNAcBac-[^{3}H]Gal retention time (29-30 minutes) (Figure 3-13).

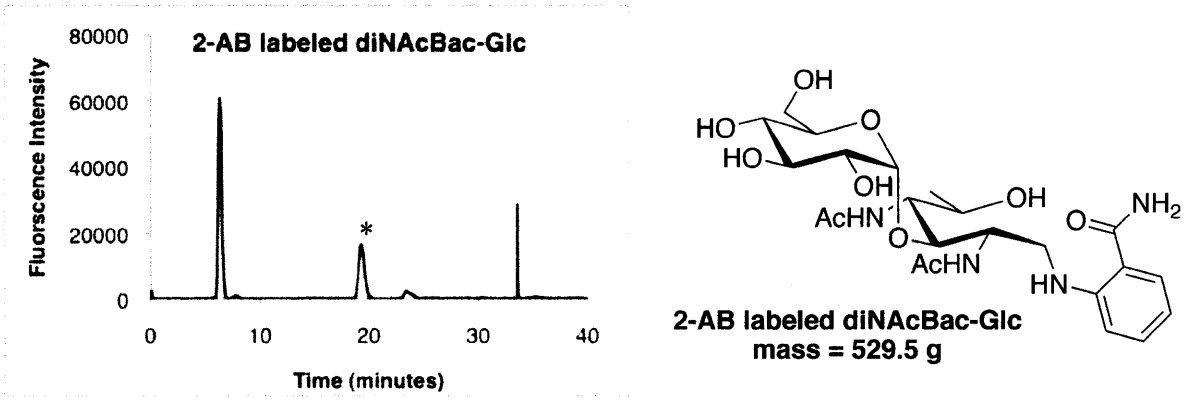


Figure 3-12. NP-HPLC trace of 2-AB labeled diNAcBac-Glc.
The diNAcBac-Glc peak is designated with an asterisk and the mass of the labeled sugar was confirmed by MALDI MS. The first peak at 8 minutes is the unreacted 2-AB. In addition, random spikes (such as the one observed at 34 minutes) are often observed by the fluorescence detector in the presence and absence of sample and seem to be a malfunction of the detector.

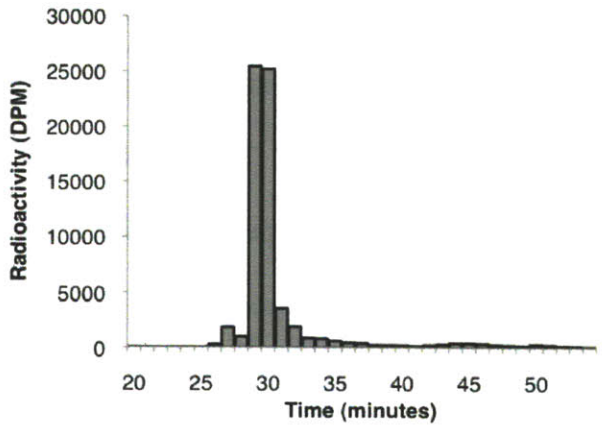


Figure 3-13. Normal phase HPLC separation of the product of PglH.
The radiolabeled Und-PP-diNAcBac-[^{3}H]Glc had retention time of 29-30 minutes. Fractions were eluted at 1 mL/min and radioactivity (DPM) was determined by scintillation counting.

Prior *in vivo* evidence suggested that unlike the *C. jejuni* disaccharide, the PglH product was not further modified by the third glycosyltransferase PglE.[13] This result was validated by the *in vitro* specificity assay described above for PglE establishing that PglE was unable to transfer Gal to Und-PP-diNAcBac-Glc (Figure 3-9).

Significant glycan diversity in *N. gonorrhoeae*

The *C. jejuni* and *N. gonorrhoeae* pathways diverge after the synthesis of Und-PP-diNAcBac. The *C. jejuni* pathway continues to N-linked glycan assembly with the successive addition of five α-(1,4)-linked GalNAc units and a branching Glc unit. However, while the *C. jejuni* N-linked heptasaccharide is highly conserved, *N. gonorrhoeae* strains display high O-linked glycan diversity. Strains have been identified which contain not only O-linked disaccharide (Gal-α-(1,3)-diNAcBac) and trisaccharide (Gal-β-(1,4)-Gal-α-(1,3)-diNAcBac) produced by PglA and PglE, but also an alternate disaccharide (Glc-α-(1,3)-diNAcBac) produced by PglH.[13] Further glycan modification occurs from the addition of O-acetyl groups by PglI.[10] In

addition, an alternate allele (*pglB2*) has been identified in *N. meningitidis* that contains a domain proposed to transfer a glyceroyl moiety instead of an acetyl group to produce 4-glyceramido-2-acetamido-2,4,6-trideoxy-α-D-hexose (GATDH).[36] This combination of biosynthetic enzymes allows neisserial strains to display a glycan repertoire with at least 13 identified glycan permutations.[13] Additional glycan variation can occur within a single strain as phase variation of the genes encoding for PglA, PglE and PglH acts as another mode of glycan regulation.[10]

In light of the significant amount of protein glycan variation present within strains of *N. gonorrhoeae*, it is surprising that the glycosyltransferases display such strict specificity (Figures 3-7 and 3-8). We have demonstrated that PglB, PglA, PglE, and PglH are specific for the native substrate (UDP-diNAcBac, UDP-Gal or UDP-Glc) and will not accept any other nucleotide-activated sugar commonly found *in vivo* under these assay conditions, even though one of the alternate substrates contains only a single stereochemical change (UDP-Glc vs. UDP-Gal) and another contains only an additional acetamido group (UDP-GalNAc vs. UDP-Gal). These results suggest that glycan identity is regulated at the level of biosynthesis and that these enzymes have evolved to selectively catalyze reactions in the milieu of intracellular NDP-sugars.

Along the same lines, PglE transferred a Gal unit onto native Und-PP-diNAcBac-Gal, but it showed little activity with Und-PP-diNAcBac-Glc, the alternate PglH disaccharide (Figure 3-9). PglE has evolved to detect the stereochemical difference between Glc and Gal, which is consistent with the model that PglA and PglE have evolved in tandem to produce a trisaccharide that is structurally distinct from the disaccharide produced by PglH.[13] In contrast, it is surprising that PglE would recognize the *C. jejuni* substrate Und-PP-diNAcBac-GalNAc (Figure 3-9), but it is consistent with the hypothesis that the glycosyltransferases exhibit specificity relative to other substrates present in the organism. PglE may not have developed selectivity against the

additional acetamido group in the *C. jejuni* disaccharide, because it is not found in the native *N. gonorrhoeae* glycome.

Conclusions

The O-linked protein glycosylation pathway in *N. gonorrhoeae* expands the repertoire of undecaprenol-dependent processes in bacteria. This chapter describes the characterization of *N. gonorrhoeae* glycosyltransferases, which build glycans on undecaprenyl-diphosphate prior to protein transfer. In addition, the biochemical functions of three enzymes required for biosynthesis of UDP-diNAcBac are established and the stereochemistry of this unusual sugar is shown for the first time to be the same as the first sugar in the N-linked protein glycosylation pathway in *C. jejuni*. The *N. gonorrhoeae* constitutes an unusual glycosylation system, not only because it produces an O-linked glycan through an undecaprenyl-diphosphate intermediate, but also because it has a high amount of glycan diversity demonstrated here by the characterization of PglH, which produces an alternate undecaprenyl-diphosphate disaccharide. Further study of this fascinating pathway is required to understand the physiological role of glycan diversity in *N. gonorrhoeae* pathogenesis and survival.

Acknowledgements

Michael Morrison was a wonderful collaborator on this project and he expertly characterized the enzymes PglD, PglC and PglB-ATD described in this chapter. In addition, the Norwegians, Professor Michael Koomey, Finn Erik Aas, and Bente Børud, from the University

of Oslo have been invaluable collaborators and we are glad that they introduced us to the new world of O-linked protein glycosylation in bacteria. I would also like to thank Dr. Matthieu Sainlos, Dr. Langdon Martin, and Dr. Cliff Stains for obtaining the MALDI MS data. In addition, Dr. Jeff Simpson of the MIT Department of Chemistry Instrumentation Facility was extremely helpful in obtaining the NMR characterization of UDP-diNAcBac. Finally, I very grateful for the many productive discussions of this manuscript and chapter with members of the Imperiali lab including Dr. Jerry Troutman, Marcie Jaffee, and Michelle Chang.

Experimental Procedures

Materials

All radioactive materials and undecaprenol were obtained from American Radiolabeled Chemicals. UDP-4-amino and UDP-diNAcBac were prepared as previously described[20] by enzymes from *C. jejuni* or as described herein by PglD, PglC and PglB. All other chemicals were obtained from Sigma-Aldrich unless stated otherwise. Radioactivity was determined using a LS6500 Beckman Scintillation Counter; organic samples were dried and resuspended in 200 μL Solvable™ (Perkin-Elmer) and 5 mL of scintillation fluid (Opti-Fluor, Perkin-Elmer). Aqueous samples were mixed with 5 mL of Ecolite(+)™ (MP Biomedicals) prior to scintillation counting.

Preparation of genetic constructs

The genes *pglD*, *pglC*, *pglB*, and *pglA* were PCR amplified from the *N. gonorrhoeae* strain MS11,[6, 10, 11] while *pglE* was amplified from *N. gonorrhoeae* strain FA 1090 and *pglH* was amplified from the *N. meningitidis* strain Z2491. The PCR products of *pglD*, *pglC*, *pglB*, *pglA*, *pglO,* and *pglE* were cloned into BamH I/Xho I in the pET-24a(+) vector (Novagen). The *pilE* and *pglH* genes were cloned into Nde I/Xho I in the pET-24a(+) vector (Novagen). The Xho I site was inserted prior to the stop codon to encode for a His_6 tag at the C-terminal end of each protein. The constructs were sequenced to verify the success of the PCR reactions.

The acetyltransferase domain of PglB (PglB-ATD) was identified through sequence homology with the related *C. jejuni* protein, PglD(*Cj*). The gene encoding the domain was amplified from the full-length gene using the forward primer 5'-CGCGGATCCATGGCGGGGAATCGCAAACTCG-3' and the reverse primer 5'-GCAACCCGGCAAAGCCCCTTTAGCTCGAGCGG-3' to generate a gene encoding the acetyltransferase domain. The gene was inserted into BamH I/Xho I in a modified pET-30b(+)

vector which contains an N-terminal His_8 tag followed by a Tobacco Etch Virus (TEV) protease site. Also, *pglH* was amplified by PCR and inserted into BamH I/Xho I in the pMAL-c2X vector. This construct encoded for the addition of an N-terminal maltose binding protein (MBP).

Expression of proteins

In general, all proteins (PglD, PglC, PglB, PglB-ATD, PglA, PglE, PglH, and *S. mutans* undecaprenol kinase[35]) were expressed heterologously in *E. coli* BL21 cells (Agilent). PglD, PglC and PglB-ATD were expressed in the BL21(DE3) pLysS strain; all other proteins were expressed in the BL21-Gold (DE3) strain. A typical expression protocol involved preparation of an overnight culture of cells (5 mL), which was used to inoculate 1 L of LB media with shaking at 37 °C. After the cells reached an optical density of ~0.8 absorbance units at 600 nm, the temperature was lowered to 16 °C and the cells were induced with 0.5 mM iso-β,D-thiogalactosylpyranoside (IPTG). After 16-18 hours of incubation, the cells were harvested and the pellets were stored at -80 °C.

Protein purification

In general, all steps of protein purification were carried out at 4 °C. Protein concentrations were determined with the appropriate extinction coefficients at a UV absorbance of 280 nm.

A cell pellet generated from 1 L culture of the soluble proteins, PglC or PglB-ATD, was resuspended in 50 mL of ice-cold 50 mM HEPES (pH 7.4) and 100 mM NaCl (Buffer A), supplemented with 30 mM imidazole. The cells were lysed by two intervals of sonication for 1 minute each at an amplitude of 55% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. In the case of PglC, 200 μM pyridoxal-5'-phosphate was also added to the buffer. The lysate was

cleared by centrifugation (145,000 x g) for 45 min. Cleared lysate was mixed with 2 mL of Ni-nitrilotriacetic acid (Ni-NTA) resin (Qiagen), tumbled for 4 hours, and then packed into a 20 mL Econo-Pac polypropylene column (BioRad). Using gravity flow, the resin-bound protein was washed with 10 column volumes of Buffer A containing 30 mM imidazole. The resin was further washed with 20 column volumes of Buffer A supplemented with 40 mM imidazole and then 10 column volumes of Buffer A containing 60 mM imidazole. The protein was eluted in Buffer A supplemented with 250 mM imidazole and 1 mL fractions were collected. Fractions containing purified material were assessed by SDS-PAGE (12%) and Western blot analysis probing for the His_6 tag (Figure 3-3). Pooled fractions of PglC and PglB-ATD were dialyzed against Buffer A, concentrated, supplemented with a final glycerol concentration of 15% and frozen at -80 °C.

Purification of the glycosyltransferase PglA was similar to PglC and PglB-ATD with a few exceptions. A buffer containing 50 mM Tris (pH 8.0) and 150 mM NaCl (Buffer B) was used instead of Buffer A and the cells were incubated with 1% Triton X-100 for 20 minutes immediately following lysis and prior to centrifugation. In addition, 5% glycerol was added to all buffers. Following elution, the most concentrated 1.5 mL fraction as determined by SDS-PAGE was desalted using a Hi-Trap desalting cartridge (GE Healthcare) with Buffer B and stored at -20°C in 30% glycerol (Figure 3-3).

To purify the membrane-associated proteins (PglD, PglB, and PglE), a cell envelope fraction (CEF) was first prepared. The cells were thawed in 40 mL of buffer per L of cell culture. The cells were lysed on ice by two intervals of sonication for 75 seconds at an amplitude of 70% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. PBS supplemented with 200 μM NAD^+ was

used for PglD and Buffer B with 1 mg/mL lysozyme was used for PglB and PglE. Cellular debris was cleared by centrifugation at 9000 x g for 45 minutes. The resulting supernatant was transferred to a clean centrifuge tube and subjected to centrifugation at 145,000 x g for 65 min to pellet the CEF. For PglB and PglE, the CEFs were resuspended in half the volume of the unlysed cell pellet weight (i.e. 1.5 mL was used for 3 g cell pellet). The CEF was aliquoted and stored at -80 °C (Figure 3-3). The CEFs of PglB and PglE were used in all glycosyltransferase assays.

PglD was further purified from the CEF. The CEF was homogenized in 10 mL of PBS (pH 7.0) containing 1% Triton X-100 and 200 μM NAD^+ per liter of cell culture. Each CEF was incubated with detergent for several hours and then centrifuged again (145,000 x g) to remove insoluble material. The resultant supernatants were incubated with 0.5-2 mL of Ni-NTA resin for 1-2 hours; the resins were washed as previously described with the addition of 0.1% Triton X-100 to the wash and elution buffers. The proteins were eluted from the resin in 1 mL fractions. Pooled fractions of PglD were dialyzed against PBS containing 200 μM NAD^+ and 0.1% Triton X-100, supplemented with a final glycerol concentration of 30%, and frozen at -80°C (Figure 3-3). For the biosynthesis of UDP-diNAcBac, PglB was purified in a manner similar to PglD, except that NAD^+ was not added to the purification buffers.

PglH was expressed as an MBP-fusion protein and purified as described elsewhere[13] (Figure S1, lane 8). In addition, the undecaprenol kinase from *S. mutans* was expressed and purified as cell envelope fractions as described previously.[35]

The glycosyltransferase PglH was expressed as a maltose binding protein-fusion protein and the protein was purified using amylose resin. A 1 L cell pellet was solubilized in 40 ml of buffer containing 50 mM Tris (pH 8.0), 150 mM NaCl, and 5% glycerol and then lysed by

sonication. The lysate was cleared by centrifugation (145,000 x g) for 65 min at 4°C. Cleared lysate was mixed with 2 mL of amylose resin (New England Biolabs) and tumbled for 30 minutes at 4°C and then packed into a 20 mL Econo-Pac polypropylene column (BioRad). Using gravity flow, the resin-bound protein was washed with 20 column volumes of the lysate buffer. The protein was then eluted in the same buffer supplemented with 10 mM maltose and 10 mL fractions were collected. Fractions containing purified material were assessed by SDS-PAGE (12%). The purified protein contains two lower molecular weight bands that are both immunoreactive with anti-MBP antibody suggesting that they are truncation products produced during overexpression. The full-length protein is the most abundant band and consists of >60% of the total protein content. The first three fractions were concentrated to ~3-4 ml of 30 μM protein concentration with Amicon Ultra-15 Centrifugal Filter units (Millipore) and stored at -20°C in the presence of 30% glycerol.

Acetyltransferase (PglB-ATD) Activity Assay

Determination of the kinetic constants for PglB-ATD was carried out using Ellman's reagent, 5,5'-dithio-*bis*-(2-nitrobenzoic acid), in a continuous fashion. Ellman's reagent was utilized to quantify substrate turnover as monitored by measuring conversion of acetyl coenzyme A (AcCoA) to CoASH using the released TNB chromophore (λ_{max} = 412 nm, $\varepsilon_{max,\ pH\ 8.0}$ = 14,150 $M^{-1}\ cm^{-1}$). The *in vitro* assay contained 50 mM HEPES (pH 7.4), 1 mM EDTA, 1 mM DTNB, and 25 nM PglB-ATD in a quartz cuvette. The substrate concentrations of AcCoA and UDP-4-amino were varied separately to determine kinetic parameters using initial velocity measurements while keeping the other substrate at saturation. The reaction was initiated with the UDP-4-amino substrate and took place at room temperature over a 200 second time period. The

absorbance change at 412 nm was measured. A blank reaction lacking UDP-4-amino was prepared as a background control. Steady-state rate parameters were calculated from equation **1** using the program GraFit 6.0.12 (Erithacus Software).

$$v = V_{max}[S]/K_m + [S] \qquad (1)$$

Aminotransferase (PglC) Activity Assay

The aminotransferase reaction was assayed by coupling generation of UDP-4-amino from the PglC reaction to the acetyltransferase activity of PglD from *C. jejuni* (*Cj*) producing CoASH, which was detected by Ellman's reagent in a similar fashion to the PglB-ATD assay. In a flat bottom 96-well plate (Nunc), 50 mM HEPES (pH 7.4), 1 mM PglD(*Cj*), 400 mM AcCoA, and 400 nM PglC were added. Since PglC activity was coupled to the turnover of the acetyltransferase PglD(*Cj*), addition of excess PglD(*Cj*) ensured that the initial velocity measurements were dependent only upon PglC activity. The concentrations of L-glutamate and UDP-4-keto were varied separately to determine kinetic parameters using initial velocity measurements while keeping the other substrate at saturation. Interference of Ellman's reagent with PglC activity required the implementation of a discontinuous assay in which reactions were initiated with L-glutamate and quenched over a 30 minute time period with 20% n-propanol, 2 mM DTNB, and 1 mM EDTA. The absorbance at 415 nm was followed on an Ultramark EX microplate imaging system (BioRad). A blank reaction without L-glutamate was set up as a background control.

Biosynthesis and stereochemical assignment of UDP-DATDH

In order to biosynthesize UDP-DATDH, 0.3 mg of PglD (bound to Ni-NTA resin) was added to 15 mL buffer containing 50 mM HEPES (pH 7.4), 100 mM NaCl, 200 mM NAD^+, and 30 mg UDP-GlcNAc. The reaction was carried out at room temperature for 12 hours with gentle

rocking. Once conversion to the UDP-4-keto sugar was complete, as verified by capillary electrophoresis as described previously,[20] the reaction was filtered and the flow-through collected. The filtrate containing the UDP-4-keto sugar was supplemented with 15 mg PglC (bound to Ni-NTA resin), 20 mM L-glutamate, and 200 μM pyridoxal-5'-phosphate. This reaction was filtered after rocking for 18 hours at room temperature and reaching 80% conversion to the UDP-4-amino sugar. The filtrate was supplemented with 0.2 mg purified, full-length PglB (bound to Ni-NTA resin) and 1.2 mM AcCoA and allowed to react at room temperature with rocking for 12 hours. The slurry was filtered and the flow-through containing the UDP-DATDH sugar was collected. Purification and NMR characterization of the final UDP-DATDH product was completed as previously described.[20]

Preparation of radiolabeled Und-PP-linked substrates

In general, radiolabeled Und-PP-substrates were prepared at two different specific activity levels. A higher specific activity was used for the OTase assay and analysis by normal phase-high performance liquid chromatography (NP-HPLC) and a lower specific activity was appropriate for the glycosyltransferase assays.

Und-PP-[^{3}H]diNAcBac was enzymatically synthesized using *S. mutans* undecaprenol kinase[37] and PglB. An undecaprenol kinase from *N. gonorrhoeae* has not been characterized, and thus the undecaprenol kinase from *S. mutans*[35] was used a tool to affect the undecaprenol phosphorylation *in situ*. A typical reaction contained 3% DMSO, 1% Triton X-100, 50 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.0), 500 μM undecaprenol, 1 mM ATP, 500 μM UDP-4-amino, 500 μM [^{3}H]AcCoA (20 mCi/mmol), 15-20 μL of undecaprenol kinase CEF, 15-20 μL PglB CEF and water to a final volume of 100 μL. The reaction was modified to prepare Und-PP-[^{3}H]diNAcBac with high specific activity by adjusting the undecaprenol and UDP-4-amino

concentrations to 100 μM and the [^{3}H]AcCoA concentration to 4.5 μM (20 Ci/mmol). After incubation at room temperature for 2 hours, the reactions were quenched into 1 mL 2:1 $CHCl_3$:MeOH and extracted three times with 400 μL of an aqueous extract prepared by dissolving 1.83 g of potassium chloride in 235 mL water, 240 mL chloroform, and 15 mL methanol. The organic layer containing the Und-PP-[^{3}H]diNAcBac product was dried down and purified using NP-HPLC as described below.

Und-PP-diNAcBac-[^{3}H]Gal was prepared in a similar manner to Und-PP-diNAcBac. The reaction components are as described above for Und-PP-diNAcBac with the following exceptions; 500 μM UDP-diNAcBac was added instead of UDP-4-amino and AcCoA, and 2 μM PglA and 500 μM UDP-[^{3}H]Gal (20 mCi/mmol) were added to affect the transfer of the galactosyl unit, which is the second sugar in the glycan. To prepare Und-PP-diNAcBac-[^{3}H]Gal with higher specific activity, undecaprenol and UDP-diNAcBac concentrations were lowered to 100 μM and the UDP-[^{3}H]Gal concentration was adjusted to 4.5 μM (20 Ci/mmol). The reactions were quenched after two hours and extracted as described above.

The synthesis of Und-PP-diNAcBac-Gal-[^{3}H]Gal utilized unlabeled Und-PP-diNAcBac-Gal, which was prepared as described above with the exception that UDP-Gal was not radioactive. A typical biosynthesis reaction contained 3% DMSO, 0.05% Triton X-100, 50 mM $MnCl_2$, 50 mM HEPES (pH 7.5), 20 μM Und-PP-diNAcBac-Gal, 20 μM UDP-[^{3}H]Gal (20 mCi/mmol), 20 μL PglE CEF and water to a final volume of 100 μL. To prepare the substrate with higher specific activity, the UDP-Gal concentration was lowered to 4.5 μM UDP-[^{3}H]Gal (20 Ci/mmol). The reactions were quenched after two hours and extracted as described above.

Und-PP-diNAcBac-[^{3}H]Glc was prepared from unlabeled Und-PP-diNAcBac. The reaction contained 3% DMSO, 0.1% n-β-D-dodecylmaltoside (DDM), 50 mM $MgCl_2$, 30 mM

Tris (pH 8.0), 20 μM Und-PP-diNAcBac, 20 μM UDP-[^{3}H]Glc (20 mCi/mmol), 10 μM PglH and water to a final volume of 100 μL. The substrate was also prepared with a higher specific activity by lowering the UDP-[^{3}H]Glc concentration to 4.5 μM (20 Ci/mmol).

The *C. jejuni* substrates (Und-PP-diNAcBac-[^{3}H]GalNAc and Und-PP-diNAcBac-GalNAc-[^{3}H]GalNAc) for the OTase reactions were prepared as previously described[17-19, 34, 35] with similar specific activities (20 Ci/mmol) to the *N. gonorrhoeae* OTase substrates.

Normal phase HPLC purification of Und-PP-linked substrates

The dried Und-PP-linked substrates were purified via NP-HPLC with a Varian Microsorb column using the previously described gradient.[34] The substrates were resolubilized in 100 μL of 4:1 $CHCl_3$:MeOH for injection onto the column. Fractions of 1 mL were collected and 10 μL of each fraction was mixed with 200 μL Solvable™ for detection of radioactivity. The fractions containing substrate were combined, aliquoted and stored at -20 °C.

To obtain the NP-HPLC analytical traces, Und-PP-linked glycan fractions were re-solubilized in 4:1 $CHCl_3$:MeOH and 100 μL of the appropriate sample was injected onto the column. The 1 mL elution fractions were dried completely and resuspended in 200 μL Solvable™ for scintillation counting.

Preparation and analysis of 2-AB labeled oligosaccharides

Unlabeled versions of Und-PP-diNAcBac-Gal and Und-PP-diNAcBac-$(Gal)_2$ were prepared in an identical manner as the radiolabeled substrates, except that unlabeled substrates were used in all reactions. The oligosaccharides were labeled with 2-aminobenzamide as previously described[18, 38] and purified using the GlykoNSep column (Prozyme). The appropriate peaks were collected and matrix-assisted laser desorption ionization mass spectrometry (MALDI MS) was used to determine the mass of the 2-AB labeled glycans.

Glycosyltransferase substrate specificity assays

To determine the UDP-sugar specificity of PglB, PglA, PglE, and PglH, radioactivity-based were performed with a variety of UDP-linked sugar substrates.

The ability of PglB to transfer UDP-diNAcBac, UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc was analyzed. The activity of PglB was coupled to the action of *S. mutans* undecaprenol kinase in order to provide the undecaprenyl-phosphate *in situ.*[35] The specificity assays included 3% DMSO, 1% Triton X-100, 50 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.0), 20 μM undecaprenol, 1 mM ATP, 2 μM UDP-[^{3}H]sugar (20 mCi/mmol), 20 μL of undecaprenol kinase CEF, 5 μL PglB CEF and water to a final volume of 100 μL. In the case of UDP-diNAcBac transfer, 2 μM UDP-4-amino and 2 μM [^{3}H]AcCoA (20 mCi/mmol) were included to assay both activities of the bifunctional PglB, which carries out transfer of the acetyl group to UDP-4-amino and transfer of phospho-diNAcBac to undecaprenyl-phosphate. The reactions were initiated with a mixture of ATP and PglB and were monitored by quenching 15 μL aliquots at 20, 40, 60, 80, and 100 seconds. The radioactivity present in the organic and aqueous layers was determined as described above.

The ability of PglA, PglE, and PglH to transfer UDP-Glc, UDP-Gal, UDP-GlcNAc, and UDP-GalNAc was also analyzed. PglH was also assayed with the GDP-linked Man, since it was known to transfer a hexose. In the case of PglA, the assays contained 3% DMSO, 0.05% Triton X-100, 50 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.0), 10 μM Und-PP-diNAcBac, 2 μM UDP-sugar (20 mCi/mmol), 0.1 μM PglA and water to a final volume of 100 μL. The reaction was monitored by quenching 15 μL at 1, 2, 3, 4, and 5 minutes. For PglE, the assays contained 3% DMSO, 0.05% Triton X-100, 50 mM $MnCl_2$, 50 mM HEPES (pH 7.5), 2 μM Und-PP-diNAcBac-Gal, 5 μM UDP-sugar (20 mCi/mmol), 5 μL PglE CEF and water to a final volume of

100 μL. The reaction was monitored by quenching 15 μL at 2, 4, 6, 8, and 10 minutes. For PglH, the assays contained 3% DMSO, 0.05% dodecyl maltoside, 50 mM $MgCl_2$, 30 mM Tris-Acetate, pH 8.0, 10 μM Und-PP-diNAcBac, 2 μM NDP-sugar (20 mCi/mmol), 9 μM MBP-PglH, and water to a final volume of 100 μl. The reactions were initiated by the addition of enzyme and were monitored by quenching 15 μl aliquots at 2, 4, 6, 8, and 10 minutes. For all three enzymes, the reactions were initiated by addition of enzyme and the radioactivity present in the organic and aqueous layers was determined as described above.

In addition, a coupled reaction was performed to test the ability of PglB and PglA to distinguish between UDP-4-amino and UDP-diNAcBac. The reaction components contained 3% DMSO, 1% Triton X-100, 40 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.0), 50 μM undecaprenol, 1 mM ATP, 500 μM UDP-4-amino, 500 μM AcCoA, 90 nM UDP-Gal (20 Ci/mmol), 5 μL undecaprenol kinase CEF, 10 μL PglB CEF and 1 μM PglA with a final volume of 100 μL. The reaction was initiated with a mixture of ATP and PglB. The extent of the reaction was monitored by quenching 15 μL aliquots at 1, 3, 5, 7, and 9 minutes. To test if PglB and PglA could recognize UDP-4-amino, a second reaction was prepared in which the AcCoA was omitted. The quenched aliquots were extracted as described above and the radioactivity present in the organic fraction was determined by scintillation counting.

References

1. Hansen, J. K.; Forest, K. T., Type IV pilin structures: insights on shared architecture, fiber assembly, receptor binding and type II secretion. *J. Mol. Microbiol. Biotechnol.* **2006,** 11, (3-5), 192-207.
2. Chamot-Rooke, J.; Mikaty, G.; Malosse, C.; Soyer, M.; Dumont, A.; Gault, J.; Imhaus, A. F.; Martin, P.; Trellet, M.; Clary, G.; Chafey, P.; Camoin, L.; Nilges, M.; Nassif, X.; Dumenil, G., Posttranslational modification of pili upon cell contact triggers *N. meningitidis* dissemination. *Science* **2011,** 331, (6018), 778-82.
3. Vik, A.; Aas, F. E.; Anonsen, J. H.; Bilsborough, S.; Schneider, A.; Egge-Jacobsen, W.; Koomey, M., Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. U S A* **2009,** 106, (11), 4447-52.
4. Karlyshev, A. V.; Everest, P.; Linton, D.; Cawthraw, S.; Newell, D. G.; Wren, B. W., The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **2004,** 150, (Pt 6), 1957-64.
5. Bieber, D.; Ramer, S. W.; Wu, C. Y.; Murray, W. J.; Tobe, T.; Fernandez, R.; Schoolnik, G. K., Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* **1998,** 280, (5372), 2114-8.
6. Banerjee, A.; Wang, R.; Supernavage, S. L.; Ghosh, S. K.; Parker, J.; Ganesh, N. F.; Wang, P. G.; Gulati, S.; Rice, P. A., Implications of phase variation of a gene (*pgtA*) encoding a pilin galactosyl transferase in gonococcal pathogenesis. *J. Exp. Med.* **2002,** 196, (2), 147-62.
7. Power, P. M.; Ku, S. C.; Rutter, K.; Warren, M. J.; Limnios, E. A.; Tapsall, J. W.; Jennings, M. P., The phase-variable allele of the pilus glycosylation gene *pglA* is not strongly associated with strains of *Neisseria gonorrhoeae* isolated from patients with disseminated gonococcal infection. *Infect. Immun.* **2007,** 75, (6), 3202-4.
8. Szymanski, C. M.; Burr, D. H.; Guerry, P., *Campylobacter* protein glycosylation affects host cell interactions. *Infect. Immun.* **2002,** 70, (4), 2242-4.
9. Banerjee, A.; Ghosh, S. K., The role of pilin glycan in neisserial pathogenesis. *Mol. Cell. Biochem.* **2003,** 253, (1-2), 179-90.
10. Aas, F. E.; Vik, A.; Vedde, J.; Koomey, M.; Egge-Jacobsen, W., *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol. Microbiol.* **2007,** 65, (3), 607-24.
11. Hegge, F. T.; Hitchen, P. G.; Aas, F. E.; Kristiansen, H.; Lovold, C.; Egge-Jacobsen, W.; Panico, M.; Leong, W. Y.; Bull, V.; Virji, M.; Morris, H. R.; Dell, A.; Koomey, M., Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of *Neisseria gonorrhoeae* type IV pili. *Proc. Natl. Acad. Sci. U S A* **2004,** 101, (29), 10798-803.
12. Stimson, E.; Virji, M.; Makepeace, K.; Dell, A.; Morris, H. R.; Payne, G.; Saunders, J. R.; Jennings, M. P.; Barker, S.; Panico, M.; et al., Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.* **1995,** 17, (6), 1201-14.
13. Børud, B., Viburiene, R., Hartley, M. D., Paulsen, B. S., Egge-Jacobsen, W., Imperiali, B., and Koomey, M., A novel glucosyltransferase, PglH, further expands the high glycan diversity in O-linked protein glycosylation of *Neisseria* species. **2011**, submitted.
14. Kumar, M.; Balaji, P. V., Comparative genomics analysis of completely sequenced microbial genomes reveals the ubiquity of N-linked glycosylation in prokaryotes. *Mol. Biosyst.* **2011**.

15. Szymanski, C. M.; Yao, R.; Ewing, C. P.; Trust, T. J.; Guerry, P., Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **1999,** 32, (5), 1022-30.
16. Young, N. M.; Brisson, J. R.; Kelly, J.; Watson, D. C.; Tessier, L.; Lanthier, P. H.; Jarrell, H. C.; Cadotte, N.; St Michael, F.; Aberg, E.; Szymanski, C. M., Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* **2002,** 277, (45), 42530-9.
17. Glover, K. J.; Weerapana, E.; Chen, M. M.; Imperiali, B., Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2006,** 45, (16), 5343-50.
18. Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. U.S.A.* **2005,** 102, (40), 14255-9.
19. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* **2005,** 12, (12), 1311-5.
20. Olivier, N. B.; Chen, M. M.; Behr, J. R.; Imperiali, B., *In vitro* biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* **2006,** 45, (45), 13659-69.
21. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
22. Takeuchi, K.; Taguchi, F.; Inagaki, Y.; Toyoda, K.; Shiraishi, T.; Ichinose, Y., Flagellin glycosylation island in *Pseudomonas syringae pv. glycinea* and its role in host specificity. *J. Bacteriol.* **2003,** 185, (22), 6658-65.
23. Horzempa, J.; Carlson, P. E., Jr.; O'Dee, D. M.; Shanks, R. M.; Nau, G. J., Global transcriptional response to mammalian temperature provides new insight into *Francisella tularensis* pathogenesis. *BMC Microbiol.* **2008,** 8, 172.
24. van Sorge, N. M.; Bleumink, N. M.; van Vliet, S. J.; Saeland, E.; van der Pol, W. L.; van Kooyk, Y.; van Putten, J. P., N-glycosylated proteins and distinct lipooligosaccharide glycoforms of *Campylobacter jejuni* target the human C-type lectin receptor MGL. *Cell. Microbiol.* **2009,** 11, (12), 1768-81.
25. Verma, A.; Arora, S. K.; Kuravi, S. K.; Ramphal, R., Roles of specific amino acids in the N terminus of *Pseudomonas aeruginosa* flagellin and of flagellin glycosylation in the innate immune response. *Infect. Immun.* **2005,** 73, (12), 8237-46.
26. Sharon, N.; Jeanloz, R. W., The diaminohexose component of a polysaccharide isolated from *Bacillus subtilis*. *J. Biol. Chem.* **1960,** 235, 1-5.
27. Sharon, N., Celebrating the golden anniversary of the discovery of bacillosamine, the diamino sugar of a *Bacillus*. *Glycobiology* **2007,** 17, (11), 1150-5.
28. Schaffer, C.; Scherf, T.; Christian, R.; Kosma, P.; Zayni, S.; Messner, P.; Sharon, N., Purification and structure elucidation of the N-acetylbacillosamine-containing polysaccharide from *Bacillus licheniformis* ATCC 9945. *Eur. J. Biochem.* **2001,** 268, (3), 857-64.
29. Schoenhofen, I. C.; Vinogradov, E.; Whitfield, D. M.; Brisson, J. R.; Logan, S. M., The CMP-legionaminic acid pathway in Campylobacter: biosynthesis involving novel GDP-linked precursors. *Glycobiology* **2009,** 19, (7), 715-25.

30. Parge, H. E.; Forest, K. T.; Hickey, M. J.; Christensen, D. A.; Getzoff, E. D.; Tainer, J. A., Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature* **1995,** 378, (6552), 32-8.
31. Schoenhofen, I. C.; McNally, D. J.; Vinogradov, E.; Whitfield, D.; Young, N. M.; Dick, S.; Wakarchuk, W. W.; Brisson, J. R.; Logan, S. M., Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.* **2006,** 281, (2), 723-32.
32. Vijayakumar, S.; Merkx-Jacques, A.; Ratnayake, D. B.; Gryski, I.; Obhi, R. K.; Houle, S.; Dozois, C. M.; Creuzenet, C., Cj1121c, a novel UDP-4-keto-6-deoxy-GlcNAc C-4 aminotransferase essential for protein glycosylation and virulence in *Campylobacter jejuni*. *J. Biol. Chem.* **2006,** 281, (38), 27733-43.
33. Bigge, J. C.; Patel, T. P.; Bruce, J. A.; Goulding, P. N.; Charles, S. M.; Parekh, R. B., Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. *Anal Biochem* **1995,** 230, (2), 229-38.
34. Troutman, J. M.; Imperiali, B., *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **2009,** 48, (12), 2807-16.
35. Hartley, M. D.; Larkin, A.; Imperiali, B., Chemoenzymatic synthesis of polyprenyl phosphates. *Bioorg. Med. Chem.* **2008,** 16, (9), 5149-5156.
36. Chamot-Rooke, J.; Rousseau, B.; Lanternier, F.; Mikaty, G.; Mairey, E.; Malosse, C.; Bouchoux, G.; Pelicic, V.; Camoin, L.; Nassif, X.; Dumenil, G., Alternative *Neisseria* spp. type IV pilin glycosylation with a glyceramido acetamido trideoxyhexose residue. *Proc. Natl. Acad. Sci. U S A* **2007,** 104, (37), 14783-8.
37. Lis, M.; Kuramitsu, H. K., The stress-responsive *dgk* gene from *Streptococcus mutans* encodes a putative undecaprenol kinase activity. *Infect. Immun.* **2003,** 71, (4), 1938-43.
38. O'Reilly, M. K.; Zhang, G.; Imperiali, B., *In vitro* evidence for the dual function of Alg2 and Alg11: essential mannosyltransferases in N-linked glycoprotein biosynthesis. *Biochemistry* **2006,** 45, (31), 9593-603.

Chapter 4: Characterization of the O-linked oligosaccharyltransferase from the protein glycosylation pathway in *Neisseria gonorrhoeae*

Most of the work in this chapter has been published in the following reference:

Hartley, M. D.; Morrison, M. J.; Aas, F.E.; Børud, B.; Koomey, M.; and Imperiali, B., Biochemical characterization of the O-linked glycosylation pathway in *Neisseria gonorrhoeae* responsible for biosynthesis of protein glycans containing N,N'-diacetylbacillosamine. *Biochem.* **2011**, in press.

Introduction

There has recently been a surge of interest and a number of published reports on bacterial protein glycosylation pathways including the N-linked glycosylation pathway in *Campylobacter jejuni* and the O-linked glycosylation pathway in *Neisseria gonorrhoeae*, which was introduced in the preceding chapter. Both pathways culminate in the key step of the pathway, in which the oligosaccharyltransferase (OTase) transfers the glycan from a polyprenyl diphosphate membrane anchor to a periplasmic protein (Figure 4-1).[1-4]

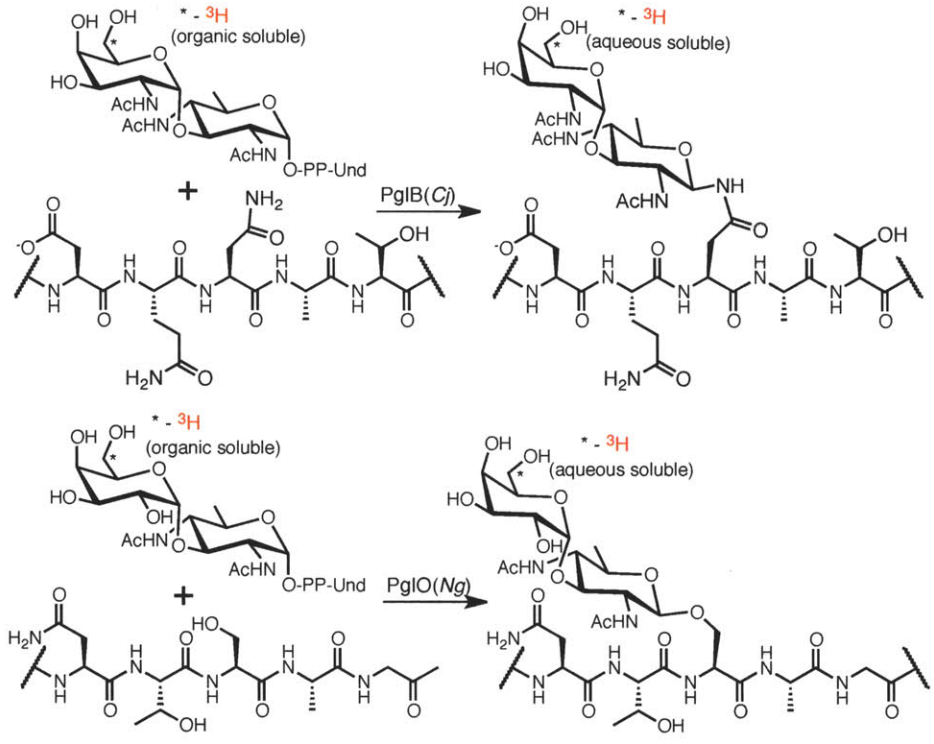


Figure 4-1. Enzymatic reactions mediated by PglB(*Cj*) and PglO(*Ng*).
In both cases, a disaccharide is depicted as the transferred glycan, but *in vivo*, PglB(*Cj*) transfers a heptasaccharide and PglO(*Ng*) transfers mono-, di- and trisaccharides. Top, PglB(*Cj*) catalyzes the transfer of GalNAc-α-(1,3)diNAcBac to the Asn carboxamide side chain. Bottom, PglO(*Ng*) transfers Gal-α-(1,3)-diNAcBac to Ser or Thr hydroxyl side chains.

The most important difference between these pathways is that the OTase in *C. jejuni*, PglB, generates asparagine-linked glycoproteins and the OTase in *N. gonorrhoeae*, PglO,

produces serine and threonine-linked glycoproteins. As such, these proteins have much lower homology (12%) than other related enzymes in the two pathways, which have sequence identities ranging from 21% - 52% (Table 3-2). PglB(*Cj*) is homologous to the STT3 catalytic subunit of the multimeric eukaryotic OTase, whereas PglO(*Ng*) resembles WaaL O-antigen polymerases.[1, 5, 6] Despite these differences, the predicted topology of these enzymes is remarkably similar containing 9-13 amino-terminal transmembrane domains followed by one or two soluble domains at the C-terminus (Figure 4-2). This architecture may represent a general paradigm for enzymes responsible for glycan transfer to protein from polyprenyl-linked donors.

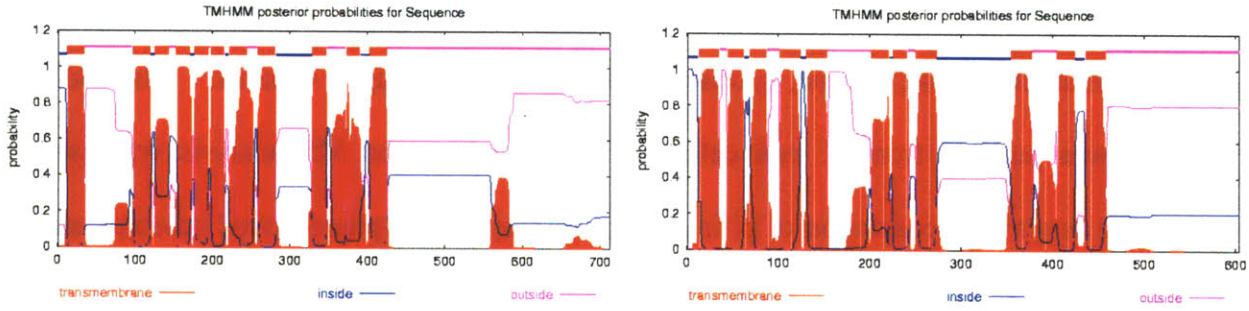


Figure 4-2. Transmembrane topology predictions for PglB(*Cj*) and PglO(*Ng*).
The TMHMM prediction program[7] was used to analyze the PglB(*Cj*) (left) and PglO(*Ng*) (right) sequences.

The bacterium *Neisseria meningitidis* contains an O-linked glycosylation pathway that is highly homologous to the pathway in *N. gonorrhoeae*. The related enzymes in the two pathways have sequence identities ranging from (85-95%, Table 3-2). In fact, the enzymes that are most similar are the OTases; the *N. meningitidis* OTase, termed PglL(*Nm*), has 95% identity with PglO(*Ng*). Recent studies have examined the glycan specificity of PglL(*Nm*) and the protein substrate specificity of PglO(*Ng*) and it would be expected that the conclusions would apply to both OTases.[2, 3]

Much recent interest has focused on the bacterial OTases since these enzymes are attractive tools for glycoengineering. Towards this end, the glycan specificity of both PglB(*Cj*) and PglL(*Nm*) has been tested *in vivo* to determine if these proteins are good candidates for glycoprotein production.[2, 3] The OTases and appropriate protein substrates were heterologously co-expressed in *E. coli* with a locus encoding for the biosynthesis of an undecaprenyl-linked glycan from multiple O-antigen biosynthetic clusters. Importantly, the O-antigen ligases were not expressed in the heterologous system, resulting in accumulation of the undecaprenyl-diphosphate glycans within the bacterial membranes. Under these conditions, both PglB(*Cj*) and PglL(*Nm*) were shown to transfer a variety of complex glycan substrates to protein.[2, 3] The O-antigen subunits utilized in these experiments comprise many types of saccharide units including Glc, Gal, Man, rhamnose, fucose, glucuronic acid and other sugar variants.[2, 3] In addition, PglL(*Nm*) was purified from *E. coli* cells and Western blot and MS analysis was used to demonstrate *in vitro* glycosylation of pilin using a farnesyl-diphosphate-linked substrate analog that contained the *E. coli* O86 O-antigen cluster.[2]

The second aspect of OTase specificity concerns the protein substrate. The N-linked OTase, PglB(*Cj*), prefers the sequon D/E-X-N-X-S/T, which has been established by a variety of *in vitro* and *in vivo* methods.[8, 9] In addition, it has been shown that this sequon can be incorporated into protein loops that are recognized and glycosylated by PglB(*Cj*).[8] In contrast, the binding determinants governing how PglO(*Ng*) interacts with protein substrates are not known. O-linked protein glycans in *N. gonorrhoeae* and *N. meningitidis* were originally discovered on Type IV pili and were thought to be a novel modification of this protein.[10, 11] Pilin glycosylation occurs at Ser 63 in a structurally defined short alpha helix (Figure 4-3). However, it was recently established that the *pgl* locus in *N. gonorrhoeae* encodes a general O-linked

protein glycosylation pathway that mediates glycosylation of many periplasmic proteins.[5] Unlike N-linked protein glycosylation, the discovery of multiple protein substrates for PglO(*Ng*) has not revealed any type of local consensus sequence. Furthermore, the glycosylation sites identified in the non-pilin protein substrates are located within regions predicted to have low structural complexity, in contrast to pilin, which is glycosylated on a defined alpha helical turn. It is unclear how PglO(*Ng*) recognizes both pilin and non-pilin substrates.

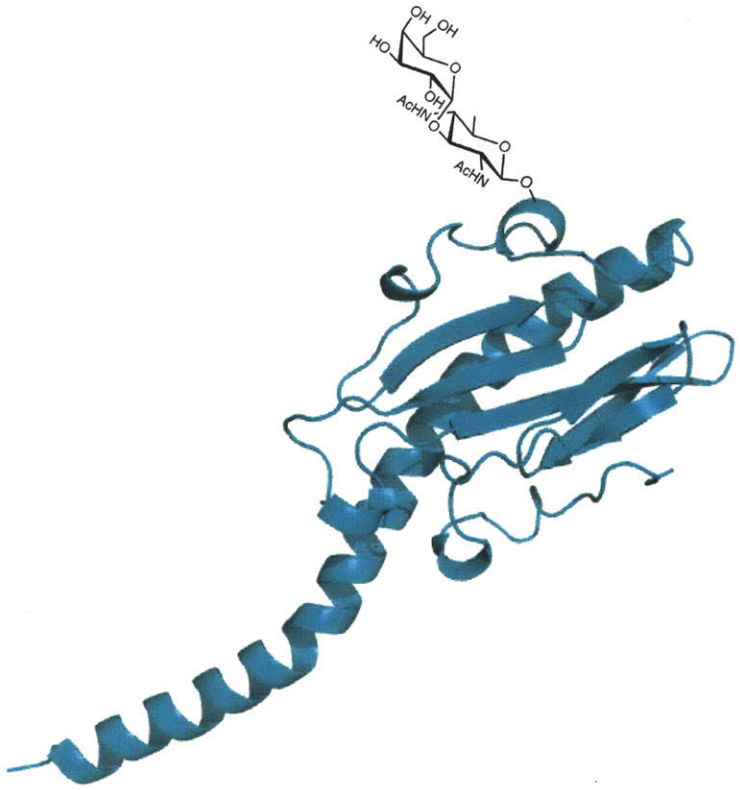


Figure 4-3. A pilin monomer from *N. gonorrhoeae* modified at Ser 63 by Gal-α-(1,3)-diNAcBac.
The original structure (PDB ID: 2PIL) was solved with a disaccharide thought to be Gal-α-(1,3)-GalNAc, although it was later shown to be the diNAcBac-containing glycan depicted above.

This chapter will address our efforts to characterize the glycan and protein specificity of the *N. gonorrhoeae* OTase, PglO(*Ng*), using *in vitro* radioactivity-based assays and Western blot analysis. In addition, this chapter describes parallel assays examining the glycan substrate

specificity of the N-linked OTase PglB(*Cj*). In general, our results show that both PglO(*Ng*) and PglB(*Cj*) are specific for glycans native to *N. gonorrhoeae* or *C. jejuni* and that PglO(*Ng*) preferentially glycosylates pilin protein.

Results and Discussion

Functional characterization of oligosaccharyltransferase, PglO

PglO and pilin are integral membrane proteins. Both proteins were heterologously expressed in *E. coli* and the target proteins were first isolated by ultracentrifugation, which afforded cell envelope fractions (CEFs). The proteins were then extracted from the CEFs with Triton X-100 and purified to homogeneity (Figure 4-4, lanes 2 and 3). To assay for OTase activity, purified PglO (1 μM) was incubated with pilin (8 μM) and the radiolabeled Und-PP-diNAcBac-[^{3}H]Gal glycosyl donor (20 nM) as described in Figure 4-1. After overnight incubation, the reaction mixture was bound to Ni-NTA resin and washed thoroughly to remove most (>99%) of the unreacted Und-PP-diNAcBac-[^{3}H]Gal donor. The pilin protein was then eluted with imidazole and the radioactivity associated with the wash and elution fractions was determined by scintillation counting (Figure 4-5). Under these assay conditions, in which pilin protein is in excess (400-fold) over Und-PP-diNAcBac-[^{3}H]Gal, PglO transferred ~60% of the sugar substrate to pilin (Figure 4-5).

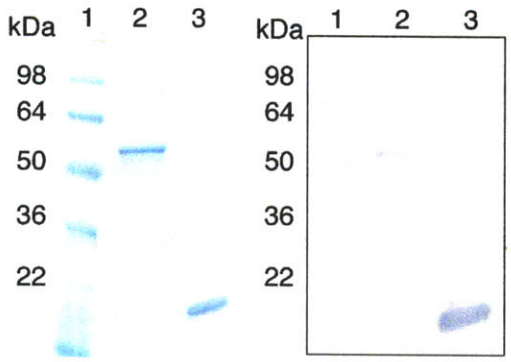


Figure 4-4. Anti-His_4 Western blot and SDS-PAGE analysis of PglO(*Ng*) and pilin.
Lane 1, molecular weight standards; lane 2, PglO(*Ng*) (67 kDa); and lane 3, pilin (18 kDa).

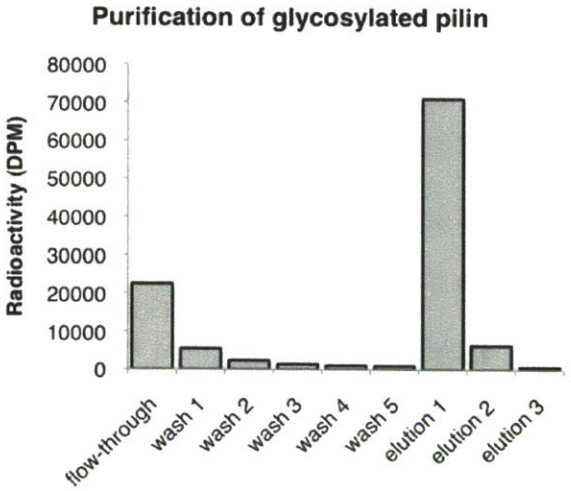


Figure 4-5. Purification of glycosylated pilin by Ni-NTA resin.
The radioactivity present in the flow-through, wash and elution steps was determined by scintillation counting. The extent of the reaction was calculated by comparing the amount of radioactivity in the elution fractions to the total radioactivity in the starting reaction. As seen above, thorough washing removes >99% of unreacted substrate. The specific activity of the Und-PP-[^{3}H]-glycans is 20 Ci/mmol, i.e., 10,000 DPM is equivalent to 4.5 pmoles of product.

These results were further verified via Western blot analysis utilizing a monoclonal antibody recognizing a diNAcBac-associated glycopeptide epitope that was developed previously.[12] For the Western blot analysis, the pilin glycosylation reaction was performed with equimolar amounts of protein substrate and Und-PP-diNAcBac donor and under these conditions ~13% of the pilin protein was associated with glycan as measured by the radioactive-transfer

method described in Figure 4-5. The diNAcBac-specific antibody showed strong staining with the glycosylated pilin, but was unreactive with the unmodified pilin (Figure 4-6). It should be noted here that the pilin protein is endogenously expressed as a pre-pilin protein, which undergoes a protease-catalyzed removal of the first seven amino acid residues to generate the functional pilin protein. When pilin was heterologously expressed in *E. coli* for the experiments described here, the proteolytic cleavage was incomplete and both the pilin and pre-pilin proteins are observable in the SDS-PAGE and Western blot analyses.

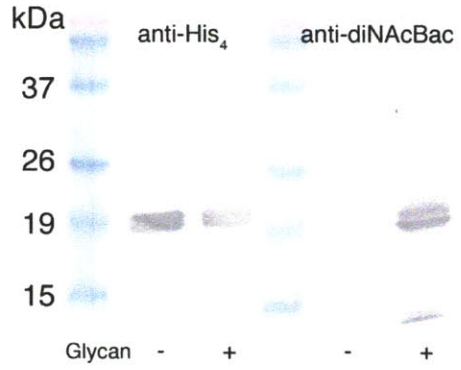


Figure 4-6. Western blot analysis of unmodified pilin and PglO-glycosylated pilin.
The doubling of the pilin band is caused by the presence of both the pre-pilin protein and the proteolytically cleaved pilin protein. *Left*, As a positive control, the His_4 antibody was used to detect the pilin protein and demonstrate that similar amounts are loaded in both lanes. *Right*, To probe for glycosylation, a diNAcBac-epitope recognizing monoclonal antibody termed npg1 was used to demonstrate that only the glycosylated pilin is reactive with the sugar-specific antibody.

The *C. jejuni* and *N. gonorrhoeae* pathways culminate in transfer of the oligosaccharide to protein; in the bacterial N-linked glycosylation pathway, PglB(*Cj*) transfers a heptasaccharide *en bloc* to the amide side chain of asparagine residues. Importantly, we demonstrate here that PglO(*Ng*) acts in a similar *en bloc* manner to transfer glycans to hydroxyl side chains of serine or threonine residues. The action of PglO(*Ng*) differs from all other known O-linked

oligosaccharyltransferases, which perform sequential addition of sugar units to proteins, and thus this enzyme represents a new functional class of enzymes.

Glycan donor specificity of PglO(*Ng*)

To further characterize PglO(*Ng*), a screen of various oligosaccharyl donors was performed. Previous studies on PglL(*Nm*),[2, 3] the homologous OTase found in *N. meningitidis*, suggested that these enzymes exhibit relaxed substrate specificity *in vivo* and can transfer oligosaccharides composed of different sugars, linkages and lengths. Thus, PglO(*Ng*) was assayed with three native substrates (Figure 4-7, the products of PglA, PglE and PglH, **2**, **3** and **4**) and with two substrates from the *C. jejuni* pathway, Und-PP-diNAcBac modified with one or two GalNAc residues (Figure 4-7, **5** and **6**). Und-PP-diNAcBac (Figure 4-7, **1**), which is common to both pathways, was also tested.

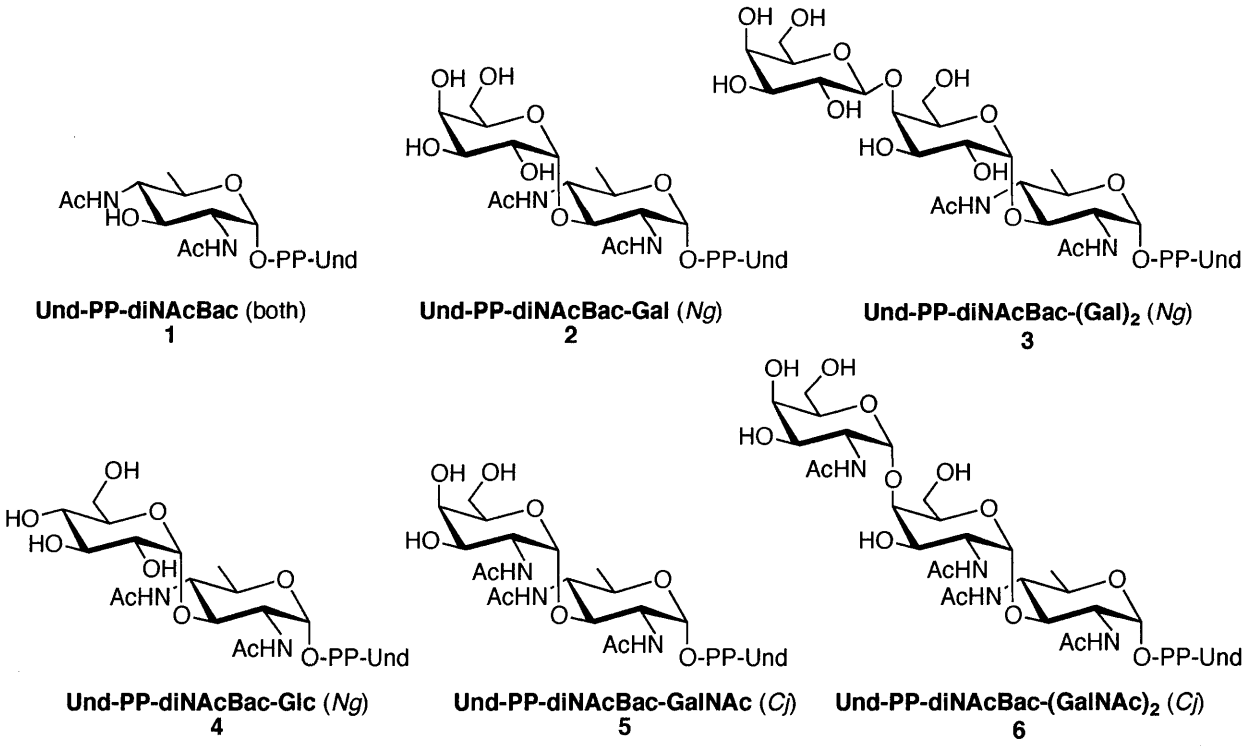


Figure 4-7. Und-PP-linked glycans assayed with PglO(*Ng*) and PglB(*Cj*).

Surprisingly, and in contrast to the *in vivo* studies with PglL(*Nm*),[2, 3] PglO(*Ng*) was only able to transfer the four native substrates under these assay conditions; the two *C. jejuni* substrates had <3% turnover (Figure 4-8). In this case, it is likely that the steric bulk of the additional C-2 acetamido groups in the second sugar of the *C. jejuni* glycan interfered with binding of the undecaprenyl-diphosphate-linked substrate. To verify that the *C. jejuni* substrates were functional as glycan donors, all six substrates were assayed with PglB(*Cj*). Pilin was used as the protein substrate in these assays as well, because it contains an N-linked glycosylation sequon (^{59}EN**N**T**S**^{63}) adjacent to the site of O-linked glycosylation.[9, 13] As seen in Figure 4-8, PglB(*Cj*) transferred the native *C. jejuni* substrates in addition to the diNAcBac-Gal (**2**) disaccharide from *N. gonorrhoeae*; it showed low reactivity with **3** and **4**. This pattern of specificity is not unexpected as PglB(*Cj*) transfers diNAcBac-GalNAc (**5**) natively and loss of the acetamido group on Gal does not impose an additional steric effect that would prevent binding. However, the *N. gonorrhoeae* trisaccharide, **3,** was a poor substrate for PglB(*Cj*), since the terminal β-(1,4) Gal linkage may interfere with binding. In addition, PglB(*Cj*) did not accept Und-PP-diNAcBac-Glc (**4**), suggesting that it is specific for the stereochemistry of the second sugar.

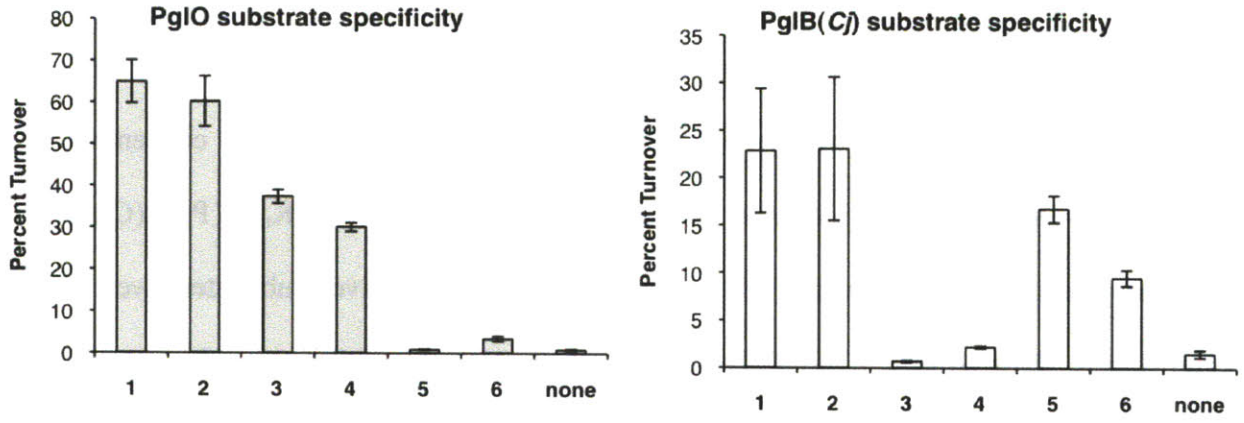


Figure 4-8. Glycan donor specificity of PglB(*Cj*) and PglO(*Ng*).
Glycan donors **1**, **2**, **3**, and **4** are native to *N. gonorrhoeae* and **1**, **5**, and **6** are native to *C. jejuni*.

Comparison of the *in vitro* glycan specificity results presented in Figure 4-8 to other studies reveals that glycan specificity may have been obscured in the in vivo studies that were performed previously. In particular, *in vivo* analyses revealed that both PglL(*Nm*), which has 95% sequence identity with PglO(*Ng*), and the N-linked OTase PglB(*Cj*) were promiscuous enzymes that transferred a variety of Und-PP-linked O-antigen substrates to serine or asparagine residues in proteins, respectively.[2, 3] In addition, previous studies showed that an *N. gonorrhoeae* strain with heterologously expressed PglA(*Cj*) contained proteins modified by the *C. jejuni* disaccharide, diNAcBac-GalNAc, implying that PglO(*Ng*) can recognize this glycan *in vivo*.[6] However, in this study, we have found that the OTases do not show a comparable substrate promiscuity *in vitro*; in fact, it appears that PglO(*Ng*) and PglB(*Cj*) are both specific for their native substrates under these assay conditions (Figure 4-8).

The previous *in vivo* studies on PglO(*Ng*), PglL(*Nm*) and PglB(*Cj*) were performed in *E. coli* in the absence of native substrate and under conditions in which the non-native undecaprenyl-linked glycan accumulated in the membrane.[2, 3] Thus, the local concentration of the substrate within the two-dimensional plane of the membrane was likely to be much higher

than in the *in vitro* assay, which would promote reaction with the OTase. Additionally, in the context of a phospholipid bilayer, the role of the membrane-bound undecaprenyl moiety may play a greater role in enzyme recognition of the substrate. In our assay, the concentration of undecaprenyl substrate (10-20 nM) was well below 2.7 μM, the apparent K_m of PglB(*Cj*) for **5**,[13] and thus specificity differences between native and non-native substrates were easily distinguished. In addition, it should be noted that the OTases exhibit substrate specificity in native cellular contexts. In the native bacteria, PglO(*Ng*), PglL(*Nm*), and PglB(*Cj*) selectively transfer the correct oligosaccharide to pilin in the presence of other undecaprenyl-linked substrates, including those involved in capsular polysaccharide biosynthesis in *N. meningitidis*[14] and *C. jejuni*[15] and the peptidoglycan component in the cellular membranes of all three organisms.[16]

Pilin glycosylation by PglB(*Cj*) and PglO(*Ng*)

Since pilin protein was utilized in the assays of both PglB(*Cj*) and PglO(*Ng*), it was important to confirm the identity of the glycosylated residues. Pilin variants were prepared with alanine mutations at the expected sites of glycosylation for PglB(*Cj*) and PglO(*Ng*), Asn 61 and Ser 63, respectively. PglB(*Cj*) was unable to glycosylate pilin-N61A, validating this residue as the N-glycan acceptor site (Figure 4-9). However, PglO showed ~85% of normal activity with the pilin S63A mutant suggesting that another site, potentially Thr 62, can be a glycosyl acceptor site in the absence of Ser 63 (Figure 4-9). Further mutational analysis confirmed this hypothesis; pilin-T62A exhibited normal glycosylation, whereas glycosylation was greatly reduced in the double alanine mutant (T62A/S63A) (Figure 4-9).

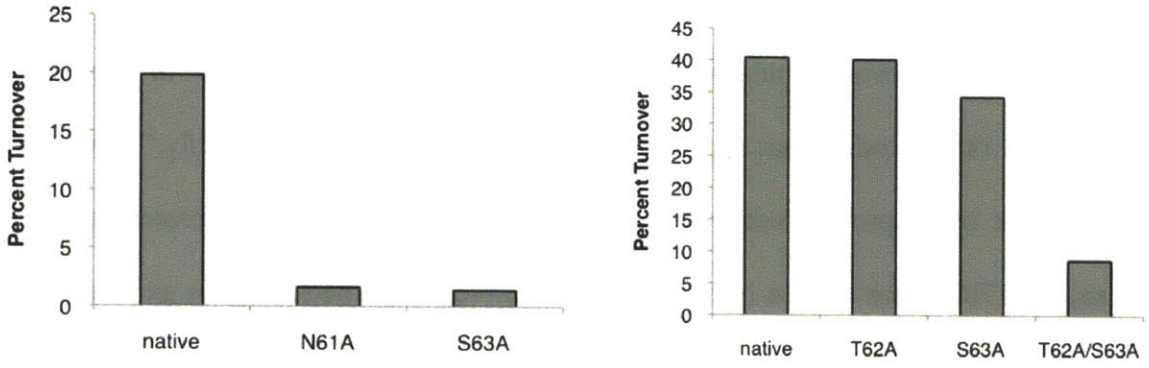


Figure 4-9. Verification of glycosylated residues in pilin protein substrate.
Left, PglB(*Cj*) is inactive when either Asn 61 or Ser 63 is mutated, which is expected as both are critical for N-linked glycosylation. *Right*, In contrast, it appears that PglO(*Ng*) can glycosylate either Ser 63 or Thr 62, but when both are mutated to alanine, glycosylation is greatly diminished.

The glycosylated residues, Asn 61 and Ser 63, are located within the α-β loop of pilin (Figure 4-3), which is surface exposed in the assembled pili structure[17] and highly modified by other post-translational modifications including phosphoethanolamine and phosphocholine.[18] It is interesting that this highly modified loop also contains a bacterial N-linked sequon, ^{59}ENNTS^{63}, that overlaps the O-linked glycosylation site.[8, 9] The presence of the N-linked sequon in the pilin structure is interesting from an evolutionary standpoint suggesting that post-translational modification of this conserved loop may be important for pilin structure and function and that in other bacteria, this loop could contain potentially be elaborated with N-linked glycans. A survey of sequenced bacterial genomes that contain a PglB(*Cj*) OTase homolog revealed at least one bacterium, *Nitrosococcus halophilus*, with a predicted pilin homolog containing the requisite N-linked glycosylation consensus sequence.

Protein substrate specificity of PglO(*Ng*)

Recent work by our collaborators has resulted in the identification of numerous non-pilin protein substrates glycosylated by the *N. gonorrhoeae* O-linked pathway; all of the identified

proteins were either periplasmic or located on the cell surface and most were predicted lipoproteins with unknown or putative functions.[5] It has been established in the previous experiments and other studies that PglO(*Ng*) glycosylates pilin protein readily,[2-5] but little is known about the other protein substrates. Two of the recently identified protein substrates, Ng1225 and Ng1769, were chosen on the basis of relatively high heterologous expression in *E. coli* (data not shown). These proteins are annotated as a peptidyl prolylisomerase and a cytochrome-c peroxidase, respectively, and are predicted lipoproteins, which were purified in the presence of detergent using Ni-NTA chromatography.

PglO(*Ng*) was assayed with pilin, Ng1225 and Ng1769 in the manner described above (Figure 4-5). An average of three experimental trials demonstrated that under these assay conditions, PglO is only able to glycosylate the pilin protein; both Ng1225 and Ng1769 showed <1% turnover after overnight reaction (Figure 4-10). These results were substantiated via Western blot analysis utilizing the diNAcBac-reactive antibody. For the Western blot analysis, the pilin glycosylation reactions were performed with equimolar amounts of protein substrate and Und-PP-diNAcBac donor. Analysis of equivalent amounts of reacted and unreacted protein substrates showed that only the glycosylated pilin showed strong staining with the anti-diNAcBac antibody (Figure 4-11).

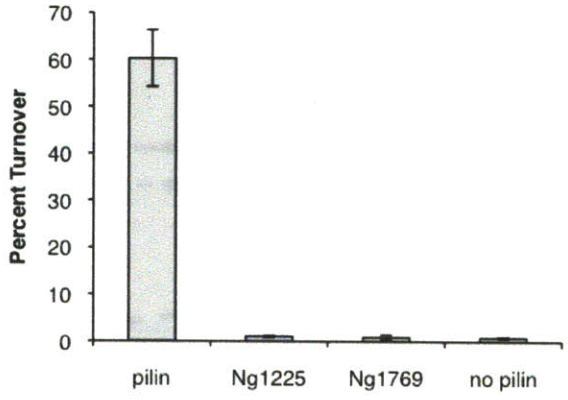


Figure 4-10. Activity of PglO(*Ng*) with pilin, Ng1225, and Ng1769.

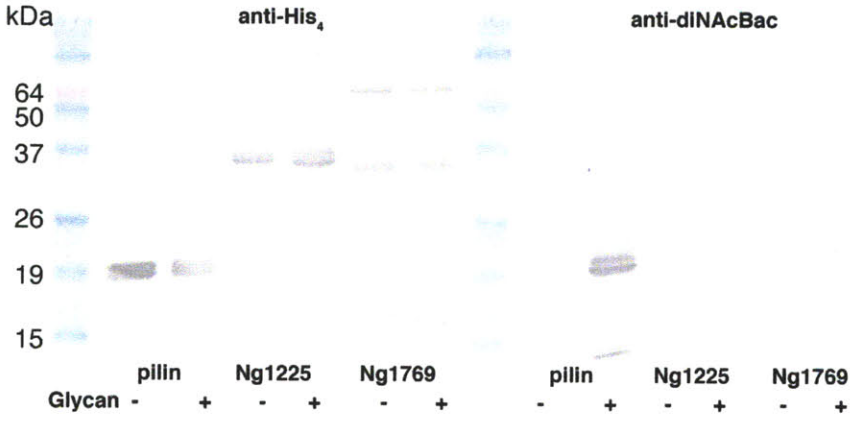


Figure 4-11. Western blot analysis of PglO(*Ng*) protein substrate specificity.
Only the glycosylated pilin is immunoreactive with the antibody recognizing the diNAcBac-epitope.[12] The molecular weight of Ng1225 is 29 kDa and the molecular weight of Ng1769 is 42 kDa.

Finally, PglO(*Ng*) activity was examined in the presence of a peptide substrate. In the N-linked glycosylation pathway, it is well established that PglB(*Cj*) glycosylates short peptides containing the established consensus sequence.[1, 8] However, when tested in a radioactivity-based glycosyltransferase activity, PglO(*Ng*) was unable to glycosylate a short peptide based on the pilin glycosylation sequence (Figure 4-12), whereas PglB(*Cj*) was active with the same peptide since it also contained a sequon for N-linked glycosylation.

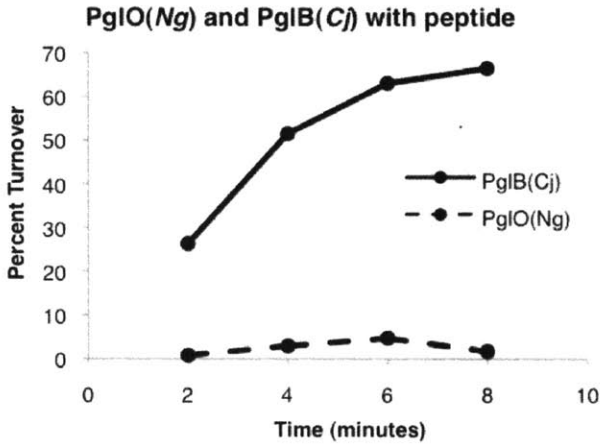


Figure 4-12. PglB(*Cj*) and PglO(*Ng*) activity with a pilin peptide substrate.
Both OTases were assayed with 200 μM peptide, KKENNTSAGNASNph, based on the pilin substrate, where Nph is 4-nitrophenylalanine, which is useful for quantification of the peptide. For PglB(*Cj*), **5** was used as the glycan donor and for PglO(*Ng*), **2** was used (Figure 4-7). The peptide assay followed the previously described protocol.[1]

PglO(*Ng*) was originally identified as the enzyme responsible for glycosylation of *N. gonorrhoeae* Type IV pili, but has since been shown to glycosylate a wide variety of periplasmic and extracellular lipoproteins. In all examples of non-pilin glycosylation substrates, the acceptor serine or threonine residues are present in loop regions predicted to have undefined structures rich in Ala, Ser, and Pro residues.[5] However, under the *in vitro* assay conditions described in this chapter, PglO(*Ng*) was only capable of glycan transfer to pilin. No activity was observed in the presence of a pilin-based peptide or the other identified protein substrates.

For the peptide substrate, we were able to verify that the assay and substrates were viable, since PglB(*Cj*) efficiently glycosylated the same peptide (Figure 4-12). However, in the case of Ng1225 and Ng 1769, it is difficult to verify if these purified lipoproteins are correctly folded or active, because the functions of these proteins have not been biochemically defined. Thus, it is difficult to interpret the results observed in Figures 4-10 and 4-11. PglO(*Ng*) inactivity could originate from improperly folded lipoprotein substrates. On the other hand, it is possible that the

reaction of PglO(*Ng*) with the lipoproteins requires another protein factor that is not present in the *in vitro* assay. Additionally, in the context of the native lipid bilayer, it is possible that the undecaprenyl moiety of the glycan donor may facilitate interaction between the lipoproteins and PglO(*Ng*) in a manner that is not recapitulated in detergent-based assays. This is an intriguing possibility and is consistent with the inability of PglO(*Ng*) to glycosylated short soluble peptides (Figure 4-12). Finally, it seems likely that PglO(*Ng*) has evolved primarily to glycosylate pilin, and as such, the pilin is a much better substrate than the other protein substrates identified, which may be reflected in our activity results. Further biochemical and structural analyses are needed to understand how PglO(*Ng*) recognizes both pilin and non-pilin protein substrates.

Conclusions

In conclusion, this chapter summarizes our study of the OTase from the O-linked glycosylation pathway in *N. gonorrhoeae*. Characterization of the OTases from *N. gonorrhoeae* and *C. jejuni* has suggested that these enzymes prefer their respective native glycans relative to closely related oligosaccharides under the described *in vitro* assay conditions. In addition, this study illuminated the interesting question of how PglO(*Ng*) recognizes its protein substrates. In the *in vitro* assay, PglO(*Ng*) glycosylated pilin protein, but not pilin-based peptides or alternative protein substrates identified *in vivo*. It is unclear how PglO(*Ng*) recognizes both pilin and non-pilin protein substrates. Further studies are required to shed light on how PglO(*Ng*) displays glycan and protein promiscuity *in vivo*, while showing more stringent specificity *in vitro*.

Acknowledgements

We are grateful to Dr. Finn Erik Aas from the laboratory of Professor Michael Koomey at the University of Oslo for preparing the original DNA constructs for PglO, pilin, Ng1225 and Ng1769 used in this study. I am also grateful to members of the Imperiali lab including Dr. Angelyn Larkin and Marcie Jaffee for sharing their OTase expertise along with the occasional glycan substrate. Finally, I would like to thank Dr. Jerry Troutman, Marcie Jaffee, Mike Morrison, and Michelle Chang for their helpful reviews and comments of this manuscript and chapter.

Experimental Procedures

Preparation of DNA constructs

The genes *pglO*, *pilE*, *Ng1225*, and *Ng1769* were PCR amplified from the *N. gonorrhoeae* strain MS11.[6, 18, 19] The genes were cloned into BamH I/Xho I in the pET-24a(+) vector (Novagen). The Xho I site was inserted prior to the stop codon to encode for a His_6 tag at the C-terminal end of each protein. The pilin alanine variants were prepared using site directed mutagenesis based on the QuickChange method (Stratagene). The double mutant was prepared using the DNA construct for the single mutant (S63A). The primers used in the PCR reactions are shown in Table 4-1. Preparation of the desired DNA constructs was verified by sequencing.

Table 4-1. Primers for generation of pilin alanine mutants. The mutated residues are italicized.

Pilin mutant	Primer
Fwd-N61A	CAAATGGCCGGAAAAC***GCC***ACTTCTGCCGGCGTG
Rev-N61A	CACGCCGGCAGAAGT***GGC***GTTTTCCGGCCATTTG
Fwd-S63A	CGGAAAACAACACT***GCG***GCCGGCGTGGCATCC
Rev-S63A	GGATGCCACGCCGGC***CGC***AGTGTTGTTTTCCG
Fwd-T62A	CCGGAAAACAAC*GCG*TCTGCCGGCGTGGC
Rev-T62A	GCCACGCCGGCAGA***CGC***GTTGTTTTCCGG
Fwd-T62A/S63A	CCGGAAAACAAC***GCGGCG***GCCGGCGTGGC
Rev-T62A/S63A	GCCACGCCGGC***CGCCGC***GTTGTTTTCCGG

Expression of proteins

PglO, PilE, Ng1225, and Ng1769 were expressed heterologously in *E. coli* BL21 cells (Agilent). A typical expression protocol involved preparation of an overnight culture of cells (5 mL), which was used to inoculate 1 L of LB media with shaking at 37 °C. After the cells reached an optical density of ~0.8 absorbance units, the temperature was lowered to 16 °C and the cells

were induced with 0.5 mM iso-β,D-thiogalactosylpyranoside (IPTG). After 16-18 hours of incubation, the cells were harvested and the pellets were stored at -80 °C.

Protein purification

In general, all steps of protein purification were carried out at 4 °C. Protein concentrations of Ng1225 and Ng1769 were determined using the appropriate extinction coefficients at a UV absorbance of 280 nm ($\varepsilon_{280,Ng1225}$ = 13370 $M^{-1}cm^{-1}$ and $\varepsilon_{280,Ng1769}$ = 37410 $M^{-1}cm^{-1}$). PglO and pilin were quantified with the MicroBCA Assay (Pierce) due to the presence of the highly absorbent detergent Triton X-100. BSA was used as the protein standard.

The cell pellets generated from the expression of the predicted lipoproteins proteins, Ng1225 and Ng1769, were resuspended in 40 mL of lysis buffer containing 50 mM Tris (pH 8.0) and 150 mM NaCl. The cells were lysed at 4 °C by two intervals of sonication for 1 minute each at an amplitude of 55% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. Following sonication, the cell lysates were incubated with 1% β-D-dodecyl maltoside (DDM) for 20 minutes. The lysates were cleared by centrifugation (145,000 x g) for 65 min. The supernatant was mixed with 1-2 mL of Ni-nitrilotriacetic acid (Ni-NTA) resin (Qiagen), tumbled for 2 hours, and then packed into a K 9/15 column (GE Healthcare). Using gravity flow, the resin-bound protein was washed with 20 column volumes of lysis buffer containing 20 mM imidazole and 0.1% DDM. The resin was further washed with 20 column volumes of lysis buffer supplemented with 45 mM imidazole and 0.05% DDM. The protein was eluted in lysis buffer supplemented with 250 mM imidazole and 1 mL fractions were collected. Fractions containing

purified material were assessed by SDS-PAGE and pooled fractions were dialyzed against lysis buffer, concentrated, and frozen at -80 °C.

To purify the membrane-associated proteins (PglO and pilin), a cell envelope fraction (CEF) was first prepared. The cells were thawed in 40 mL of lysis buffer per L of cell culture supplemented with 1 mg/mL of lysozyme. The cells were lysed at 4 °C by two intervals of sonication for 75 seconds each at an amplitude of 70% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. Cellular debris was cleared by centrifugation at 9000 x g for 45 minutes. The resulting supernatant was transferred to a clean centrifuge tube and subjected to centrifugation at 145,000 x g for 65 min to pellet the cell envelope fraction (CEF). The CEF was homogenized in 10 mL of buffer containing 1% Triton X-100 per liter of cell culture. Each CEF was incubated with detergent for several hours and then centrifuged again (145,000 x g) to remove insoluble material. The resultant supernatants were incubated with 0.5 mL of Ni-NTA resin for 1-2 hours; the resin was washed as previously described with the addition of 0.1% Triton X-100 instead of DDM to the wash and elution buffers. The proteins were eluted from the resin in 1 mL fractions. The most concentrated fractions of PglO and pilin were desalted using the Hi-Trap desalting cartridge (GE Healthcare) with lysis buffer and stored at -80 °C.

Oligosaccharyltransferase assays

The OTase reactions were performed with a variety of Und-PP-linked glycosyl donors. The Und-PP-glycan donors were prepared at a high specific activity to maximize the assay sensitivity in the manner as described in Chapter 3. In general, the reactions contained 5% DMSO, 0.7% Triton X-100, 50 mM $MnCl_2$, 25 mM HEPES (pH 7.5), 70 mM sucrose, 10-20 nM

Und-PP-substrate (20 Ci/mmol), 8 μM pilin and 1 μM PglO in 100 μL reaction volume. The reactions were incubated overnight at room temperature with shaking. The glycosylated pilin protein was isolated via Ni-NTA purification. Briefly, the reaction was incubated with 15 μL of Ni-NTA resin for several hours in a 1.5 mL eppendorf tube. The tube was briefly centrifuged and the supernatant was removed. The resin was then washed five times with 500 μL Buffer A containing 30 mM imidazole and 0.1% Triton X-100. For each wash, the buffer was added to the eppendorf tube, the resin was mixed thoroughly with buffer, and the supernatant was removed following a brief centrifugation. The protein was eluted in three fractions of 500 μL of Buffer A containing 300 mM imidazole and 0.1% Triton X-100. Scintillation fluid (Ecolite(+)™, MP Biomedicals) was added to all flow-through, wash and elution fractions and the radioactivity of each sample was determined.

Glycosylated protein samples for Western blot analysis were prepared and purified in the same manner, except that unlabeled versions of Und-PP-diNAcBac was used at concentrations of 10-24 μM. Parallel reactions with radioactive substrates were performed at identical concentrations in order to determine reaction yields. The Western blot analysis was performed following standard protocols. Following transfer of the protein bands to nitrocellulose membrane, the membrane was blocked for one hour with 3% BSA solution in TBS-T buffer. TBS-T contains 0.05 % Tween in TBS buffer (pH 7.4). The membrane was then incubated with an antibody specific for His_4 (Qiagen) at the recommended concentration. After washing the membrane three times for 5 to 15 minutes with TBS-T, the membrane was incubated with a goat anti-mouse alkaline phosphatase secondary antibody (Pierce) at a 5000X dilution. Finally, the membrane was washed twice with TBS-T and once with TBS for 10 minutes each. The Western blot was developed by incubating the membrane with 1-Step NBT/BCIP Solution (Pierce) for

five minutes or until the bands appeared. A diNAcBac-epitope monoclonal antibody termed npg1, which was previously described,[20] was used to detect diNAcBac-modified protein following a similar Western blot protocol with the exception that an anti-rabbit alkaline phosphatase secondary antibody was used.

References

1. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* **2005,** 12, (12), 1311-5.
2. Faridmoayer, A.; Fentabil, M. A.; Haurat, M. F.; Yi, W.; Woodward, R.; Wang, P. G.; Feldman, M. F., Extreme substrate promiscuity of the *Neisseria* oligosaccharyl transferase involved in protein O-glycosylation. *J. Biol. Chem.* **2008,** 283, (50), 34596-604.
3. Faridmoayer, A.; Fentabil, M. A.; Mills, D. C.; Klassen, J. S.; Feldman, M. F., Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J. Bacteriol.* **2007,** 189, (22), 8088-98.
4. Power, P. M.; Seib, K. L.; Jennings, M. P., Pilin glycosylation in *Neisseria meningitidis* occurs by a similar pathway to wzy-dependent O-antigen biosynthesis in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **2006,** 347, (4), 904-8.
5. Vik, A.; Aas, F. E.; Anonsen, J. H.; Bilsborough, S.; Schneider, A.; Egge-Jacobsen, W.; Koomey, M., Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. U S A* **2009,** 106, (11), 4447-52.
6. Aas, F. E.; Vik, A.; Vedde, J.; Koomey, M.; Egge-Jacobsen, W., *Neisseria gonorrhoeae* O-linked pilin glycosylation: functional analyses define both the biosynthetic pathway and glycan structure. *Mol. Microbiol.* **2007,** 65, (3), 607-24.
7. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
8. Chen, M. M.; Glover, K. J.; Imperiali, B., From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* **2007,** 46, (18), 5579-85.
9. Kowarik, M.; Young, N. M.; Numao, S.; Schulz, B. L.; Hug, I.; Callewaert, N.; Mills, D. C.; Watson, D. C.; Hernandez, M.; Kelly, J. F.; Wacker, M.; Aebi, M., Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **2006,** 25, (9), 1957-66.
10. Stimson, E.; Virji, M.; Makepeace, K.; Dell, A.; Morris, H. R.; Payne, G.; Saunders, J. R.; Jennings, M. P.; Barker, S.; Panico, M.; et al., Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.* **1995,** 17, (6), 1201-14.
11. Parge, H. E.; Forest, K. T.; Hickey, M. J.; Christensen, D. A.; Getzoff, E. D.; Tainer, J. A., Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature* **1995,** 378, (6552), 32-8.
12. Børud, B.; Aas, F. E.; Vik, A.; Winther-Larsen, H. C.; 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), 2816-29.
13. Chen, M. M.; Weerapana, E.; Ciepichal, E.; Stupak, J.; Reid, C. W.; Swiezewska, E.; Imperiali, B., Polyisoprenol specificity in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2007,** 46, (50), 14342-8.
14. Masson, L.; Holbein, B. E., Role of lipid intermediate(s) in the synthesis of serogroup B *Neisseria meningitidis* capsular polysaccharide. *J. Bacteriol.* **1985,** 161, (3), 861-7.

15. St Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A., The structures of the lipooligosaccharide and capsule polysaccharide of Campylobacter jejuni genome sequenced strain NCTC 11168. *Eur J Biochem* **2002,** 269, (21), 5119-36.
16. Hill, D. J.; Griffiths, N. J.; Borodina, E.; Virji, M., Cellular and molecular biology of *Neisseria meningitidis* colonization and invasive disease. *Clin. Sci. (Lond.)* **2010,** 118, (9), 547-64.
17. Craig, L.; Volkmann, N.; Arvai, A. S.; Pique, M. E.; Yeager, M.; Egelman, E. H.; Tainer, J. A., Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol. Cell.* **2006,** 23, (5), 651-62.
18. Hegge, F. T.; Hitchen, P. G.; Aas, F. E.; Kristiansen, H.; Lovold, C.; Egge-Jacobsen, W.; Panico, M.; Leong, W. Y.; Bull, V.; Virji, M.; Morris, H. R.; Dell, A.; Koomey, M., Unique modifications with phosphocholine and phosphoethanolamine define alternate antigenic forms of *Neisseria gonorrhoeae* type IV pili. *Proc. Natl. Acad. Sci. U S A* **2004,** 101, (29), 10798-803.
19. Banerjee, A.; Wang, R.; Supernavage, S. L.; Ghosh, S. K.; Parker, J.; Ganesh, N. F.; Wang, P. G.; Gulati, S.; Rice, P. A., Implications of phase variation of a gene (*pgtA*) encoding a pilin galactosyl transferase in gonococcal pathogenesis. *J. Exp. Med.* **2002,** 196, (2), 147-62.
20. Børud, B., Viburiene, R., Hartley, M. D., Paulsen, B. S., Egge-Jacobsen, W., Imperiali, B., and Koomey, M., A novel glucosyltransferase, PglH, further expands the high glycan diversity in O-linked protein glycosylation of *Neisseria* species. **2011**, submitted.

Chapter 5: Adaptation of Nanodisc technology to the study of the N-linked protein glycosylation pathway in *Campylobacter jejuni*

Introduction

N-linked glycosylation is a critical protein modification that occurs in eukaryotes, archaea, and prokaryotes. In eukaryotic systems, N-linked glycans are important for protein folding and function, intracellular communication and immunological responses,[1] whereas in prokaryotes, glycosylation of extracellular proteins is thought to contribute to bacterial pathogenicity.[2, 3] The N-linked protein glycosylation pathway is conserved in all three kingdoms of life and comprises a series of membrane-associated and integral-membrane glycosyltransferases acting sequentially on a membrane-bound polyprenol to form a polyprenyl-diphosphate glycan.[4] The oligosaccharide is then transferred to the nitrogen of an asparagine side chain amide on an acceptor protein by the action of the oligosaccharyl transferase (OTase). In yeast, the eukaryotic model for N-linked protein glycosylation, the multimeric OT is composed of eight integral membrane proteins and has been mostly intractable to biochemical characterization.[5, 6] The bacterial and archaeal OTases are relatively simpler proteins comprising a single subunit with 9-13 predicted transmembrane regions.[4] Since all of the transformations involved in N-linked glycosylation occur at the membrane interface, synergistic biophysical and biochemical techniques are needed to shed light on the surface interactions occurring *in vivo*.

The complexity of the proteins in the eukaryotic pathway has presented a considerable biochemical challenge, and the recent discovery of N-linked protein glycosylation (Pgl) in the Gram-negative bacterium *Campylobacter jejuni*[7, 8] has resulted in complete biochemical characterization of the pathway enzymes.[9-11] The *C. jejuni* pathway is initiated with the biosynthesis of uridine diphosphate-N,N'-diacetylbacillosamine (UDP-diNAcBac) by PglF, PglE, and PglD.[12] Undecaprenyl-phosphate (Und-P) acts as a membrane anchor for glycan biosynthesis and PglC is a phospho-glycosyltransferase, which has been shown to transfer

phospho-diNAcBac to form undecaprenyl-diphosphate-diNAcBac (Und-PP-diNAcBac).[11] Four subsequent glycosyltransferases (PglA, PglJ, PglH and PglI) catalyze the assembly of Und-PP-heptasaccharide (Figure 5-1).[9] A putative flippase (PglK) then translocates the heptasaccharide to the periplasmic face of the inner membrane,[13] where the OTase (PglB) transfers the glycan onto an asparagine residue of a periplasmic protein.[10, 14, 15]

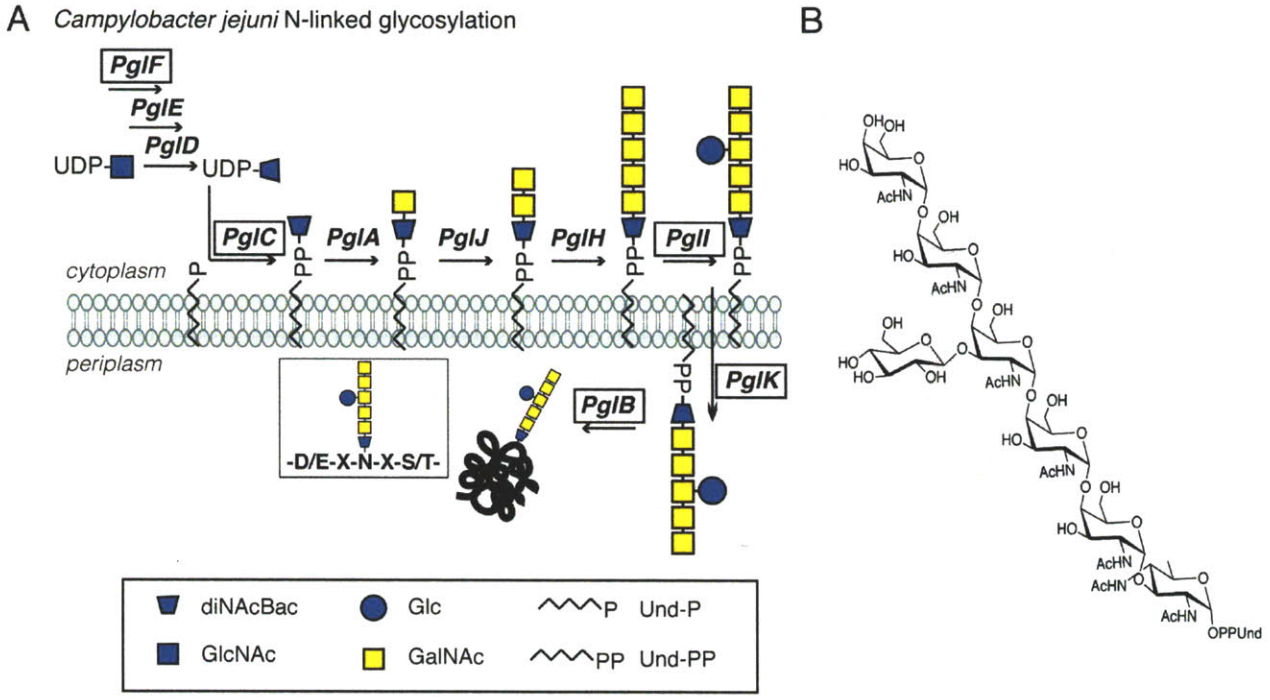


Figure 5-1. N-linked protein glycosylation in *C. jejuni.*
A, Schematic of the *C. jejuni* N-linked protein glycosylation pathway. Boxed enzymes indicate the predicted presence of integral transmembrane helices. *B*, Chemical structure of the Und-PP-heptasaccharide produced by the cytoplasmic enzymes.

Previously, the biochemical functions of the *C. jejuni* enzymes involved in N-linked glycosylation were demonstrated using detergent micelle-based *in vitro* assays.[9-11] Five of the enzymes involved in N-linked glycosylation (PglF, PglC, PglI, PglK and PglB) are predicted to be integral membrane proteins by the transmembrane prediction program TMHMM[16] and several

of the other proteins (PglA, PglJ and PglH) are predicted to have hydrophobic patches. Detergents, like Triton X-100 or β-D-dodecylmaltoside (DDM), solubilize the polyprenyl-linked substrates and hydrophobic enzymes by forming micelles, in which the detergent molecules bind to exposed hydrophobic surfaces. In the absence of detergent, the polyprenyl substrates and many of the pathway enzymes, even those that do not have predicted transmembrane domains, form insoluble aggregates. Detergent micelles have many advantages and provide a reliable, reproducible method of achieving solubility to assess biochemical function. In addition, the facile mixing of detergent micelles allows the enzymes to encounter all substrates similarly to soluble enzymes in aqueous solutions.[17] However, detergent micelles do not resemble the defined hydrophobic and hydrophilic regions of native lipid bilayers (Figure 5-2) and thus, the results of studies performed in detergent micelles may not accurately reflect enzymatic action *in vivo*. Native-like membrane systems such as liposomes[18] and bicelles[19] are frequently used to study membrane proteins, but these systems also have some disadvantages, including poor reproducibility and non-homogeneous preparations.[20] In addition, small liposome vesicles have an additional disadvantage in that vesicle curvature can negatively impact enzyme activity and related studies.[20]

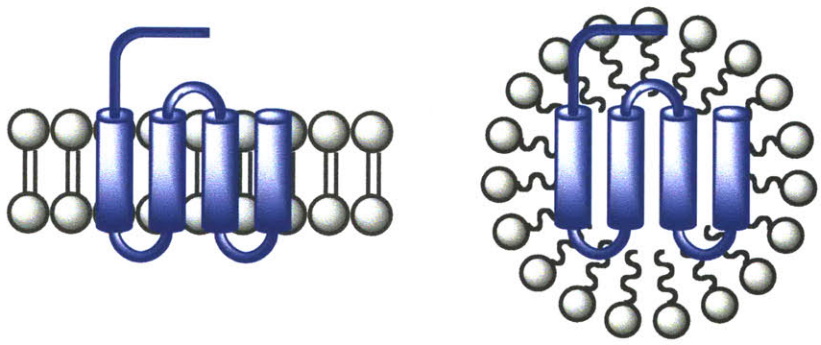


Figure 5-2. Generic integral membrane protein in lipid bilayer or detergent micelle.
Comparison of the environments experienced by a generic integral membrane protein (blue) in a lipid bilayer (left) or solubilized in a detergent micelle (right).

To address the role of the membrane interface in N-linked protein glycosylation, we have turned to Nanodiscs, a new membrane-mimetic system introduced by the Sligar laboratory.[21] To develop this technology, Sligar and coworkers exploited the unique features of apolipoprotein A1 (ApoA-1), which is involved in solubilizing cholesterol for transport in the blood.[22, 23] It was known that ApoA-1 forms a discoidal shape when transporting phospholipids and alters to a more globular form when cholesterol is present.[21] In addition, the X-ray crystal structure of an ApoA-1 truncation formed a large loop (Figure 5-3) and was quite dissimilar to the compact helical bundle formed by the full-length protein (Figure 5-3). These unique features were utilized to adapt the truncated ApoA-1 protein to form a new type of membrane mimetic, termed a Nanodisc, in which ApoA-1 encircles and solubilizes the hydrophobic edge of a flat bilayer of lipids (Figure 5-3).[23] An important innovation of Nanodiscs is that the length of ApoA-1 ensures that a given population of Nanodiscs will be relatively monodispersed. Further engineering of ApoA-1 by insertion of additional helical domains resulted in the generation of Nanodiscs with diameters ranging from 7 nm to 12 nm.[22] More recent work by Sligar and others has developed Nanodiscs with diameters of 16 nm,[24] 20 nm, and 25 nm,[25] but these larger discs have increased population size heterogeneity.



Figure 5-3. Representations of full-length ApoA-1, truncated ApoA-1 and a Nanodisc. *Left*, X-ray structure of full-length ApoA-I (PDB: 2A01). *Middle*, X-ray structure of truncated ApoA-1 (PDB: 1AV1). *Right*, A modeled representation of Nanodiscs from a side view adapted from Nath, et al.[21] shows truncated ApoA-1 forming a discoidal particle that contains phospholipids. The two blue loops represent two copies of the ApoA-1 protein and the space-filling model in gray represents the lipid bilayer.

Nanodiscs are extremely useful tools for the study of integral membrane proteins. Several types of membrane proteins have been incorporated into Nanodiscs including the G-protein coupled receptor rhodopsin,[26] bacterial chemoreceptors,[27] and cytochrome P450s.[28-30] These studies have revealed important nuances of protein oligomeric state as well as substrate binding and activity. Nanodiscs are ideal systems for characterizing enzymatic function of integral membrane proteins since Nanodisc preparations are highly reproducible relative to other membrane mimics. Nanodiscs comprise homogeneous solutions with membrane discs of the same size with the desired copy number of the membrane protein, and in addition, they provide a reliable, native-like system for investigating the physical role of the membrane in protein function.

The application of Nanodiscs to N-linked protein glycosylation constitutes a powerful approach to addressing the role of the membrane in Und-PP-heptasaccharide biosynthesis (Figure 5-1). It is intriguing that some, but not all of the enzymes in the *C. jejuni* N-linked glycosylation pathway contain predicted transmembrane domains. Of the eight cytoplasmic

enzymes that synthesize the Und-PP-glycan, only PglF, PglC, and PglI are integral membrane proteins. It is our hypothesis that these three enzymes may serve as anchors to facilitate the formation of a macromolecular enzyme complex responsible for efficient glycan biosynthesis (Figure 5-4).

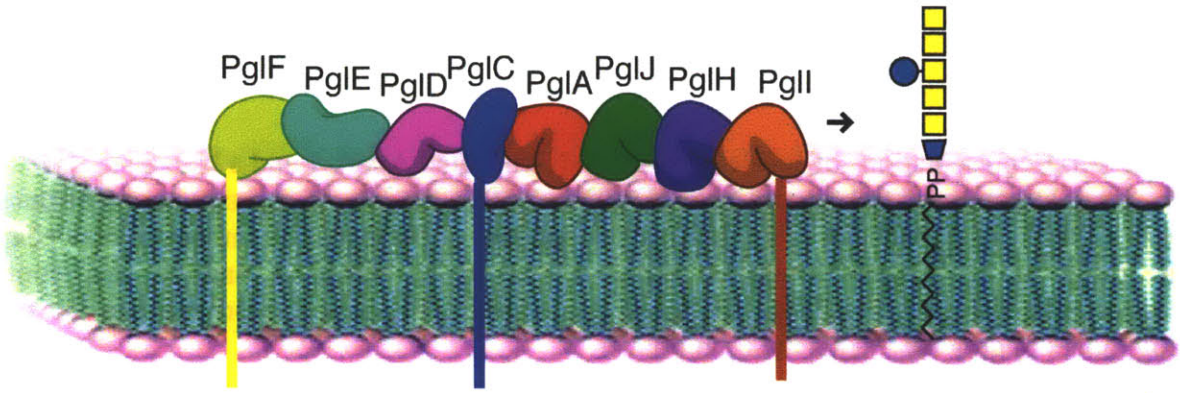


Figure 5-4. A model of Und-PP-heptasaccharide biosynthesis at the membrane interface. In this model, the eight cytoplasmic enzymes in the *C. jejuni* N-linked protein glycosylation pathway form a complex and act in a sequential fashion to produce the Und-PP-glycan product.

Several lines of reasoning support the model shown in Figure 5-4. First, the three integral membrane-bound proteins are responsible for key transition points in the pathway. PglF initiates UDP-diNAcBac biosynthesis, PglC facilitates the transition from soluble UDP-diNAcBac to membrane-bound Und-PP-diNAcBac, and PglI completes Und-PP-heptasaccharide biosynthesis and presumably presents the completed glycan to the flippase, PglK, for translocation across the lipid bilayer.[9, 11, 12] In particular, the membrane dependency of PglF lends credence to the model, since this enzyme acts on a soluble substrate, UDP-GlcNAc and produces a soluble UDP-sugar. Thus, the membrane-spanning domains may act to anchor the sugar biosynthetic reactions in the membrane and facilitate interactions with the membrane-bound undecaprenyl carrier. Secondly, the *Neisseria gonorrhoeae* O-linked glycosylation pathway studied in Chapters 3 and 4 has functional homology to the *C. jejuni* pathway and contains a bifunctional enzyme that carries out

the same reactions as PglD and PglC. This bifunctional enzyme effectively bridges the transition from the aqueous UDP-sugar biosynthesis to the membrane-bound undecaprenyl-dependent portion of the pathway, suggesting that these two enzymes may preferentially interact in the *C. jejuni* pathway. Finally, from the standpoint of pathway efficiency, a macromolecular enzyme complex allows the enzymes to act in a sequential fashion to produce Und-PP-heptasaccharide without dissociation of the intermediates. This would also prevent truncated glycans from encountering the flippase and undergoing premature translocation.

In order to investigate the validity of our proposed model of N-linked glycosylation, we have elected to use Nanodiscs as a platform for investigations of the interactions amongst the pathway enzymes, polyprenyl substrates and lipid bilayer surface. By analyzing enzyme interactions in the presence of a lipid bilayer mimic, we are pushing closer to a more accurate *in vitro* system that will illuminate the *in vivo* process. In this chapter, the first two enzymes in the pathway, PglC and PglA, are purified to homogeneity and incorporated into the Nanodisc platform. In addition, pull-down assays suggest that PglA preferentially interacts with PglC-containing Nanodiscs. This work provides the foundation for the studies described in Chapter 6 and 7, which focus on developing methods to study protein-protein interactions and enzyme activity in the context of Nanodiscs.

Results and Discussion

Purification of PglC

PglC is a phospho-glycosyltransferase containing one predicted transmembrane domain at the amino-terminus that is responsible for the transfer of phospho-diNAcBac to undecaprenyl

phosphate. PglC was previously expressed, purified, and annotated biochemically,[11] but the original construct (pET-24) and purification protocol yielded low amounts of aggregated protein (0.1-0.2 mg/L of cell culture). Insertion of target proteins into Nanodiscs requires monodispersed protein and minimum concentrations of 10-40 μM. Western Blot analysis of PglC purification using the original protocol suggested several areas for improvement (Figure 5-5). While overall expression levels were adequate, a great deal of protein was lost in the low and high-speed centrifugal spins required for membrane protein purification (Figure 5-5, lanes 2 and 4). In addition, most of the detergent solubilized PglC did not bind to Ni-NTA resin (Figure 5-5, lanes 5 and 7). Both of these problems may stem from protein aggregation, and thus steps were taken to improve the solubility of the protein.

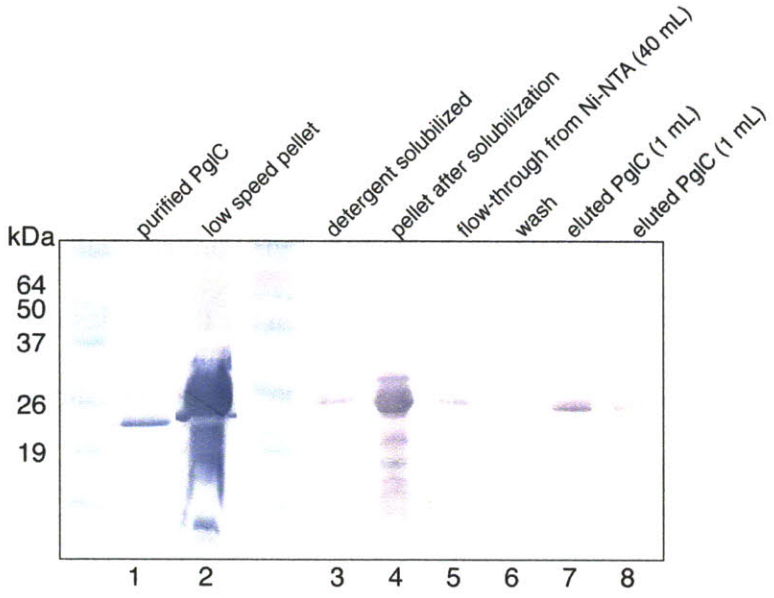


Figure 5-5. Western blot analysis of original PglC purification attempts.
For the purification steps (flow-through and elutions), similar amounts were loaded, but the sample volumes were different and these are indicated on the Western blot above.

A variety of PglC constructs were prepared with different solubility and purification tags at the N- and C-termini (Table 5-1). GST (glutathione S-transferase) and MBP (maltose binding protein) fusions have the benefit of providing an alternate method of purification. Test

expressions of all four proteins showed similar expression levels (Figure 5-6), but further purification of the GST and MBP constructs did not result in improved purification yields. In addition, the GB1 (B1 immunoglobulin) fusion was promising in terms of PglC expression levels, but the original placement of the His_6 tag at the C-terminus was troublesome since the C-terminal His_6 tag was not accessible to the Ni-NTA resin binding (Figure 5-5).

Table 5-1. PglC constructs screened for improved solubility characteristics.

Construct	N-terminus	C-terminus	Molecular Weight (kDa)
pET-24	T7	His_6	24
pET-NO	His_8-TEV	-	24
pGEX	GST	-	49
pGBH	GB1	His_6	30
pMAL	MBP	-	66
pMAL	MBP-TEV		67
pGBH-MH	His_6-GB1	-	30
pGBH-MH	FLAG-GB1	-	30

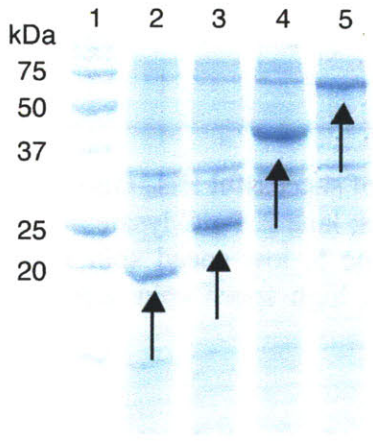


Figure 5-6. SDS-PAGE analysis of PglC constructs.
Lane 1, molecular weight standards; lane 2, T7-PglC-His_6 (24 kDa); lane 3, GB1-PglC-His_6 (30 kDa); lane 4, GST-PglC (49 kDa); lane 5, MBP-PglC (66 kDa). Arrows indicate the band representing overexpressed PglC.

To avoid these problems, a new construct (pGBH-MH) was engineered in which the His_6 tag was inserted prior to the GB1 domain and a stop codon was encoded at the C-terminus to

remove the C-terminal His_6 tag. The GB1 domain has been effective as a solubility tag[31] and was expected to display the His_6 tag in an exposed, accessible location. Subsequent purification of the pGBH-MH construct showed improved binding to Ni-NTA resin and demonstrated a significant increase of the yield of detergent purified PglC (Figure 5-7, lanes 8 and 9). In addition, supplementation of the lysis buffer with lysozyme prior to sonication improved lysis of the bacterial cell wall and increased protein yield from the first centrifugal spin (Figure 5-7, lanes 3 and 5). With these improved purification conditions, approximately 1-2 mg of PglC can be obtained from a liter of cell culture, which represents a 10-fold improvement over previous attempts.

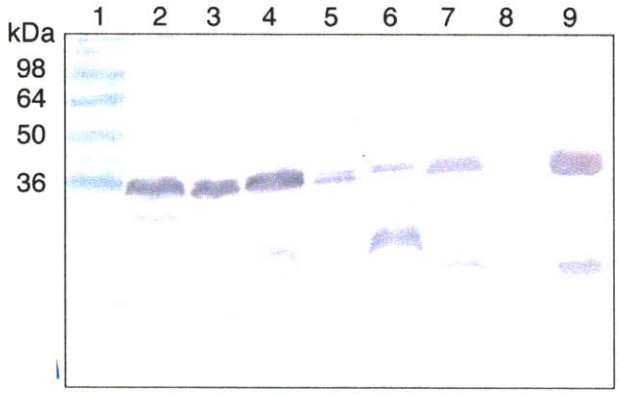


Figure 5-7. Western blot analysis of successful His_6-GB1-PglC purification.
Lane 1, molecular weight standards; lane 2, PglC lysate; lane 3, low speed spin supernatant with lysozyme; lane 4, PglC lysate; lane 5, low speed spin supernatant without lysozyme; lane 6, detergent solubilized CEF; lane 7, high speed spin supernatant; lane 8, flow-through; lane 9, purified PglC.

With the optimized PglC purification in hand, we were able to easily generate the amount of protein required for Nanodisc studies. In addition, milligram quantities of PglC made this enzyme an attractive candidate for X-ray crystallography studies, which were initiated in collaboration with Professor Karen Allen at Boston University. PglC is relatively small (23 kDa) and has only one predicted transmembrane helix (Figure 5-8, left); these features are unique

amongst most eukaryotic and prokaryotic polyprenyl phosphate:HexNAc 1-phosphate transferases, which are typically larger and contain 10 transmembrane domains with important sequence conservation in the cytoplasmic loops (Figure 5-8, right).[32, 33] Although structurally not representative of these functional homologs, PglC is an attractive model protein for studying phospho-glycosyltransferase reactions, since it is small, less hydrophobic and contains the minimal architecture needed to facilitate the transferase reaction.

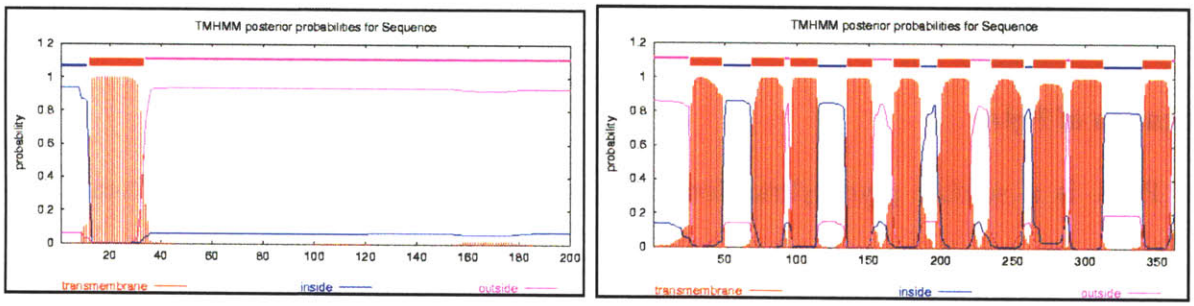


Figure 5-8. Transmembrane topology prediction of PglC and MraY (*E. coli*).
Left, the TMHMM[16] server was used to determine the predicted transmembrane domains of PglC, which has one N-terminal membrane-spanning helix. *Right*, MraY is shown, which has 10 transmembrane domains.

Unfortunately, the PglC purification protocol developed above was not ideal for X-ray crystallography, because it relied on Triton X-100 as a detergent. The heterogeneity of this detergent makes it a poor choice for obtaining well-behaved protein crystals, and a variety of different detergents and purification conditions were screened to identify conditions compatible with crystallography. In addition, X-ray crystallography requires relatively low amounts of detergent at concentrations only two to three fold greater than the critical micelle concentration.[17, 34, 35] Dodecyl maltoside (DDM) has been shown to be an ideal detergent for membrane protein crystallography, but early studies suggested that PglC was not stable at high protein concentrations (2-5 mg/ml) in buffer containing low amounts of DDM. Purification was

carried out under a variety of conditions and these details are outlined in the Appendix, but no detergent was identified that could sufficiently solubilize PglC. Important follow-up work by Vinita Lukose suggested that the solubility of PglC was maintained when PglC was co-eluted with PglD, which is the soluble acetyltransferase proceeding it in the *C. jejuni* Pgl pathway. Unfortunately, co-crystallization screens of these two proteins have not yet produced any structural data, but current efforts are focused on exploring the soluble domain of PglC as a structural target and establishing a physical interaction between PglD and PglC.

Purification of PglA

PglA is a soluble glycosyltransferase that catalyzes transfer of GalNAc from UDP-GalNAc to form an α-(1,3) glycosidic linkage to Und-PP-diNAcBac. PglA was biochemically characterized in previous studies.[9] However, size exclusion chromatography (SEC) analysis of this protein purified under the original conditions revealed that the protein formed an aggregate that eluted in the void volume, even though it is not predicted to contain transmembrane domains. The aggregation problem was alleviated by the inclusion of low amounts of detergent to the lysis and purification buffers, and SEC verified PglA monodispersity (Figure 5-9). β-D-dodecylmaltoside (DDM) was chosen as the detergent, because it successfully maintains PglA solubility and does not interfere with UV-based quantification. Previous studies of PglA utilized the pET-24 construct, but it was useful for the studies described here to prepare PglA in the pET-NO vector (Table 5-1). Both constructs rely on Ni-NTA affinity purification, but pET-NO also encodes a proteolytic site for the Tobacco Etch Virus (TEV) protease making it possible to cleave the His_6 tag after purification and generate an almost native protein with only two extra residues ("GS") on the N-terminus.

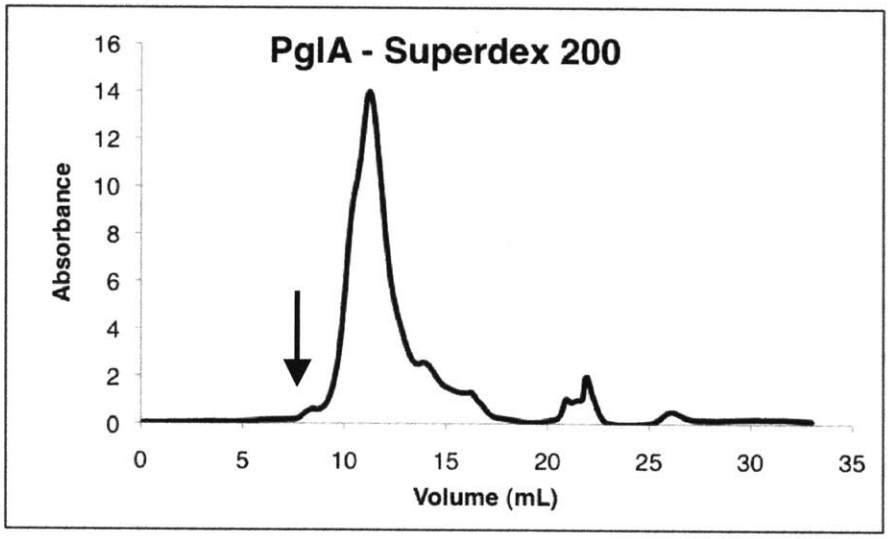


Figure 5-9. SEC analysis of detergent-solubilized PglA sample.
The arrow indicates the elution volume of the void. SDS-PAGE analysis was used to confirm that the dominant peak correlates with PglA.

Incorporation of PglC into Nanodiscs

The integral membrane protein, PglC, was incorporated into Nanodiscs following procedures established by the laboratory of Professor Stephen Sligar at the University of Illinois.[36] Each Nanodisc contains two copies of the membrane scaffold protein (MSP), which encircle the lipid bilayer (Figure 5-3). The original MSP sequences were extended with the addition of extra alpha helices resulting in discs with diameters of 11 nm, 12 nm, 13 nm and 16 nm, although the DNA construct for the largest disc is not yet available (Table 5-2).[22, 24] Initially, the MSPD1E2 construct (11 nm diameter) was utilized, because the His_6-cleaved version of the largest MSP (MSPD1E3, 12 nm) migrated identically with PglC on SDS-PAGE (12% acrylamide), which complicated analysis of Nanodisc formation. The MSP constructs used in the following experiments were either provided by the Sligar laboratory or expressed and purified in

the lab following established protocols.[37] For most of the following experiments, the TEV-cleaved version of MSP was used, in which the His_6 tag has been proteolytically removed.

Table 5-2. Membrane scaffold protein (MSP) constructs and Nanodisc sizes.

MSP construct	Lipid bilayer diameter	Nanodisc diameter	Ref
MSP1D1E1	8.5 nm	10.6 nm	22
MSP1D1E2	9.8 nm	11.9 nm	22
MSP1D1E3	10.8 nm	12.9 nm	22
MSP2N2/MSP2N3	14-15 nm	16-17 nm	24

A variety of lipids have been utilized for Nanodisc preparation including variants of phosphatidylcholine (PC) and phosphatidylethanolamine (PE) as well as native *E. coli* lipid extract. We chose to use *E. coli* lipid extract, which has a ratio of 67:23:9 PE:PC:cardiolipin, as preparation conditions for this lipid mixture were already established.[36] In addition, *E. coli* native phospholipid composition is expected to be similar to *C. jejuni*, since they are both Gram-negative bacteria. Thus, the *E. coli* lipid extract would optimally mimic the native membrane environment.

The first step of Nanodisc preparation is incubation of the protein components (MSP and detergent-solubilized PglC) with the detergent solubilized lipids at precise ratios (Figure 5-10). Subsequent detergent removal by dialysis or Biobeads initiates spontaneous Nanodisc formation, in which the MSP effectively encapsulates membrane discs with or without PglC (Figure 5-10). Dialysis-based methods are only effective with some detergents that have low critical micellar concentrations (CMCs) such as sodium cholate. Since PglC and PglA are solubilized with Triton X-100 and DDM, which do not dialyze easily, it is necessary to use Biobeads to initiate Nanodisc formation. Biobeads are small polystyrene bead that selectively bind and remove hydrophobic small molecules, like detergents, from solution. Exact protein:lipid ratios are

essential for optimal Nanodisc formation and for incorporation of the desired number of copies of PglC. Initial lipid amounts were based on previous work that established the optimal *E. coli* lipid:MSP ratios as 105:1 lipid:MSPD1E2 and 120:1 lipid:MSPD1E3.[36] Preliminary screens in which the lipid amounts were slightly varied (± 10-20%) did not show significant enhancement of Nanodisc formation.

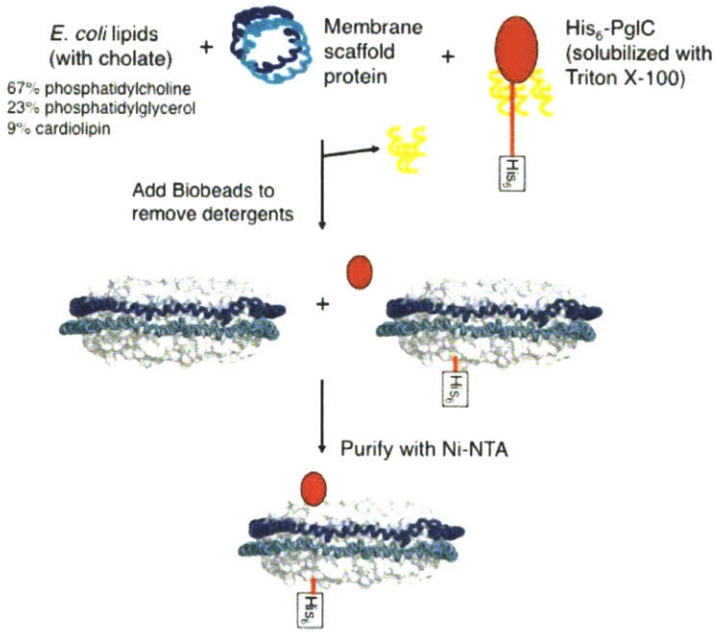


Figure 5-10. Schematic of Nanodisc preparation.
After pre-incubation of the detergent solubilized *E. coli* lipids, MSP and PglC, the detergent is removed by Biobeads and Nanodiscs spontaneously form. Subsequent purification of PglC-containing Nanodiscs was performed by Ni-NTA affinity chromatography.

Nanodisc preparations containing one copy of PglC per disc are preferred, because this simplifies downstream studies on protein-protein interactions in the *C. jejuni* pathway. Using the smaller MSPD1E2 construct (12 nm), Nanodiscs were prepared with MSP:PglC ratios of 2:1, 5:1 and 10:1. After Biobead-mediated detergent removal, the PglC-containing discs were purified from the empty discs using Ni-NTA affinity chromatography (Figure 5-11). Gel densitometry was performed to quantify the amount of PglC and MSP present in the eluted sample using UV-

quantified protein standards of PglC and MSP (Figures 5-11 and 5-12). As shown in Table 5-3, an MSP:PglC ratio of 10:1 was effective for obtaining approximately 0.9 copies of PglC per disc.

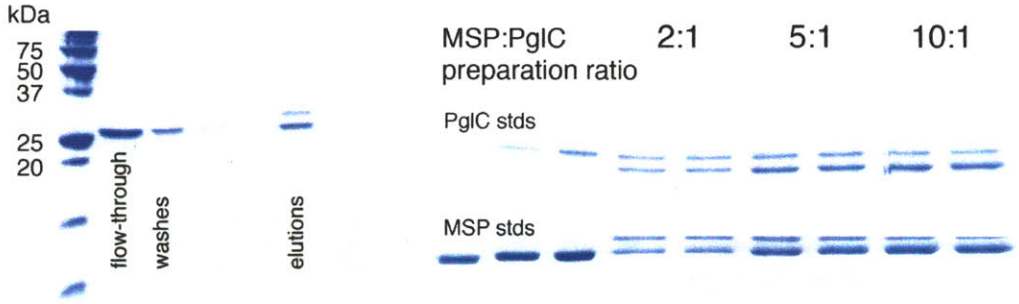


Figure 5-11. Purification and quantification of PglC-containing Nanodiscs.
Left, SDS-PAGE analysis of Ni-NTA purification of PglC-containing Nanodiscs. *Right*, SDS-PAGE analysis for gel densitometry analysis. The top gel shows MSP standards and the lower gel shows PglC standards followed by samples of the MSP:PglC preparations at 2:1, 5:1, and 10:1.

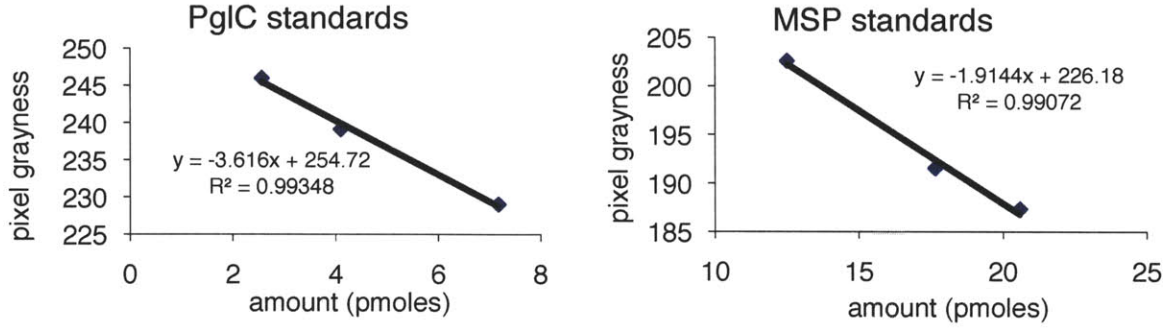


Figure 5-12. Standard curves from the gels depicted in Figure 5-11.
Gel densitometry is performed by correlating pixel grayness to the amount of protein standard loaded on the gel.

Table 5-3. PglC and MSP quantification in 12 nm Nanodiscs using gel densitometry.

MSP:PglC	MSP (grayness)	MSP (pmoles)	PglC (grayness)	PglC[a] (pmoles)	PglC/MSP	PglC/Nanodiscs
2:1	222.6	1.8	229.9	8.2	4.5	8.9
	219.8	3.3	227.7	9.0	2.7	5.4
5:1	197.1	15.2	224.0	10.2	0.7	1.3
	197.5	15.0	223.4	10.4	0.7	1.4
10:1	186.7	20.6	225.9	9.5	0.5	0.9
	187.1	20.4	228.1	8.8	0.4	0.9

[a]For the PglC quantification gel in Figure 5-11 (top right), the six Nanodisc lanes contain only 10 µL of sample, whereas all other lanes contain 12 µL of sample. The amounts, shown in pmole, in this column take into account that correction factor.

At this point, the reproducibility of PglC-insertion into Nanodiscs suggested that further protein-protein interaction studies would be feasible and thus it was desirable to use the largest MSP construct available at that time (MSPD1E3, 13 nm). As previously mentioned, His_6-cleaved MSPD1E3 migrates very similarly to PglC, but it was discovered that separation of the two proteins could be attained using a SDS-PAGE gel with 15% acrylamide. Therefore, it was necessary to identify the appropriate ratio to obtain one copy of PglC per Nanodisc for the larger MSP construct. Nanodiscs were prepared with MSP:PglC ratios of 5:1, 8:1 and 10:1 and were quantified by the gel densitometry method previously described (Table 5-4). A ratio of 10:1 was sufficient for ensuring one copy of PglC per disc.

Table 5-4. PglC and MSP quantification in 13 nm Nanodiscs using gel densitometry.

MSP:PglC	MSP (pmoles)	PglC (pmoles)	PglC/MSP	PglC/Nanodiscs
5:1	9.3	12.2	2.6	1.3
	9.4	12.9	2.8	1.4
8:1	6.2	7.7	2.5	1.2
	6.1	8.2	2.7	1.3
10:1	5.7	5.6	2.0	1.0
	5.4	4.8	1.8	0.9

SEC analysis (Figure 5-13), native agarose gels, and dynamic light scattering (DLS) further verified formation of PglC-Nanodiscs in the large MSP construct. By SEC, formation of a large assembly with an approximate diameter of ~15-16 nm was observed (Figure 5-13) and incorporation of both PglC and MSP was verified via SDS-PAGE analysis. Although the dimensions of a PglC-containing Nanodisc are unknown, the SEC analysis was consistent with the insertion of a membrane protein with a soluble domain into a Nanodisc, which has a 13 nm diameter in the absence of an incorporated protein. Similarly, native agarose gel analysis confirmed that PglC and MSP alone migrated distinctly from a PglC-Nanodisc complex, although this method required too much sample for it to be an effective analytical method. Finally, DLS is an effective means of determining the size distribution and population of a biological sample. DLS has proven to be a fast and reliable method to demonstrate that Nanodisc preparations contain particles that are 13-14 nm in diameter.

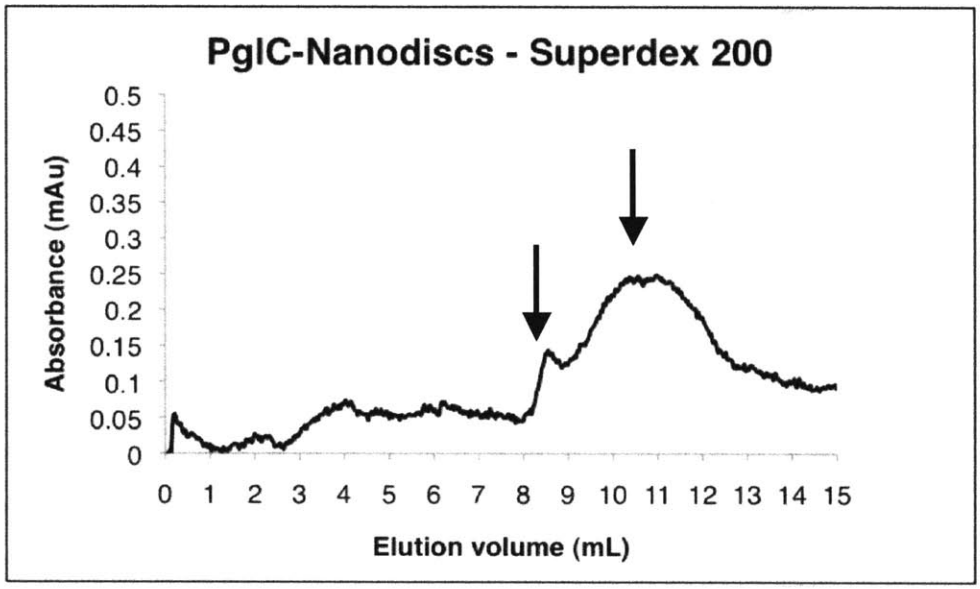


Figure 5-13. Size exclusion chromatography of PglC-containing Nanodiscs.
The first arrow indicates the void volume and the second arrow indicates the PglC-Nanodiscs.

PglA-containing Nanodiscs

Although PglA has no predicted transmembrane domains, it is still expected to associate with the membrane interface to interact with the membrane-bound undecaprenyl diphosphate carrier. In addition, as described previously, purification of PglA in a monodispersed state requires the presence of a small amount of detergent. Cross-linking of detergent solubilized PglA verifies that most of it remains monodispersed (Figure 5-14, lane 4). The presence of detergent complicates protein-protein interaction studies, because PglA cannot be titrated into the sample since the detergent in the sample would disassemble the Nanodiscs. Thus, PglA must be included in the Nanodisc assembly mixture. As an initial test of PglA incorporation, it was interesting to determine if PglA would incorporate into the discs in the absence of PglC. Nanodiscs were prepared as described for PglC with the exception that PglA contained the His_6 tag instead of PglC (Figure 5-10). Interestingly, it was observed that MSP co-eluted with PglA, suggesting PglA can interact with the membrane surface of Nanodiscs (Figure 5-14).

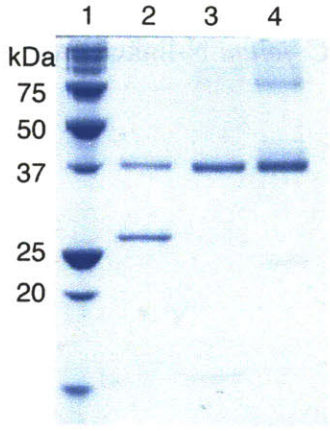


Figure 5-14. SDS-PAGE analysis of PglA-containing Nanodiscs.
Lane 1, molecular weight standards; lane 2, PglA-containing Nanodiscs; lane 3, purified PglA; lane 4, cross-linked PglA demonstrating that the PglA sample is mostly monodispersed.

Interaction of PglC and PglA in Nanodiscs

To determine if PglC and PglA would interact in the context of the Nanodisc membrane interface, a pull-down assay was performed in which PglC contained a His_6 tag, but PglA and MSP did not. Nanodiscs were prepared as described above for PglC (Figure 5-10) with the larger MSPD1E3 (13 nm), but detergent solubilized PglA was included in the preparation mixture at a 1:1:10 ratio of PglC:PglA:MSPD1E3. This is equivalent to a PglC:PglA:Nanodisc ratio of 1:1:5. Statistical distribution of PglA amongst empty and PglC-containing Nanodiscs would predict that 20% of the PglC-containing discs would also contain PglA. When the Nanodisc mixture was prepared and PglC was eluted from the Ni-NTA resin, both PglA and MSP were present. Subsequent quantification of the PglC-PglA-Nanodiscs by gel densitometry analysis suggested that 70% of PglC-containing Nanodiscs also contained PglA, which is much higher than the predicted random distribution of 20% (Figure 5-15 and Table 5-5). This result constitutes the first piece of evidence that PglC and PglA may have a specific interaction at the lipid bilayer surface. Furthermore, it validates use of Nanodiscs for further investigations into complex formation amongst the enzymes of the *C. jejuni* N-linked protein glycosylation pathway.

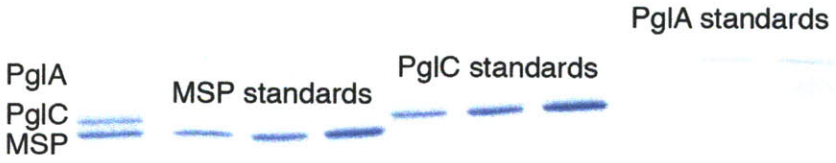


Figure 5-15. SDS-PAGE analysis of PglC-PglA-Nanodisc quantification by gel densitometry. The PglC and PglA Nanodiscs are shown in the first lane, and the quantified MSP, PglC, and PglA standards are shown in the subsequent lanes. It should be noted that PglA staining by Coomassie Blue is much lighter than MSP or PglC for a similar amount of protein.

Table 5-5. PglC, PglA and MSP quantification in 13 nm Nanodiscs using gel densitometry.

Protein	Amount (pmoles)		Ratios
MSP	13.9	PglC/MSP	2.6
PglC	8.9	PglC/Nanodisc	1.3
PglA	6.1	PglA/PglC	0.7

Conclusions

Establishing a viable and reproducible membrane mimetic system is essential for biochemical evaluation of how Und-PP-heptasaccharide biosynthesis occurs at the membrane interface (Figure 5-4). This chapter describes the initial validation of Nanodiscs as a method for exploring the protein-protein interactions in the *C. jejuni* protein glycosylation pathway. Protein purifications of PglC and PglA were optimized to obtain better yields (1-2 mg/L cell culture) of soluble, monodispersed protein. In addition, it was established that PglC can be efficiently incorporated into Nanodiscs and importantly, that PglA preferentially binds to PglC-containing discs. Further work described in Chapter 6 will focus on development of a biophysical method to probe the interaction between PglC and PglA in Nanodiscs. Importantly, we believe that this system may serve as a tractable model to elucidate general insights into polyprenyl-dependent biosynthesis, which is responsible for essential biomolecules such as bacterial peptidoglycan and eukaryotic N-linked protein glycans.

Acknowledgements

We would like to gratefully acknowledge the gift of MSP from the Sligar laboratory and the excellent technical assistance provided by Yelena Grinkova. In addition, I am thankful for the continued efforts of Vinita Lukose directed towards structural studies of PglC. We would also like to thank Professor Karen Allen and Drew Lynch for directing their X-ray crystallography expertise towards PglC. Finally, I am grateful to Vinita Lukose and Dr. Philipp Schneggenburger for reviewing this chapter and making helpful comments.

Experimental Procedures

Materials

The Sligar laboratory generously provided the MSPD1E2 and MSPD1E3 protein for the initial experiments described here. Subsequently, the DNA construct for MSPD1E3 was obtained through Addgene and the protein was expressed and purified as described below. The polar lipid extract from *E. coli* was obtained from Avanti Polar Lipids. Biobeads were obtained from BioRad and Amberlite XAD-2 (an more economical alternative to Biobeads) was obtained from Sigma. All other chemicals were obtained from Sigma unless otherwise noted.

DNA constructs

PglC was previously inserted into the pET-24 construct, which contained an N-terminal T7 tag (MASMTGGQQMG) and a C-terminal His_6 tag.[11] Alternate constructs were prepared (pET-NO, pGBH, pMAL and pGEX, Table 5-1) using the BamH I/Xho I restriction pair for digestion and ligation. Furthermore, new constructs were prepared that contained a His_6 tag or a FLAG tag N-terminal to the GB1 domain. PCR was used to amplify the *GB1-pglC* gene and insert the basepairs encoding the alternative purification tags at the N-terminus of GB1 (Table 5-6). The amplified region was inserted into the original pGBH vector using the Nde I/Xho I restriction pair. Finally, DNA encoding the TEV proteolytic sequence was inserted prior to the *pglC* gene by PCR amplification (Table 5-6). The *TEV-pglC* gene was inserted into the pMAL and pGBH-MH vectors using the restriction enzyme pair BamH I/Xho I.

Table 5-6. Primers used for preparation of PglC DNA constructs.
Restriction enzyme sites are italicized and inserted sequences are underlined.

Construct	Primers (5' to 3')
Fwd-His_6-GB1	GGAATCC*CATATG*CACCACCACCACCACCACCAGTACAAGCTTG CTCTGAACGGTAAAAC
Fwd-FLAG-GB1	GGGAATTC*CATATG*GACTACAAGGATGACGACGACAAGCAGTA CAAGCTTGCTCTGAACGGTAAAAC
Fwd-TEV-PglC	CGC*GGATCC*GAAAATCTGTATTTTCAGAGCATGTATGAAAAAG TTTTTAAAAGAATTTTTG
Rev-PglC-stop	CGC*CTCGAG*TCAGTTCTTGCCATTAAATTTC

Protein expression

PglC and MSP were expressed in the BL21-Gold (DE3) strain and PglA was expressed in the BL21-Codon Plus (DE3) RIL strain. For PglC and PglA, a 5 mL overnight culture of cells was prepared and used to inoculate 1 L of LB media with shaking at 37°C. After the cells

reached an optical density of ~0.8 absorbance units, the temperature was lowered to 16°C and the cells were induced with 0.5 mM IPTG. After 16-18 hours, the cells were harvested and the pellets were stored at -80°C.

For MSP, 2.5 mL of overnight culture was used to inoculate 0.5 L of Terrific Broth (TB) media at 37°C. The cells were induced with 0.5 mM IPTG at an optical density of 0.8 absorbance units and incubated with shaking for another 3-4 hours at 37°C. The cells were promptly harvested and the pellets were stored at -80°C.

Purification of PglC

To purify the integral membrane protein PglC, a cell envelope fraction (CEF) was first prepared. The cells were thawed in 40 ml of cold 50 mM Tris (pH 8.0), 150 mM NaCl (Buffer A) with 1 mg/ml lysozyme per L of cells. The cells were lysed at 4 °C by two intervals of sonication for 75 seconds each at an amplitude of 70% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. Cellular debris was cleared by centrifugation at 9000 x g for 45 minutes. The resulting supernatant was transferred to a clean centrifuge tube and subjected to centrifugation at 145,000 x g for 65 min to pellet the CEF. The CEF corresponding to 1 L of cell culture was resuspeneded in 1-5 ml of Buffer A and frozen at -80°C.

Alternatively, PglC was purified from the CEF on the same day. In either case, PglC CEF corresponding to 1 L of cell culture was homogenized in 10 mL of Buffer A containing 1% Triton X-100 (Sigma) using a glass homogenizer. Biorad assays suggested that 10 mL per L of cells was optimal for acheiving the recommended 3-5 mg/mL total protein recommended during detergent solubilization. The CEF was incubated with detergent for 1-2 hours and then

centrifuged again (145,000 x g) to remove insoluble material. The resultant supernatant was incubated with 0.5 ml of Ni-NTA resin per liter of cell culture for 1-2 hours. The resin was washed thoroughly with 40 column volumes of Buffer A supplemented with 20 mM imidazole and 0.05% Triton X-100 and 40 column volumes of Buffer A supplemented with 45 mM imidazole and 0.05% Triton X-100. The protein was eluted in Buffer A containing 300 mM imidazole and 0.05% Triton X-100 in 5 ml fractions. The most concentrated fractions were combined, dialyzed against Buffer A overnight and concentrated 10-fold. The purified PglC was aliquoted and stored at -80°C to avoid multiple freeze/thaw events. PglC was quantified by two methods: the MicroBCA assay (Pierce) based on comparison with a BSA standard and gel densitometry based on comparison with a PglC standard. The PglC standard was purified in an identical manner using a non-absorbent detergent (hydrogenated Triton X-100) and was quantified by UV using the extinction coefficient (ε_{280} = 31,400 $M^{-1}cm^{-1}$). The standard was stored at -80°C and used for gel densitometry quantification experiments.

Purification of PglA

PglA cell pellets from 2 L of cell culture were thawed in 80 mL of Buffer A. The cells were lysed at 4 °C by two intervals of sonication for 60 seconds each at an amplitude of 55% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. Following sonication, the lysate was incubated with 1% Triton X-100 for 20 minutes at 4 °C. The lysate was then centrifuged at 145,000 x g to remove cellular debris. The resultant supernatant was incubated with 1 mL of Ni-NTA resin for 1-2 hours at 4 °C. The resin was then washed with 40 mL of Buffer A containing 20 mM imidazole and 0.1% β-D-dodecylmaltoside (DDM, Anatrace) and 40 mL of Buffer A containing 45 mM imidazole and 0.05% DDM. The protein was eluted in 9 mL of Buffer A

containing 300 mM imidazole and 0.05% DDM. Purified PglA was then dialyzed overnight into 2-4 L of Buffer A. The next day PglA was concentrated to 1-2 mL, quantified and stored at -80 °C or the His_6 tag was removed from PglA by TEV proteolytic cleavage.

PglA TEV cleavage

Immediately following PglA purification and dialysis, a TEV cleavage reaction was prepared containing 9 mL of PglA protein (~5 μM), 500 μL of 20X TEV cleavage buffer (1 M Tris-HCl (pH 8.0) and 0.5 mM EDTA), 3.5 μL of β-mercaptoethanol and 500 μL of 200-500 μM TEV protease. The TEV protease was purified in the Imperiali laboratory using standard purification protocols. The reaction proceeded at 4 °C for 24-48 hours and the reaction progress was monitored by SDS-PAGE analysis. After the reaction has reached completion, PglA was dialyzed three times into 2-4 L of Buffer A over 36 hours. Following dialysis, PglA was incubated with 0.5-1 mL of Ni-NTA resin for 1-2 hours, the flow-through was then eluted and washed four times with 3 mL fractions of Buffer A. SDS-PAGE analysis was used to determine which fractions contain PglA and those fractions were combined, concentrated, and stored at -80 °C. PglA was quantified by UV absorption using the molar extinction coefficient (42,814 $M^{-1}cm^{-1}$). PglA monodispersity was verified by size exclusion chromatography performed using the Superdex 200 10/300 column (GE Healthcare). The running buffer contained Buffer A with 0.03% DDM.

Purification of MSP

MSP was purified following published protocols.[37] Cell pellets from 1 L of cell culture were thawed in 40 mL of lysis buffer containing 50 mM Tris-HCl (pH 8.0) and 300 mM NaCl. The cells were lysed at 4 °C by two intervals of sonication for 60 seconds each at an amplitude of 55% with 1 second on/off pulses using the sonicator (Sonics Vibra Cell). The cells were incubated on ice for five minutes in between rounds of sonication. Following sonication, the lysate was incubated with 1% Triton X-100 for 20 minutes at 4 °C. The lysate was then centrifuged at 145,000 g to remove cellular debris. The resultant supernatant was incubated with 3-4 mL of Ni-NTA resin for 1-2 hours at 4 °C. The resin was then washed with three different solutions. The resin was first washed with 5-10 column volumes of lysis buffer (pH 8.0) containing 1% Triton X-100, then 5-10 column volumes of lysis with 50 mM Na-cholate and 20 mM imidazole, and finally 5-10 column volumes of lysis buffer with 50 mM imidazole. The protein was eluted with lysis buffer containing 400 mM imidazole. The fractions containing protein determined by SDS-PAGE were combined and dialyzed against buffer containing 50 mM Tris (pH 8.0) and 150 mM NaCl. MSP was then concentrated to 50-100 μM and stored at -80 °C. Subsequent TEV cleavage of MSP was performed in the same way as TEV cleavage of PglA described above.

Lipid preparation and quantification

E. coli lipids total extract (Avanti) were used for all Nanodiscs experiments. An ampoule containing 100 mg of lipids dissolved in chloroform was distributed into 500 μL aliquots into disposable glass test tubes using a glass syringe. The chloroform was evaporated by drying under a nitrogen stream while rotating the test tube at an angle to generate a thin film of lipids. The

dried lipids were then resolubilized in 500 μL of buffer containing 50 mM Tris-HCl (pH 8.0) and 100 mM sodium cholate with vigorous vortexing and sonication. Each tube of detergent-solubilized lipids was then quantified by the following procedure and stored at -80°C.

The lipid quantification procedure was adapted from the Avanti Polar Lipids website. A solution containing 8.9 N H_2SO_4 was prepared in a volumetric flask by slowing adding 123.5 mL of concentrated H_2SO_4 to 376.5 mL of deionized water. The solution was prepared on ice to help dissipate the heat and stored at room temperature for up to 1 year. A 10% ascorbic acid solution was prepared in a 25 mL volumetric flask by dissolving 2.5 g of ascorbic acid in 25 mL total of deionized water and 2.5 % solution of ammonium molybdate (VI) tetrahydrate was prepared in a 25 mL volumetric flask by dissolving 0.625 g in 25 mL total of deionized water. Both the ascorbic acid and molybdate solutions were stored at 4°C in sealed 20 mL scintillation vials with protection from light exposure by aluminum foil.

A 0.65 mM phosphate standard (Sigma) was used to generate a standard curve. The standard curve ranged from 0 nmoles to 162.5 nmoles (250 μL) and were pipetted into separate glass tubes (100 x 13 mM). In addition, 2-3 μL of the detergent-solubilized lipids were aliquoted into glass test tubes. A heating plate with temperature control was used in conjunction with a metal test tube rack to maintain the high temperatures needed for lipid phosphate hydrolysis. Each phosphate standard and sample tube was incubated with 225 μL of 8.9 N H_2SO_4 for 25 minutes at 200-215°C. The temperature must be greater than 200°C to ensure phosphate hydrolysis. The tubes were removed from heat and allowed to cool for five minutes. Then 75 μL of H_2O_2 were added to each tube and the tubes were returned to the heating block at the same temperature for 30 more minutes. The samples were colorless at this point and were removed to cool to room temperature. Finally, 1.95 mL of deionized water was added to each tube followed

by 250 µL of the molybdate solution and 250 µL of the ascorbic acid solution. The solutions were mixed thoroughly with vortexing after each addition and incubated for an addition 7 minutes at 100°C. The standards and samples were then cooled to room temperature and the absorbance at 820 nm was determined using the spectrophotometer.

Incorporation of PglC into Nanodiscs

To prepare Nanodiscs that contain one copy of PglC, the following components were incubated at 4°C: 1 µM PglC, 10 µM MSPD1E3(-), 1.2 mM lipids (~30 mM stock solubilized in 100 mM cholate), and 25 mM sodium cholate (pH 7.0). While the mixture is incubating (15-20 minutes), the Biobeads (BioRad) or Amerlite XL beads (Sigma) are prepared in the manner previously described. A volume of beads equivalent to the volume of the Nanodisc mixture was measured and methanol (three times the volume of beads) was added. The methanol was removed by pipette and the beads were washed 2-3 more times with methanol in this manner. The beads were further washed with water with a vacuum filter. Water was added to the beads and vacuum was applied to remove the water. To avoid drying out the beads, the vacuum is removed after the water has been filtered away from the beads and the washing process is repeated 2-3 more times.

A 1 mL or 5 mL syringe is used to inject the Biobeads into the Nanodisc sample with minimal addition of water. To do this, the tip of the syringe is cut off using scissors and the water-washed Biobeads are pulled into the syringe. The syringe tip is placed onto a Kimwipe and excess water is pushed out. This process is repeated until the desired amount of Biobeads is contained in the syringe (500 µL Biobeads for a 500 µL reaction). Finally, the end of the syringe barrel is removed with scissors and the Biobeads are pushed out of the syringe into Nanodisc

pre-incubation solution. The Nanodisc-Biobeads mixture is then incubated at room temperature for 2-4 hours to ensure total removal of all detergent from solution.

In order to separate the Nanodisc mixture from the Biobeads, a 26-gauge needle was used to pierce the bottom and the top of the microcentrifuge tube containing the Nanodisc biobeads mixture. A hole was made in the cap of a 15 mL falcon tube, and the pierced tube containing Nanodiscs was placed into the cap. Brief centrifugation (1-2 minutes) allowed all of the Nanodisc solution to elute into the falcon tube.

PglC-containing Nanodiscs were purified from empty Nanodiscs by Ni-NTA affinity chromatography. A typical Nanodiscs reaction was incubated with ~100 μL of Ni-NTA resin for 1-2 hours at room temperature. The flow-through was collected and the resin was washed four times with 4 mL of PBS (pH 7.0). The PglC-containing Nanodiscs were eluted with PBS (pH 7.0) supplemented with 300 mM imidazole in 200-500 μL fractions. SDS-PAGE was used to verify Nanodisc formation.

Prior to size exclusion chromatography or other studies, the Nanodiscs were dialyzed overnight against 1-2 L of PBS (pH 7.0) to remove the imidazole. Size exclusion chromatography was performed with a Superdex 200 10/300 column and PBS was used as the running buffer. Dynamic light scattering was performed using a DynaPro Titan Dynamic Light Scatterer in the MIT Biophsyical Instrumentation Facility.

PglA-containing Nanodiscs and PglC-PglA-containing Nanodiscs were prepared in the same way as PglC-Nanodiscs, except that 1 μM PglA was included in the preparation mixture.

Gel densitometry analysis of Nanodiscs

Protein standards of PglC, PglA and MSP were quantified by UV spectroscopy using the expected extinction coefficients at 280 nm ($\varepsilon_{280,PglC}$ = 31,400 $M^{-1}cm^{-1}$, $\varepsilon_{280,PglA}$ = 42,814 $M^{-1}cm^{-1}$, $\varepsilon_{280,MSPD1E3(-)}$ = 26,600 $M^{-1}cm^{-1}$). PglC was purified with hydrogenated Triton X-100 in order to obtain a PglC sample quantifiable by UV to use as a standard for gel densitometry. Nanodisc samples were concentrated after dialysis and loaded onto 15% SDS-PAGE gels alongside known amounts of the protein standards. After staining with Instant Blue for 15-30 minutes, the gels were scanned and analyzed using Adobe Photoshop. Using the measurement tool, the average pixel grayness was recorded for the Nanodiscs and protein bands using a box of constant and defined size (slightly larger than the size of a protein band). The grayness values were used to generate a standard curve, which was used to estimate the amounts of PglC, PglA and MSP present in the Nanodisc samples.

References

1. Varki, A., Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **1993,** 3, (2), 97-130.
2. Szymanski, C. M.; Wren, B. W., Protein glycosylation in bacterial mucosal pathogens. *Nat. Rev. Microbiol.* **2005,** 3, (3), 225-37.
3. Banerjee, A.; Ghosh, S. K., The role of pilin glycan in neisserial pathogenesis. *Mol. Cell. Biochem.* **2003,** 253, (1-2), 179-90.
4. Weerapana, E.; Imperiali, B., Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology* **2006,** 16, (6), 91R-101R.
5. Dempski, R. E., Jr.; Imperiali, B., Oligosaccharyl transferase: gatekeeper to the secretory pathway. *Curr. Opin. Chem. Biol.* **2002,** 6, (6), 844-50.
6. Lennarz, W. J., Studies on oligosaccharyl transferase in yeast. *Acta Biochim Pol* **2007,** 54, (4), 673-7.
7. Szymanski, C. M.; Yao, R.; Ewing, C. P.; Trust, T. J.; Guerry, P., Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* **1999,** 32, (5), 1022-30.
8. Wacker, M.; Linton, D.; Hitchen, P. G.; Nita-Lazar, M.; Haslam, S. M.; North, S. J.; Panico, M.; Morris, H. R.; Dell, A.; Wren, B. W.; Aebi, M., N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **2002,** 298, (5599), 1790-3.
9. Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. U.S.A.* **2005,** 102, (40), 14255-9.
10. Glover, K. J.; Weerapana, E.; Numao, S.; Imperiali, B., Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* **2005,** 12, (12), 1311-5.
11. Glover, K. J.; Weerapana, E.; Chen, M. M.; Imperiali, B., Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* **2006,** 45, (16), 5343-50.
12. Olivier, N. B.; Chen, M. M.; Behr, J. R.; Imperiali, B., *In vitro* biosynthesis of UDP-N,N'-diacetylbacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* **2006,** 45, (45), 13659-69.
13. Alaimo, C.; Catrein, I.; Morf, L.; Marolda, C. L.; Callewaert, N.; Valvano, M. A.; Feldman, M. F.; Aebi, M., Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. *EMBO J.* **2006,** 25, (5), 967-76.
14. Chen, M. M.; Glover, K. J.; Imperiali, B., From peptide to protein: comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* **2007,** 46, (18), 5579-85.
15. Kowarik, M.; Young, N. M.; Numao, S.; Schulz, B. L.; Hug, I.; Callewaert, N.; Mills, D. C.; Watson, D. C.; Hernandez, M.; Kelly, J. F.; Wacker, M.; Aebi, M., Definition of the bacterial N-glycosylation site consensus sequence. *EMBO J.* **2006,** 25, (9), 1957-66.
16. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.

17. Seddon, A. M.; Curnow, P.; Booth, P. J., Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta* **2004,** 1666, (1-2), 105-17.
18. Rigaud, J. L.; Levy, D., Reconstitution of membrane proteins into liposomes. *Methods Enzymol* **2003,** 372, 65-86.
19. Sanders, C. R.; Prosser, R. S., Bicelles: a model membrane system for all seasons? *Structure* **1998,** 6, (10), 1227-34.
20. Chan, Y. H.; Boxer, S. G., Model membrane systems and their applications. *Curr Opin Chem Biol* **2007,** 11, (6), 581-7.
21. Nath, A.; Atkins, W. M.; Sligar, S. G., Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **2007,** 46, (8), 2059-69.
22. Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G., Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J Am Chem Soc* **2004,** 126, (11), 3477-87.
23. Bayburt, T. H.; Sligar, S. G., Self-assembly of single integral membrane proteins into soluble nanoscale phospholipid bilayers. *Protein Sci* **2003,** 12, (11), 2476-81.
24. Grinkova, Y. V.; Denisov, I. G.; Sligar, S. G., Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Eng Des Sel* **2010,** 23, (11), 843-8.
25. Fischer, N. O.; Blanchette, C. D.; Segelke, B. W.; Corzett, M.; Chromy, B. A.; Kuhn, E. A.; Bench, G.; Hoeprich, P. D., Isolation, characterization, and stability of discretely-sized nanolipoprotein particles assembled with apolipophorin-III. *PLoS One* **2010,** 5, (7), e11643.
26. Bayburt, T. H.; Leitz, A. J.; Xie, G.; Oprian, D. D.; Sligar, S. G., Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J Biol Chem* **2007,** 282, (20), 14875-81.
27. Boldog, T.; Grimme, S.; Li, M.; Sligar, S. G.; Hazelbauer, G. L., Nanodiscs separate chemoreceptor oligomeric states and reveal their signaling properties. *Proc Natl Acad Sci U S A* **2006,** 103, (31), 11509-14.
28. Denisov, I. G.; Grinkova, Y. V.; McLean, M. A.; Sligar, S. G., The one-electron autoxidation of human cytochrome P450 3A4. *J Biol Chem* **2007,** 282, (37), 26865-73.
29. Kimmich, N.; Das, A.; Sevrioukova, I.; Meharenna, Y.; Sligar, S. G.; Poulos, T. L., Electron transfer between cytochrome P450cin and its FMN-containing redox partner, cindoxin. *J Biol Chem* **2007,** 282, (37), 27006-11.
30. Nath, A.; Grinkova, Y. V.; Sligar, S. G.; Atkins, W. M., Ligand binding to cytochrome P450 3A4 in phospholipid bilayer nanodiscs: the effect of model membranes. *J Biol Chem* **2007,** 282, (39), 28309-20.
31. Cheng, Y.; Patel, D. J., An efficient system for small protein expression and refolding. *Biochem Biophys Res Commun* **2004,** 317, (2), 401-5.
32. Bouhss, A.; Trunkfield, A. E.; Bugg, T. D.; Mengin-Lecreulx, D., The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol. Rev.* **2008,** 32, (2), 208-33.
33. Price, N. P.; Momany, F. A., Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. *Glycobiology* **2005,** 15, (9), 29R-42R.
34. Newby, Z. E.; O'Connell, J. D., 3rd; Gruswitz, F.; Hays, F. A.; Harries, W. E.; Harwood, I. M.; Ho, J. D.; Lee, J. K.; Savage, D. F.; Miercke, L. J.; Stroud, R. M., A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat Protoc* **2009,** 4, (5), 619-37.

35. Arnold, T.; Linke, D., The use of detergents to purify membrane proteins. *Curr Protoc Protein Sci* **2008,** Chapter 4, Unit 4 8 1-4 8 30.
36. Boldog, T.; Li, M.; Hazelbauer, G. L., Using Nanodiscs to create water-soluble transmembrane chemoreceptors inserted in lipid bilayers. *Methods Enzymol* **2007,** 423, 317-35.
37. Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G., Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **2009,** 464, 211-31.

Chapter 6: Biophysical approaches to studying protein-protein interactions in Nanodiscs

Introduction

Nanodiscs represent a powerful approach to studying membrane proteins in a phospholipid bilayer. Nanodiscs are small discoidal membranes contained by a scaffold protein that encircles the disc protecting the hydrophobic edges of the bilayer.[1] Many types of integral membrane proteins have been functionally incorporated into Nanodiscs, demonstrating that Nanodiscs are a valuable platform for studying membrane protein activity in a defined biochemical environment.[1, 2] As described in Chapter 5, Nanodiscs were chosen as the model system to study the protein-protein interactions in the *C. jejuni* N-linked protein glycosylation pathway with the goal of providing evidence to support the proposed existence of the macromolecular biosynthetic complex presented in Figure 5-4. Additionally, Nanodiscs provide a discrete membrane model system in which to dissect the specific effects of polyprenyl-linked substrates on protein-protein interactions and complex formation. Initial work presented in Chapter 5 established that PglC, the initial phospho-glycosyltransferase in the membrane-committed phase of the Pgl pathway, can be readily incorporated into Nanodiscs and that PglA, the second glycosyltransferase, co-incorporates with PglC in a Nanodisc-based pull-down assay (Figure 6-1).

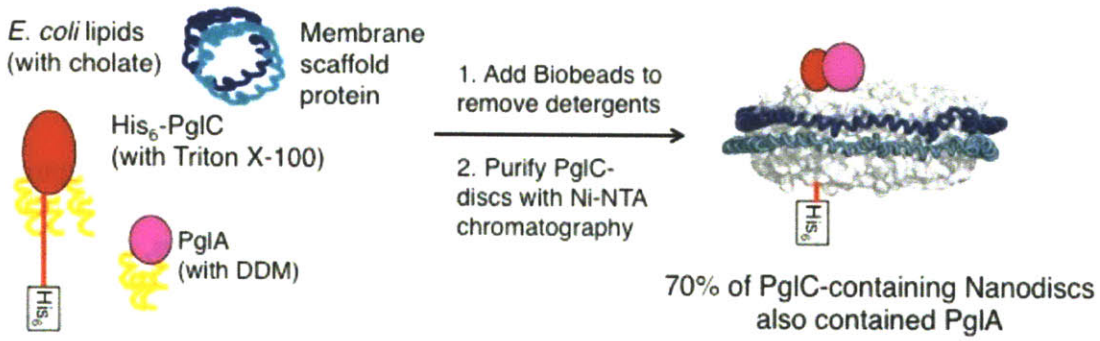


Figure 6-1. Co-incorporation of PglC and PglA into Nanodiscs.
Results from Chapter 5 are summarized, which demonstrated that PglC and PglA may interact specifically in the context of Nanodiscs.

Nanodiscs have been used to study a variety of known protein-protein interactions, including a cytochrome P450 with its related P450 reductase, epidermal growth factor (EGF) receptor with EGF, rhodopsin with transducin, and oligomeric complexes such as bacterial chemoreceptors.[3-6] In addition, Nanodiscs have also been exploited to identify previously undefined protein interactions such as that of the Syd protein, which was shown in a recent study to interact with the SecYEG complex involved in protein translocation.[7] In previous studies of this type, co-incorporation of interacting proteins was verified by Nanodisc-based pull-down assays, size exclusion chromatographic analysis of loaded Nanodiscs, and functional assays. A valuable extension of these biochemical approaches would involve spectroscopic methods such as Resonance Energy Transfer (RET), which examines protein-protein interactions and would provide more specific information on the interactions of PglC and PglA to the Nanodisc surface. RET methods rely on direct, non-radiative energy transfer from a luminescent or fluorescent donor to a fluorescent acceptor. In addition, RET is a distance dependent phenomenon such that it only occurs if the partners are in close proximity, 20-100 Å. The specific distance depends on

the fluorescence or luminescence properties of donors and acceptors that are selected for the RET system.

Two different RET approaches will be described in this chapter, fluorescence-RET (FRET)[8] and luminescence-RET (LRET).[9] FRET relies on a fluorescent donor and acceptor pair, which can be either small molecule fluorophores or encoded fluorescent proteins, and can report on distances from 20 – 80 Å. In the latter case, due to the size of the fluorophore-containing unit, distances derived are imprecise and qualitative rather than quantitative. Alternatively, LRET exploits sensitized lanthanide ion luminescence as the donor and a fluorescent small molecule as the acceptor.[10-14] LRET has a number of advantages over traditional FRET including long-lived luminescence of the donor, which allows for elimination of background excitation of the acceptor fluorophore in time-gated experiments. In addition, the unpolarized emission from the lanthanide metal allows for specific and efficient energy transfer allowing the application of LRET to measure distances up to 100 Å.

The development of genetically encodable lanthanide-binding-tags (LBTs) in the Imperiali group has made LRET an attractive and viable alternative to FRET since it circumvents the necessity to incorporate a lanthanide-binding site via chemical modification.[15, 16] LBTs are short peptide sequences that selectively bind lanthanide ions including terbium with low nM affinity. Excitation of a tryptophan residue within the LBT sensitizes bound-terbium ions resulting in a long-lived luminescence. The tryptophan-terbium pair is ideal because the emission properties of the indole of tryptophan are well matched for sensitizing terbium luminescence.[10] Recent studies in the Imperiali group have applied LBTs as LRET partners to provide Å-resolution information on the interaction distances between LBT-labeled SH2 domains and fluorescently-labeled phosphotyrosine peptide ligands (Figure 6-2).[11]

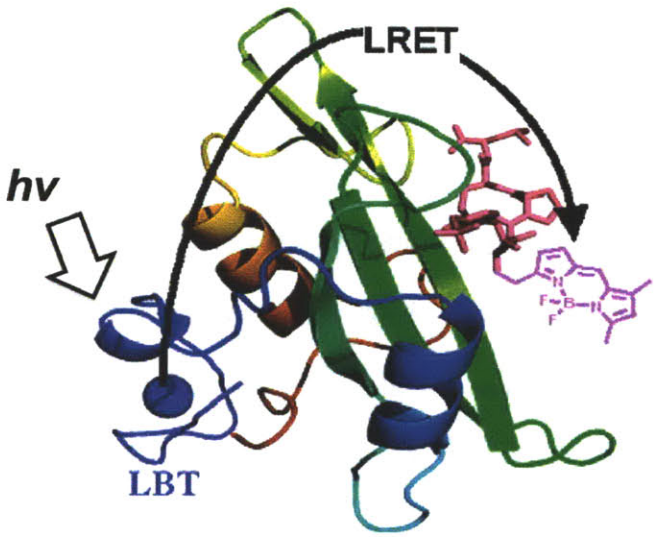


Figure 6-2. Depiction of LRET between an LBT and a Bodipy fluorophore.
A tryptophan in the LBT loop is excited thereby sensitizing the terbium ion luminescence, which can undergo energy transfer to the nearby Bodipy fluorophore. The RET only occurs when the fluorophore is incorporated into a phosphopeptide ligand specific for the LBT-labeled SH2 domain. This figure was adapted from the following reference.[11]

This chapter describes progress towards the application of FRET- and LRET-based approaches for investigating the interactions between PglC and PglA in the Nanodisc model membrane system. RET methods require proteins modified with fluorescent or luminescent labels. This chapter begins by describing a successful site-directed mutagenesis and cysteine-labeling strategy that generated two PglC Cys-containing variants that could be modified by maleimide-based fluorophore reagents. Furthermore, it is demonstrated that fluorophore-labeled PglC shows native enzymatic activity and can be readily incorporated into Nanodiscs, suggesting that it will be a viable partner in either FRET or LRET experiments.

Next, efforts toward obtaining active, fluorescently-labeled PglA for FRET-based analysis are described. Although the technical difficulties suggest that the non-specific Cys-

labeling approach undertaken described herein would not ultimately be successful for obtaining modified PglA as a potential FRET partner, strategies are identified for future initiatives toward this goal. As an alternative to FRET, the generation of two different LBT-tagged PglA proteins for LRET analysis is described. In this context, the co-incorporation of PglA-LBT and Bodipy FL-labeled PglC into Nanodiscs is established and although LRET experiments have not yet been performed, the key step of establishing the PglC/PglA LRET pair is a very important step towards this ultimate goal. Finally, this chapter presents preliminary experiments examining the potential application of amphipols as useful reagents for solubilizing PglA without the need for Nanodisc-disrupting detergents. Overall, this chapter summarizes important groundwork that has focused on adapting biophysical RET-based methodology for the identification of PglC and PglA interactions in Nanodiscs.

Results and Discussion

Fluorescent labeling of PglC

Fluorescent or luminescent modification of both PglC and PglA is required for RET studies. PglC is an excellent candidate for labeling by maleimide-based fluorophores, because it contains no native Cys residues. Thus, by making single site-directed mutations in the protein sequence, PglC proteins can be generated for site-specific labeling. Unfortunately, structural data for PglC is not yet available, which makes it more complicated to choose which residues to mutate. In order to derive a working model of the protein fold, the PglC sequence was submitted to four different secondary structure prediction programs: PROF,[17] SSPro,[18] APSSP2,[19] and PSIpred[20] (Figure 6-3). The aligned results shown in Figure 6-3 depict the predicted alpha-

helices in gray, the predicted beta-sheets in blue, and the predicted loops in yellow. The structural predictions programs suggested very similar topological features containing mostly alpha-helical regions with short interspersed β-sheet and loop structures. In addition, the predictions are consistent with transmembrane topology prediction by the program TMHMM[21] as shown previously in Figure 5-8 that suggested the presence of an N-terminal bilayer-spanning alpha helix.

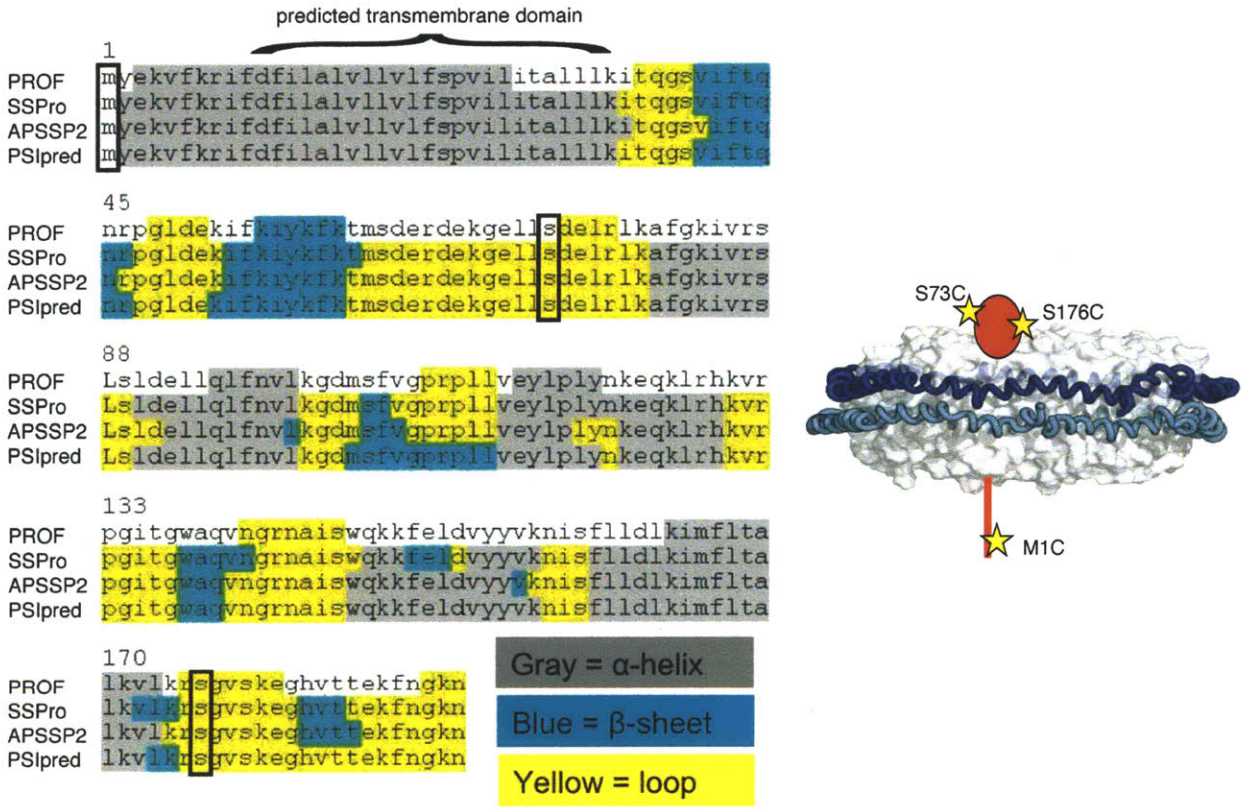


Figure 6-3. Secondary structure prediction of PglC and identification of PglC Cys mutant candidates.
Left, four different computational programs (PROF, SSPro, APSSP2, and PSIpred)[17-20] were used to predict the secondary structure of PglC. The gray regions were predicted to be α-helices, the blue regions are β-sheets, and the yellow regions are loops. The boxed regions indicate the three Cys mutants (M1C, S73C, and S176C) that were prepared. In addition, the region predicted to contain the N-terminal transmembrane domain is indicated. *Right*, a schematic representation of the three potential fluorophore-labeling sites on PglC is shown with respect to PglC predicted transmembrane topology and a Nanodisc.

From this information, three positions were chosen as sites for Cys mutations. The first site is the N-terminal Met, the second is Ser 73, which is in the center of a predicted loop region, and the last is Ser 176, found in a predicted C-terminal loop. The N-terminal Met is not essential for heterologous expression in this case, because PglC is expressed with an N-terminal GB1 fusion domain to improve solubility. The DNA constructs for the three Cys mutants were generated by site-directed mutagenesis and the PglC variants were expressed and purified using the optimized protocols described in Chapter 5. Upon purification, both PglC-M1C and PglC-S176C showed yields similar to the native protein, but the yield for PglC-S73C was 10-fold lower. In addition, activity assays using comparable amounts of the three mutant proteins suggested that the activity of PglC-S73C mutant was significantly reduced relative to the other two mutants and wild type PglC (Figure 6-4). Since Ser 73 falls within a predicted internal loop region (Figure 6-3), it is very possible that the Cys mutation may have disrupted the protein fold or protein stability in some way. Because of this result, subsequent experiments focused on the other two PglC variants, PglC-M1C and PglC-S176C, which are also called the N-terminal and C-terminal mutants, respectively.

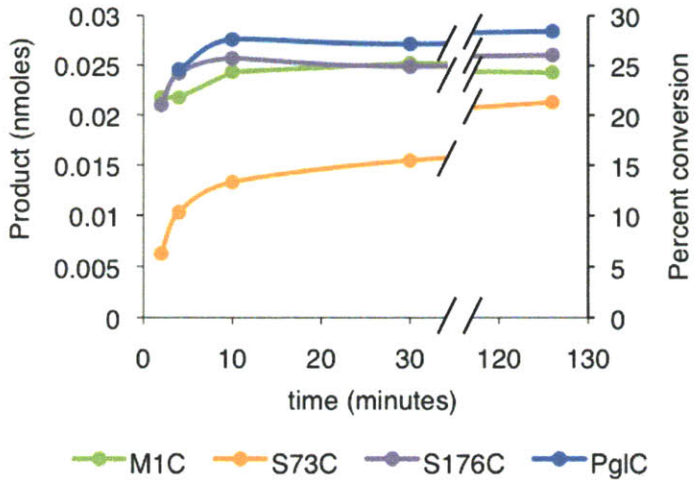


Figure 6-4. Activity of PglC and three Cys-containing PglC mutants.

The active PglC mutants were labeled by overnight incubation with Bodipy-FL maleimide, which shows excitation and emission properties similar to fluorescein, in the presence of reducing agent. Successful labeling was verified by SDS-PAGE (Figure 6-5) and excess fluorophore was removed by rebinding the protein to Ni-NTA resin and thorough washing. To determine labeling efficiency, the fluorophore was quantified by absorbance and the protein was quantified by a colorimetric protein assay. Based on the concentrations determined, the labeling efficiency of the PglC-M1C mutant was ~70% and the PglC-S176C mutant was ~50%. The activities of the fluorophore-labeled PglC proteins were compared to the activity of unlabeled PglC and were found to be unaffected by covalent modification with Bodipy-FL (Figure 6-6). These studies suggest that labeling PglC has a minimal effect on PglC structure and implies that the labeled PglC variants are good candidates for further FRET and LRET studies.

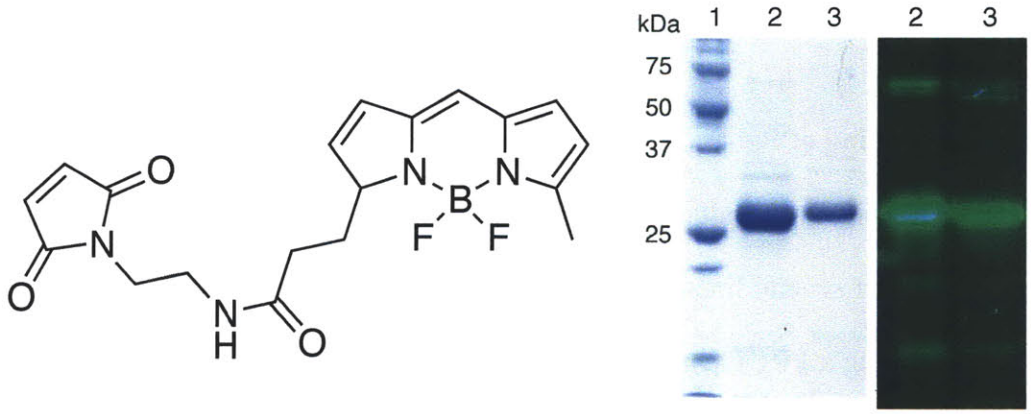


Figure 6-5. SDS-PAGE analysis of PglC labeled with Bodipy-FL.
Left, Structure of Bodipy-FL maleimide is shown. *Right*, The gel is shown stained with Coomassie blue stain or imaged using a transilluminator to detect the presence of fluorophore labeled protein. Lane 1, molecular weight standards; lane 2, PglC-M1C labeled with Bodipy-FL; lane 3, PglC-S176C labeled with Bodipy-FL.

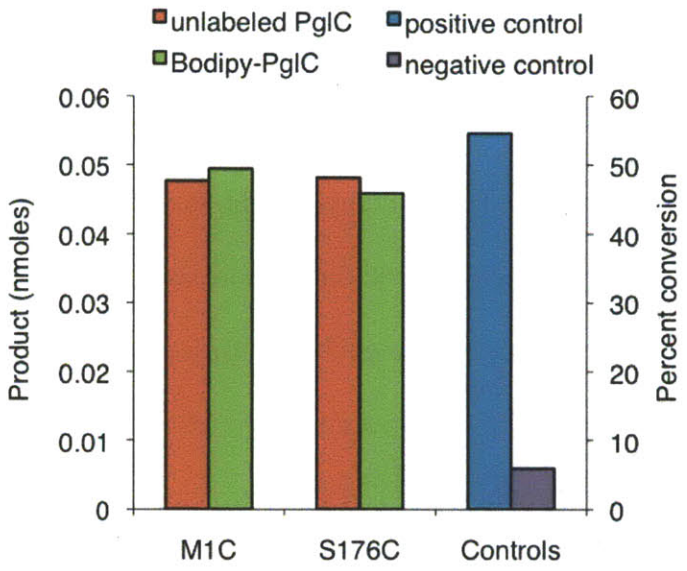


Figure 6-6. Comparison of PglC activity with and without fluorophore labeling.
The reactions were quenched after two hours and the transfer of radiolabeled glycan from the aqueous component to the organic layer was determined by scintillation counting. The activity of the modified (green) and unmodified (red) Cys-PglC variants is compared to wild type protein (blue) and a negative control without enzyme (purple)

Fluorescent labeling of PglA

PglA has six Cys residues making it less ideal than PglC for targeted Cys-based labeling strategies, however several lines of evidence suggested that this strategy might still be a feasible route for generating a FRET partner. For example, the quantitative absorption-based analysis of the reaction of PglA with 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB or Ellman's reagent) revealed the presence of only one reactive thiol in the protein. Similar to PglC, no structural data is available for PglA, but based on sequence homology evaluation, PglA is expected to have a glycosyltransferase-B (GT-B) protein fold, which contains two flexibly linked β/α/β Rossmann domains oriented toward each other.[22, 23] This fold is one of the most common structural motifs observed in most glycosyltransferases that have been structurally characterized to date.[22, 23] The predicted homology model of PglA was obtained using the I-TASSER prediction program.[24, 25]

In the derived model, the two Rossmann fold domains are clearly defined and by analogy with other GT-B fold glycosyltransferases it is predicted that the catalytic site of the protein would occur in the cleft between the two domains. Figure 6-7 shows the location of the six Cys residues within the predicted model. Interestingly, four of the Cys residues appear to be paired in two disulfide bonds in the interior of the structure, which would leave only two available Cys thiols for chemical modification. Further inspection of the model suggests that Cys 245 is on the inner side of a packed alpha helix, while Cys 146 is on an exterior loop region of the protein. This analysis suggests that only Cys 146 in PglA is likely to be surface exposed, which is consistent with DTNB labeling studies, which revealed through quantitative absorption spectroscopy measurements of the released TNB chromophore that only one thiol residue per protein was available for modification.

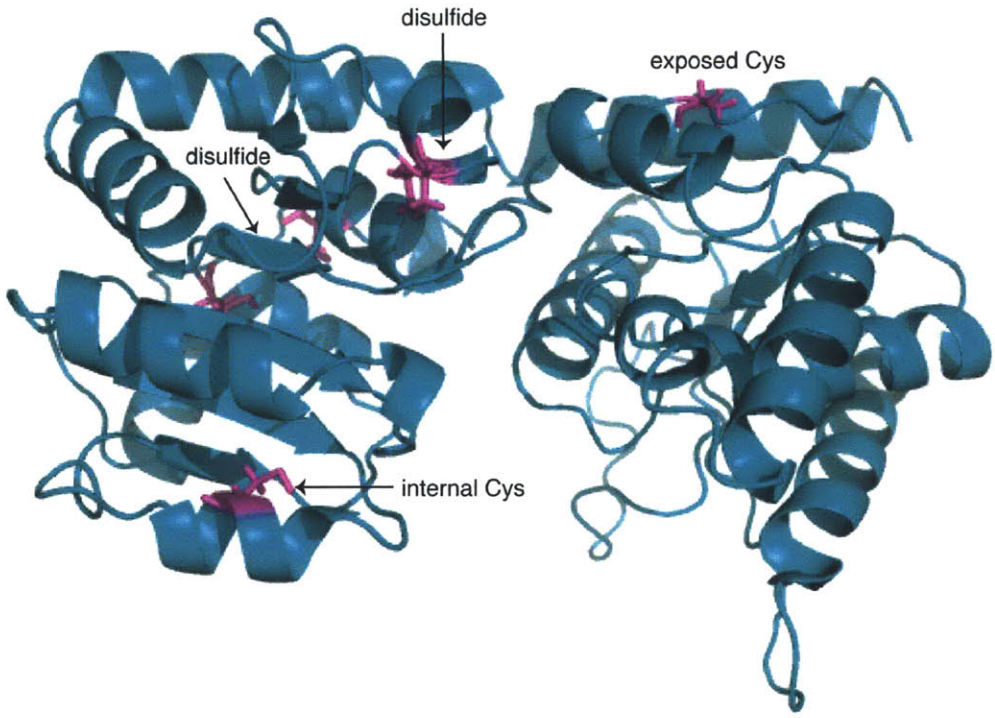


Figure 6-7. The I-TASSER predicted structure of PglA.
The six Cys residues are shown in purple. Cys 168 at the top right is the most surface exposed and proposed to the most available for modification by thiol-labeling reagents.

Based on the preliminary labeling studies and the homology model, which both suggested the presence of one available thiol, PglA was incubated with the maleimide version of AlexaFluor 568. This fluorophore was chosen a FRET partner for Bodipy-FL based on the spectral overlap of the two fluorophores and the relatively low propensity for excitation of the acceptor Alexafluor 568 at the Bodipy-FL excitation wavelengths (Figure 6-8). Overnight incubation of PglA with the fluorophore resulted in covalent modification as observed by SDS-PAGE. Quantification of PglA and the fluorophore as determined by absorbance indicated a labeling efficiency near 100%, however, it is unknown if this labeling occurred at one or multiple Cys residues. The excess fluorophore was mostly removed by purification with a desalting column (Figure 6-9).

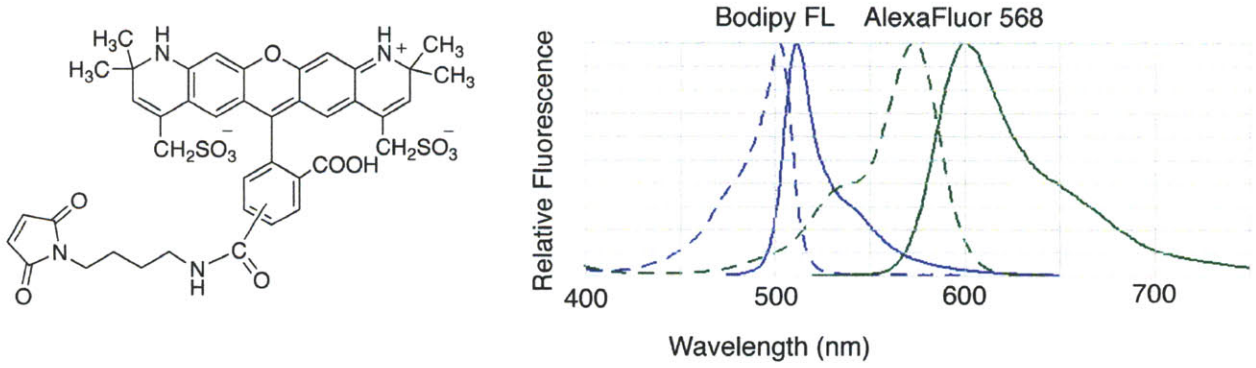


Figure 6-8. AlexaFluor 568 and Bodipy FL as FRET partners.
Left, the structure of AlexaFluor 568 is shown. *Right*, the emission (dashed) and excitation (solid) spectra of Bodipy-FL (blue) and AlexaFluor 568 (green) are shown.

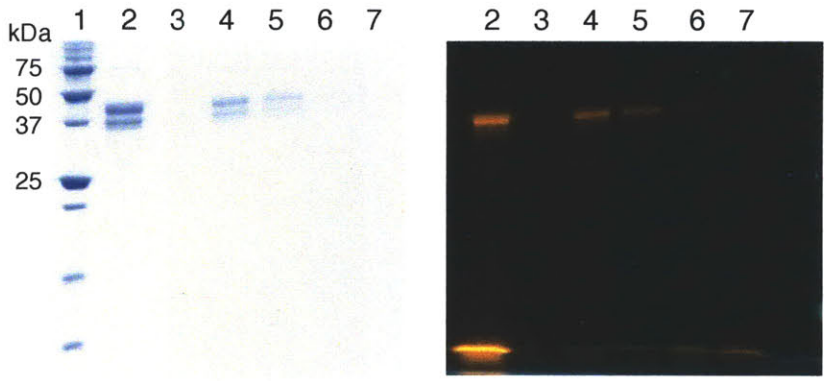


Figure 6-9. SDS-PAGE analysis of fluorophore labeled PglA.
Left, The gel is analyzed by staining with Coomassie blue. *Right*, Same gel is shown imaged using a transilluminator. Lane 1, molecular weight standards; lane 2, PglA labeling reaction; lane 3, flow-through from desalting column; lane 4-7, elution fractions from desalting column indicate that most of the excess fluorophore (lower band) has been removed from the protein containing fractions.

Unfortunately, upon labeling with Alexafluor 568, PglA showed no activity in functional assays. To further investigate the source of this inactivity, PglA was incubated with either DTNB or iodoacetamide and then assayed for activity. Interestingly, when PglA was reacted with excess iodoacetamide the enzyme showed comparable activity to the unlabeled protein, however, the DTNB-labeled PglA was completely inactive. These results suggested that blocking a catalytically important Cys residue was not the source of PglA inactivity. Instead, the loss of function was likely due to covalent modification of one or multiple Cys residues by the relatively large NTB moiety, which may have caused structural instability or unfolding. Furthermore, attempts to co-incorporate fluorophore-labeled versions of PglA and PglC into Nanodiscs were unsuccessful. Labeled-PglC readily incorporates into Nanodiscs, with yields similar to the native PglC (Figure 6-10). In contrast, very little of the fluorophore-labeled PglA was associated with

the purified PglC-containing Nanodisc fractions (Figure 6-10). Quantification by absorbance suggested that less than 10% of the PglA was incorporated relative to PglC.

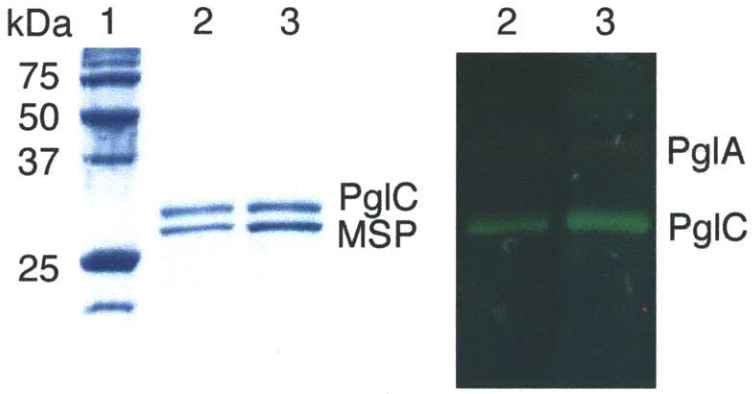


Figure 6-10. SDS-PAGE analysis of Nanodisc samples containing fluorescently-labeled versions of PglC and PglA.
Left, The gel is analyzed by staining with Coomassie blue. *Right*, Same gel is shown imaged using a transilluminator. Fluorescently-labeled PglA is only faintly visible in the PglC Nanodiscs when examined by fluorescence emission (right). Lane 1, MW standards; lane 2, N-terminal labeled PglC Nanodiscs with PglA; lane 3, C-terminal labeled PglC Nanodiscs with PglA.

As mentioned above, the fluorescent labeling procedure appeared to inactivate PglA by potentially disrupting the tertiary structure and thus it is not surprising that it showed weak incorporation into PglC-containing Nanodiscs. These studies provide some insight into future directions for the application of FRET to the study of PglC and PglA interactions. In particular, it is likely that Cys engineering is required to produce a fluorophore-labeling site that does not result in structural and hence functional perturbation. First, identification of the covalently modified Cys by tryptic-digestion and mass spectrometry would provide verification of the PglA homology model and also key information for further mutational screens. Specifically, site-directed mutants would be prepared, in which the labeled Cys is eliminated and new Cys mutations, which are less likely to interfere with the structure and activity of PglA, are

introduced. Rational identification of potential modification sites within PglA would then provide the basis for future FRET studies.

Generation of PglA-LBT constructs

As an alternative to FRET-based methods, LRET was considered an attractive option since PglA could be generated with an LBT coexpression tag via standard molecular biology techniques and this would eliminate the need for post-purification labeling approaches. Previous studies in the group resulted in the development of several LBT constructs with both N- and C-terminal single- and double-LBT fusions.[11, 15, 16, 26, 27] For the purpose of the Nanodisc experiments, it was necessary to prepare two new constructs. Both constructs required a His_6 purification tag that could be subsequently cleaved by TEV proteolysis as described in Chapter 5. In addition, it was desirable to have a single-LBT to act as the LRET donor. The double-LBT developed recently binds two terbium ions and has a brighter luminescence and other advantageous properties for NMR and X-ray applications,[26, 27] but obviously, the presence of two terbium ions would complicate LRET analysis.

The N-terminal LBT construct was prepared by modifying an existing N-terminal dLBT construct that contained a His_6-TEV cleavage sequence on the 5' end of the double-LBT. PCR was used to amplify the DNA encoding His_6-TEV followed by the first LBT sequence, which was modified to encode the single-LBT (FIDTNNDGWIEGDELLA) identified with the highest binding affinity for terbium (the K_D for Tb(III) is 18 nM). The PCR product was then ligated into the original vector to generate a construct in which genes could be readily inserted using the restriction enzyme pair BamH I and Xho I (Figure 6-11). In addition, the C-terminal LBT construct was generated in collaboration with Marcie Jaffee. For this DNA construct, the pET-

NO vector, previously described in Table 5-1, which already contained the desired His_8-TEV sequence at the N-terminus, was modified by a two-step PCR strategy to generate a pET-NO-LBT construct containing a C-terminal single-LBT construct (Figure 6-11). Steps were taken to ensure that this construct would also contain the BamH I/Xho I restriction pair for facile insertion of any protein of interest from the *C. jejuni* Pgl pathway. For incorporation of LBT-PglA (N-terminal LBT-PglA) or PglA-LBT (C-terminal PglA-LBT) into Nanodiscs, these two proteins were expressed, purified and the $His_{6/8}$ tags were removed by TEV proteolysis following the protocol presented in Chapter 5.

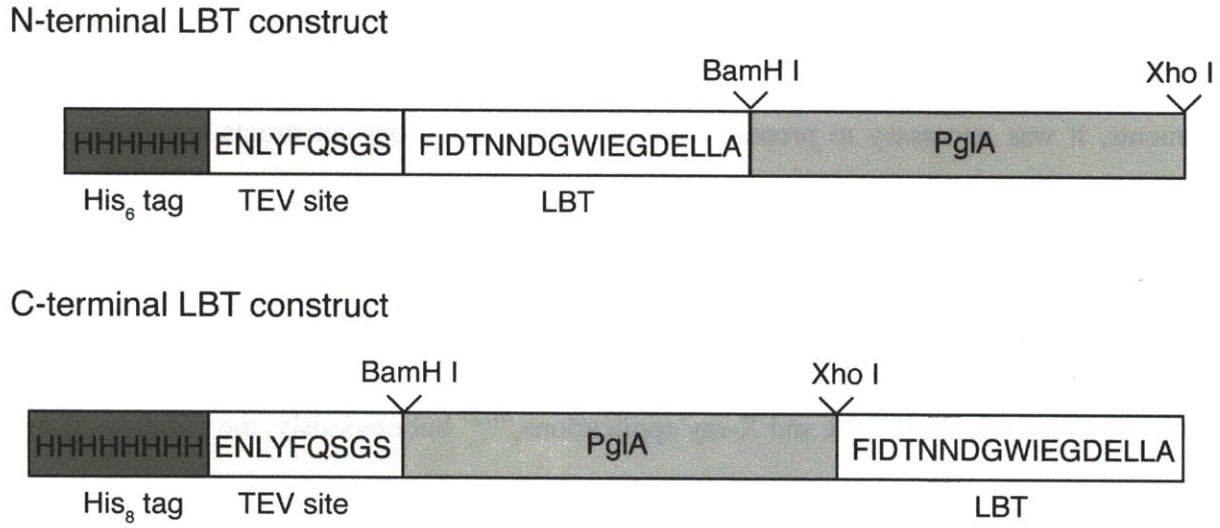


Figure 6-11. Diagram of the N-terminal and C-terminal single-LBT constructs.

Preparation of Nanodiscs with LRET partners

For RET experiments, one major obstacle is determining whether two proteins are interacting in a specific manner or if RET is observed simply because the proteins are sequestered in the same Nanodisc. We have addressed this concern with our experimental design by labeling PglC with either an N-terminal or a C-terminal fluorophore. Since the N-terminus of

PglC is a predicted transmembrane domain, the N-terminal tag is expected to be on the opposite side of the bilayer from the soluble PglC domain and the expected interaction interface with PglA (Figure 6-12). In the same way, the C-terminal PglC tag is expected to be on the same side of the Nanodisc as the PglA binding partner. With the two PglC mutants, the N-terminal label will act as an internal control that should enable us to distinguish between PglA binding non-specifically to Nanodisc surfaces and PglA binding more specifically to PglC-containing Nanodiscs, since the LRET should be greatly diminished in the case of the N-terminal PglC label. The Förster distance (R_o) for the LBT-Bodipy-FL pair has been previously calculated as 40 Å, which suggests that LRET occurring across the width of the lipid bilayer (5.5 nm) should be greatly diminished relative to two proteins interacting on the same side.

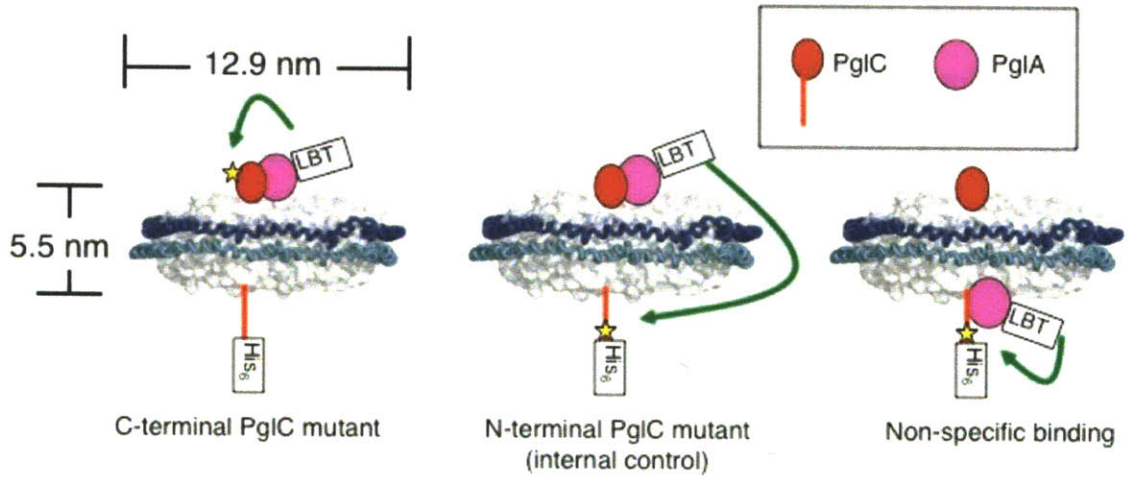


Figure 6-12. Experimental design for analysis of PglA and PglC-containing Nanodiscs by LRET.
A, C-terminally labeled PglC would be expected to show significant LRET upon specific interaction with PglA. *B*, N-terminally labeled PglC should have lower LRET if PglC and PglA are specifically interacting on the opposite side of the disc. *C*, in contrast, if PglA is interacting less specifically with PglC, then it may incorporated into either side of the Nanodisc equally and more LRET would be observed with the C-terminally labeled PglC protein.

In comparison to protein partners fluorescently labeled with small molecules, LBT-based LRET partners have several clear advantages. Like fluorescent proteins, they are genetically

encoded ensuring 100% tagging efficiency, however, unlike fluorescent proteins, the short LBTs located at the protein termini are less likely to perturb protein structure or interfere with protein binding interactions and function. PglC-Nanodiscs were prepared in the presence of native PglA, LBT-PglA, and PglA-LBT. SDS-PAGE analysis, shown in Figure 6-13, suggests that only native PglA and PglA-LBT can incorporate into PglC-containing Nanodiscs. In contrast, LBT-PglA, which has the LBT at the N-terminus of the protein, shows no association with the Nanodiscs. This suggests that the N-terminus may be an important region of PglA for interaction with the phospholipid bilayer or PglC. Interestingly, since the specific placement of the LBT at the N-terminus of PglA interferes with PglA binding to PglC Nanodiscs, this in turn provides support for the proposal that the interaction observed between native PglA and PglC-containing discs is specific. Furthermore, C-terminal PglA-LBT has been successfully co-incorporated into Nanodiscs containing either of the two fluorophore-labeled PglC mutants, which suggests that LRET studies should be feasible (Figure 6-13). Work is ongoing in the laboratory to optimize this method for investigations of the potential PglC-PglA interactions in Nanodiscs.

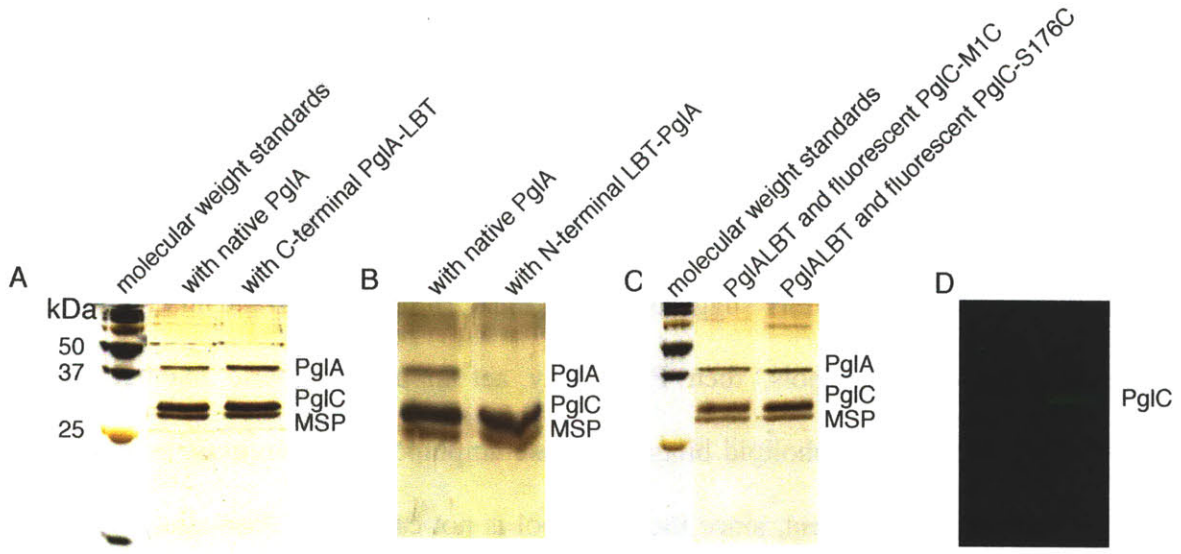


Figure 6-13. SDS-PAGE analysis of PglC-Nanodiscs with PglA, LBT-PglA and PglA-LBT. The gels are visualized with silver staining, since use of Coomassie Blue results in poorly stained PglA bands. *A,* Both native PglA or C-terminal LBT-PglA will co-incorporate into PglC-containing Nanodiscs. *B*, A comparison of native PglA or N-terminal LBT-PglA incorporation into PglC-containing Nanodiscs demonstrates that the N-terminal LBT-PglA will not incorporate. *C*, Nanodiscs are prepared with C-terminal PglA and the two fluorophore-labeled PglC variants. *D,* Same gel representation as ***C*** imaged with a transilluminator to demonstrate incorporation of fluorescent PglC.

Use of Amphipols to solubilize PglA

One major limitation of the experiments described above is the requirement for co-incorporation of PglA into the initial Nanodisc assembly mixture, because of the detergent required to solubilize PglA. In this regard, it would be advantageous to solubilize PglA in a way that would not disrupt the Nanodisc assembly, as this would greatly improve the experimental design. In this case, PglA could be added to the PglC Nanodiscs in a second, controlled step of the assembly process. This would broaden the scope of the experimental possibilities, since it would be feasible to perform titrations and attain more quantitative binding information by LRET or other biophysical methods including microscale thermophoresis.[28, 29]

To do this, PglA must be solubilized with a non-membrane-permeabilizing detergent. One interesting option developed in recent years is the amphipol class of reagents[30, 31] (Figure 6-14). These polymer-based molecules act like weak detergents in that they can bind and solubilize hydrophobic patches on proteins, but they cannot extract proteins from phospholipid bilayers. In the context of our experiments, this suggests that if PglA is solubilized and maintained in a monodispersed form by amphipols, then PglA may act in a monomeric manner without disruption of the Nanodisc phospholipid bilayer. To use amphipols, protein must be purified as normal in the presence of detergent, since the amphipol is not capable of extracting the protein from native phospholipids. After purification, a small amount of amphipol, ~4 mg amphipol per mg of protein has been shown to be optimal, is incubated with the detergent-solubilized protein. The detergent is removed with dialysis or by the Biobead method described in Chapter 5 and the resultant sample contains only protein and amphipol. When this was performed with PglA, size exclusion chromatography demonstrated that PglA was monodispersed in the presence of amphipols (Figure 6-14). Furthermore, preliminary results from dynamic light scattering experiments suggest that Nanodiscs are stable in the presence of amphipol-solubilized PglA and that this may be a viable option for solubilization of PglA for future protein-protein interaction studies in Nanodiscs.

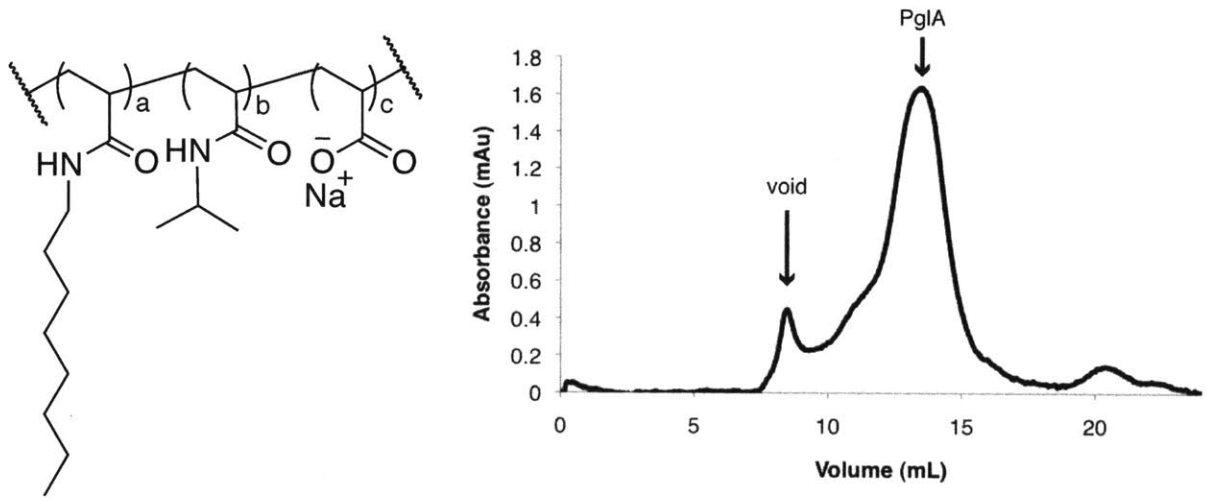


Figure 6-14. Amphipol-solubilized PglA.
Left, The structure of amphipol A8-35 is shown. The polymer is randomly composed of the three subunits, where a, b, and c comprise approximately 25%, 40% and 35%. Each polymer strand contains ~60 monomers. *Right*, Size exclusion chromatography analysis of PglA solubilized with amphipol A8-35 with UV detection at 280 nm.

Conclusions

Biophysical techniques, such as FRET and LRET, are useful tools for examining protein-protein interactions. This chapter describes efforts towards labeling PglC and PglA to generate FRET and LRET pairs that can potentially be used for RET experiments. Several Cys-variants of PglC were generated and labeled with maleimide-based Bodipy fluorophore. The labeled PglC proteins are active and can be readily integrated into Nanodisc lipid bilayers. Importantly, one of labeled versions of PglC is modified at the N-terminus, which is predicted to be on the opposite side of the Nanodisc from the PglC soluble domain and putative PglA interaction domain. This variant can help to distinguish between PglA interacting non-specifically with the lipid bilayer or PglA interacting with both the lipid bilayer and PglC.

Two approaches were undertaken to modify PglA as either a FRET or LRET partner. In the first, PglA was modified using a thiol reactive AlexaFluor for FRET studies. A non-specific

strategy was used based on preliminary labeling studies and structural modeling that suggested PglA contained only one available Cys residue. Although fluorescently-labeled PglA was prepared by this method, the enzyme was not active and did not incorporate into Nanodiscs with a similar efficiency to the unlabeled protein. In the second approach, PglA was expressed with a genetically encoded LBT at either the N-terminus or the C-terminus of the protein for LRET studies. Interestingly, only the C-terminal LBT variant incorporated into Nanodiscs at the same levels as native PglA, indicating that the N-terminal LBT interrupted the binding interaction between PglA and the PglC-containing Nanodiscs. LRET experiments are underway that will hopefully demonstrate the specific interaction of PglA and PglC in Nanodiscs.

Acknowledgements

I would like to thank Dr. Cliff Stains for his insightful comments and advice about this chapter, and in general, for sharing his knowledge about fluorescence and related experimental details. In addition, I am grateful to Dr. Langdon Martin for his past and present help with all things related to LBTs and luminescence. Finally, I want to acknowledge the collaborative effort with the excellent Marcie Jaffee in the generation of C-terminal LBT DNA construct.

Experimental Procedures

Preparation of DNA constructs encoding PglC Cys mutants, LBT-PglA, and PglA-LBT

All DNA constructs described below were verified by DNA sequencing. The PglC Cys mutants were prepared by site directed mutagenesis following the Quickchange protocol (Stratagene). The primers used to generate the PglC mutants containing M1C, S73C, or S176C are shown in Table 6-1.

Finally the C-terminal LBT vector was generated from the pET-NO vector described earlier in this thesis (Table 5-1), which contains His_6-TEV-PglA. The PglA gene was amplified by two rounds of PCR (Table 6-1, C-sLBT) to generate an insert containing BamHI-PglA-XhoI-sLBT-EcoRI. The BamH I/EcoR I restriction pair was then used to insert the DNA into the original pET-NO vector to generate the construct. This required deletion of the Xho I site already present in the pET-NO vector and insertion of an EcoR I site. First, PCR was used to amplify the PglA gene with an EcoR I site at the C-terminus (Table 6-1, EcoR I), and the amplicon was ligated into the original vector using the BamH I and Xho I restriction pair. Finally, site directed mutagenesis was used to mutate the Xho site (Table 6-1, Xho I).

Table 6-1. Primers used to generate PglC Cys mutants and PglA-LBT DNA constructs

Primer	Sequence (5' to 3')
Fwd PglC M1A	CCGAAGGATCCGAAAATCTGTATTTTCAGAGC***TGC***TATGAAAA AGTTTTTAAAAGAATTTTTG
Rev PglC M1A	CAAAAATTCTTTTAAAAACTTTTTCATA***GCA***GCTCTGAAAATA CAGATTTTCGGATCCTTCGG
Fwd PglC S73A	GAGATGAAAAGGGTCGAGTTATTA***TGC***GATGAATTGCGTTTG
Rev PglC S73A	CAAACGCAATTCATC***GCA***TAATAACTCACCCTTTTCATCTC
Fwd PglC S176A	ATGTTTTTAACAGCTTTAAAGGTTTTAAAACGA***TGC***GGGGTAA GCAAAGAAGG
Rev PglC S176A	CCTTCTTTGCTTACCCC***GCA***TCGTTTTAAAACCAAATTTGCTGT TAAAAACAT
Fwd N-sLBT	GCTCTCAAGGGCATCG***GTCGAC***GCTCTCCC
Rev N-sLBT	CGC***GGATCC***CGCCAGCAGTTCATCCCCTTCGATC

Fwd Y to F	GTATTTTCAGGGTCCGGGC***TTC***ATCGATACCAATAACGATGG
Rev Y to F	CCATCGTTATTGGTATCGA***TGA***AGCCCGGACCCTGAAAATAC
Fwd C-LBT	GTATTTTCAGAGCGGATCCATGAGAATAG
Rev C-LBT1	ATAGGGTAATTAAGAATGTACTCGAGTTCATTGATACCAATAA TGATGGCTGGATTG
Rev C-LBT2	CGATGAATTCTCACGCCAGCAGTTCATCACCTTCAATCCAGCC ATCATTATTGGTATC
Fwd EcoR I	Fwd C-LBT
Rev EcoR I	CCGCTCGAG***GAATTC***TCATACATTCTTAATTACCC
Fwd Xho I	GAATGTATGAGAATTC***CTGCAG***TGACACCACCACCACC
Rev Xho I	GGTGGTGGTGGTGTCA***CTGCAG***GAATTCTCATACATTC

The N-terminal LBT vector was prepared by modifying an existing dLBT vector that contained His_6-TEV-dLBT-PglA. Using a restriction site upstream (UP) of the gene (Sal I), an insert was amplified by PCR containing SalI-UP-His_6-TEV-sLBT-BamHI, where sLBT refers to the first of the two LBTs in the dLBT sequence originally encoded. The primers used for the PCR reaction are shown in Table 6-1 (N-sLBT). The insert was cloned into the original vector with the Sal I/BamH I restriction pair. Finally, it was necessary to make a single Y to F amino acid change in the first residue of the LBT in order to express the LBT that had the tightest binding affinity for terbium. This change was affected by site directed mutagenesis using the primers in Table 6-1 (Y to F).

DTNB and iodoacetamide labeling of PglA

A 10 mM DTNB solution or a 10 mM iodoacetamide solution was prepared in buffer containing 50 mM HEPES (pH 7.4) and 1 mM EDTA. For the DTNB quantification of free thiols, PglA was incubated at three different concentrations (0.4 μM, 0.8 μM and 1.5 μM) with 0.1 mM DTNB as well as a control reaction without PglA. Absorbance readings of the released DTNS thiol (ε_{412} = 14150 $M^{-1}cm^{-1}$) were measured at 10 minutes, 20 minutes, and 2 hours. The values were corrected for background, averaged and corrected for the dilution factor. The results

suggested that the concentration of the undiluted, modified protein was 15-16 μM, which was identical to the concentration of the unmodified protein stock solution (Table 6-2). In the case of the 0.4 μM dilution, the signal to noise ratio was quite low and thus number was expected to be less accurate. The results presented in Table 6-2 implied that PglA had one thiol available for modification.

Table 6-2. DTNB quantification of PglA thiol content.

[Protein]	Dilution factor	Corrected [DTNS]	# of Modification
0.4 μM	40	23 μM	1.5
0.8 μM	20	16 μM	1.1
1.5 μM	10	17 μM	1.1

Subsequent experiments evaluated the activity of PglA labeled with DTNB or iodoacetamide. In these experiments, PglA at 40 μM was incubated with 2 mM iodoacetamide or 2 mM DTNB for one hour at room temperature. After incubation, PglA activity was analyzed in an assay containing 3% DMSO, 0.05% Triton X-100, 50 mM $MgCl_2$, 30 mM Tris-Acetate (pH 8.0), 1 μM PglA, 1 μM Und-PP-diNAcBac, and 90 nM UDP-[^{3}H]GalNAc with a final reaction volume of 100 μL. Aliquots of 15 μL were quenched at 1, 2, 5, 10, and 60 minutes into 1 mL of 2:1 $CHCl_3$:MeOH and the organic layer was washed three times with 400 μL aqueous solution containing 240 mL of methanol, 15 mL of chloroform, and 235 mL of water with 1.83 g of potassium chloride. The organic layer was dried and solubilized with 200 μL Solvable (Perkin-Elmer). For scintillation detection, the organic layer was mixed with 5 mL of OptiFluor (Perkin-Elmer) and the aqueous layer was mixed with 5 mL of Eco-lite (MP Biomedicals).

Fluorescent labeling of PglC and PglA

In general, PglC (20-40 μM) in 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl was incubated overnight at 4 °C in the presence of 10-fold excess fluorophore solubilized in DMSO (Bodipy-FL or AlexaFluor 568, Invitrogen) and 10-fold excess TCEP (Tris(2-carboxyethyl) phosphine). PglA was labeled in a similar manner, except that TCEP was either not added or only added as a single equivalent to avoid excess labeling. After overnight reaction, the excess fluorophore was removed in one of several ways. For the Bodipy-FL fluorophore, it was quite effective to rebind the protein to Ni-NTA resin and wash away the excess label. However, this did not work for the less soluble AlexaFluor 568 as this molecule precipitated in the presence of the Ni-NTA resin. Alternatively, excess AlexaFluor 568 was removed by using a 5 mL HiTrap desalting column (GE Healthcare). This method was effective, but did not result in complete fluorophore removal. Finally, if detergent is not present in the protein sample, as in the case of amphipol-solubilized PglA, then Biobeads can be used to remove the excess fluorophore. The Biobeads were used in the same way as described previously in Chapter 5.

Solubilization of PglA with amphipols

To solubilize PglA with amphipols, PglA was purified with detergent as described in Chapter 5. Then PglA was incubated for 30-60 minutes at 4 °C with amphipol (A8-35, Avanti Polar Lipids) at a ratio of 1 mg PglA to 4 mg A8-35. The detergent was then removed using Biobeads as described previously in Chapter 5, and the amphipol-solubilized PglA solution was stored at -80 °C.

References

1. Nath, A.; Atkins, W. M.; Sligar, S. G., Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **2007,** 46, (8), 2059-69.
2. Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G., Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **2009,** 464, 211-31.
3. Duan, H.; Civjan, N. R.; Sligar, S. G.; Schuler, M. A., Co-incorporation of heterologously expressed Arabidopsis cytochrome P450 and P450 reductase into soluble nanoscale lipid bilayers. *Arch Biochem Biophys* **2004,** 424, (2), 141-53.
4. Mi, L. Z.; Grey, M. J.; Nishida, N.; Walz, T.; Lu, C.; Springer, T. A., Functional and structural stability of the epidermal growth factor receptor in detergent micelles and phospholipid nanodiscs. *Biochemistry* **2008,** 47, (39), 10314-23.
5. Bayburt, T. H.; Leitz, A. J.; Xie, G.; Oprian, D. D.; Sligar, S. G., Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. *J Biol Chem* **2007,** 282, (20), 14875-81.
6. Boldog, T.; Grimme, S.; Li, M.; Sligar, S. G.; Hazelbauer, G. L., Nanodiscs separate chemoreceptor oligomeric states and reveal their signaling properties. *Proc Natl Acad Sci U S A* **2006,** 103, (31), 11509-14.
7. Dalal, K.; Nguyen, N.; Alami, M.; Tan, J.; Moraes, T. F.; Lee, W. C.; Maurus, R.; Sligar, S. S.; Brayer, G. D.; Duong, F., Structure, binding, and activity of Syd, a SecY-interacting protein. *J Biol Chem* **2009,** 284, (12), 7897-902.
8. Stryer, L.; Haugland, R. P., Energy transfer: a spectroscopic ruler. *Proc Natl Acad Sci U S A* **1967,** 58, (2), 719-26.
9. Selvin, P. R.; Hearst, J. E., Luminescence energy transfer using a terbium chelate: improvements on fluorescence energy transfer. *Proc Natl Acad Sci U S A* **1994,** 91, (21), 10024-8.
10. Allen, K. N.; Imperiali, B., Lanthanide-tagged proteins--an illuminating partnership. *Curr Opin Chem Biol* **2010,** 14, (2), 247-54.
11. Sculimbrene, B. R.; Imperiali, B., Lanthanide-binding tags as luminescent probes for studying protein interactions. *J Am Chem Soc* **2006,** 128, (22), 7346-52.
12. Sandtner, W.; Bezanilla, F.; Correa, A. M., In vivo measurement of intramolecular distances using genetically encoded reporters. *Biophys J* **2007,** 93, (9), L45-7.
13. Posson, D. J.; Selvin, P. R., Extent of voltage sensor movement during gating of shaker K+ channels. *Neuron* **2008,** 59, (1), 98-109.
14. Vazquez-Ibar, J. L.; Weinglass, A. B.; Kaback, H. R., Engineering a terbium-binding site into an integral membrane protein for luminescence energy transfer. *Proc Natl Acad Sci U S A* **2002,** 99, (6), 3487-92.
15. Franz, K. J.; Nitz, M.; Imperiali, B., Lanthanide-binding tags as versatile protein coexpression probes. *Chembiochem* **2003,** 4, (4), 265-71.
16. Nitz, M.; Franz, K. J.; Maglathlin, R. L.; Imperiali, B., A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides. *Chembiochem* **2003,** 4, (4), 272-6.
17. Ouali, M.; King, R. D., Cascaded multiple classifiers for secondary structure prediction. *Protein Sci* **2000,** 9, (6), 1162-76.

18. Pollastri, G.; Przybylski, D.; Rost, B.; Baldi, P., Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. *Proteins* **2002,** 47, (2), 228-35.
19. Raghava, G. P. S., A combination method for protein secondary structure prediction based on neural network and example based learning. . *CASP5* **2002**, A-132.
20. Jones, D. T., Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **1999,** 292, (2), 195-202.
21. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **2001,** 305, (3), 567-80.
22. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G., Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **2008,** 77, 521-55.
23. Roychoudhury, R.; Pohl, N. L., New structures, chemical functions, and inhibitors for glycosyltransferases. *Curr Opin Chem Biol* **2010,** 14, (2), 168-73.
24. Roy, A.; Kucukural, A.; Zhang, Y., I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **2010,** 5, (4), 725-38.
25. Zhang, Y., I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **2008,** 9, 40.
26. Martin, L. J.; Hahnke, M. J.; Nitz, M.; Wohnert, J.; Silvaggi, N. R.; Allen, K. N.; Schwalbe, H.; Imperiali, B., Double-lanthanide-binding tags: design, photophysical properties, and NMR applications. *J. Am. Chem. Soc.* **2007,** 129, (22), 7106-13.
27. Silvaggi, N. R.; Martin, L. J.; Schwalbe, H.; Imperiali, B.; Allen, K. N., Double-lanthanide-binding tags for macromolecular crystallographic structure determination. *J Am Chem Soc* **2007,** 129, (22), 7114-20.
28. Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S., Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* **2010,** 1, 100.
29. Baaske, P.; Wienken, C. J.; Reineck, P.; Duhr, S.; Braun, D., Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed Engl* **2010,** 49, (12), 2238-41.
30. Popot, J. L.; Berry, E. A.; Charvolin, D.; Creuzenet, C.; Ebel, C.; Engelman, D. M.; Flotenmeyer, M.; Giusti, F.; Gohon, Y.; Hong, Q.; Lakey, J. H.; Leonard, K.; Shuman, H. A.; Timmins, P.; Warschawski, D. E.; Zito, F.; Zoonens, M.; Pucci, B.; Tribet, C., Amphipols: polymeric surfactants for membrane biology research. *Cell Mol Life Sci* **2003,** 60, (8), 1559-74.
31. Seddon, A. M.; Curnow, P.; Booth, P. J., Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta* **2004,** 1666, (1-2), 105-17.

Chapter 7: Studies towards the functional reconstitution of the PglC and PglA glycosyltransferases into Nanodiscs and incorporation of undecaprenyl-phosphate-linked substrates

Introduction

Polyprenyl-phosphates and their derivatives mediate reactions in essential glycan biosynthetic pathways in both prokaryotes and eukaryotes. As discussed in Chapter 1, polyprenyl-phosphate derivatives are essential co-substrates, but little is known about why a linear polyprenol is the preferred glycan carrier in many biosynthetic pathways. One possible explanation is that the polyprenyl-phosphate acts as an important checkpoint for the regulation of cellular processes. Another possibility is that the specific structure of the linear polyprenol plays a functional role facilitating glycan translocation across lipid bilayers, a process that occurs in all polyprenyl-dependent pathways. A third hypothesis is that polyprenyl-phosphates associate with integral membrane proteins and act to coordinate and facilitate the concerted action of enzymes in glycan biosynthetic complexes. It is this last hypothesis that has driven the Nanodisc studies described in the preceding two chapters and that will be further explored in this final chapter.

Undecaprenyl-phosphate (Und-P) is the glycan carrier utilized in all bacterial polyprenyl-dependent glycan assembly pathways and it serves as the membrane anchor in *Campylobacter jejuni* N-linked protein glycosylation.[1] The molecular structure of undecaprenyl-phosphate, which is a C55 linear polyprenol (illustrated in Figure 7-1), is 50 Å long when fully extended, which is similar to the width of a typical phospholipid bilayer. In fact, undecaprenol is one of the shortest polyisoprenols found in nature. Most eukaryotes exploit the linear polyisoprenols of the dolichol family, which may extend from 80 to 90 Å in length, and select members of the *Potentilla* family in plants have even longer polyprenols with extended lengths of 150 to 160 Å.[2] Since these extended molecular dimensions are much greater than the width of phospholipid bilayers, it is clear that the dolichols must adopt a "folded" or "coiled" structure to fit within the membrane. Structural studies using small-angle X-ray scattering, NMR and molecular modeling

have suggested that polyisoprenols adopt a three-part structure in the membrane comprising a head domain at the hydrophilic end, a central coiled structure that includes the *cis*-isoprene units, and a tail domain including the *trans* isoprene units.[3-5] In addition, the dimensions of polyprenyl-phosphates in the structures that have been calculated[3-5] imply that they would reside mostly within one leaflet of the bilayer, due to the coiled structure. Polyprenyl-phosphates have unique structures relative to phospholipids; phospholipids are composed of shorter, mostly saturated and unbranched acyl chains with only 16-20 carbons and readily form lipid bilayers. In contrast, biophysical studies have established that polyprenyl-phosphates increase the fluidity of phospholipid bilayers in which they reside and promote formation of non-bilayer geometries. It is probable that these properties contribute to their physiological functions.[3, 4, 6-8] The structural and biophysical details of linear polyprenols are important when considering their potential influence on glycan assembly.

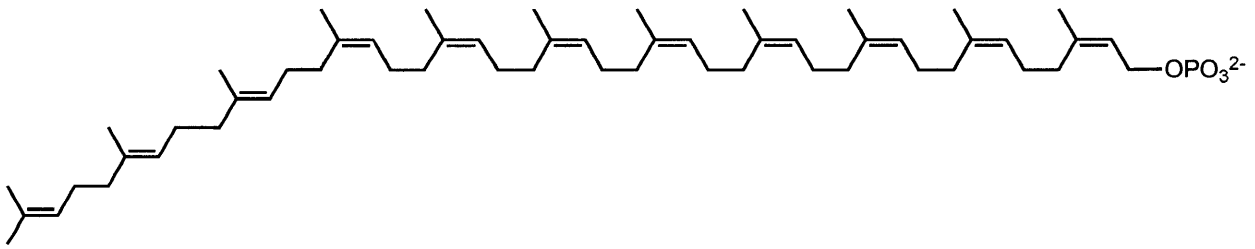


Figure 7-1. The chemical structure of bacterial undecaprenyl-phosphate (Und-P).

A proposed model of the *C. jejuni* N-linked protein glycosylation process was presented previously in Figure 5-4 and suggests the presence of a biosynthetic complex composed of the eight enzymes involved in production of the Und-PP-heptasaccharide. This chapter focuses on the functional aspects of the proposed biosynthetic complex. In principle, biosynthesis of

polyisoprenyl-linked glycans from all organisms could occur either through separate or coordinated enzymatic reactions. In the *C. jejuni* pathway, the first route would involve the glycosyltransferases performing reactions independently and releasing the Und-PP-glycan into the phospholipid bilayer following each transfer (Figure 7-2). In the second path, the enzymes would act in a coordinated manner to form the final glycan product without release of Und-PP-linked intermediates into the bulk phospholipid bilayer. This model is consistent with the presence of a biosynthetic complex all of the glycosyltransferases (Figure 7-2). These two models represent the extreme descriptions of how glycan biosynthesis could occur in this pathway, but it is also possible that some combination of these pathways exists *in vivo*.

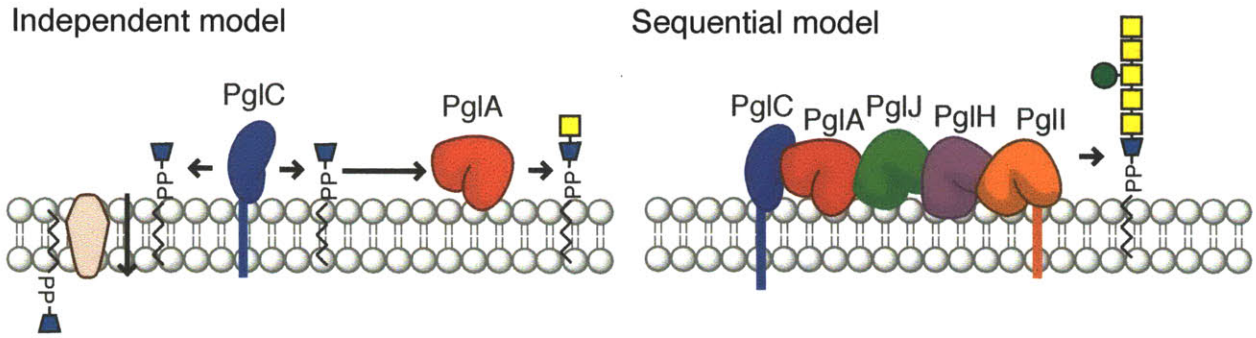


Figure 7-2. Independent and sequential models of glycosyltransferase activity in *C. jejuni*. *Left*, The first two enzymes in the *C. jejuni* pathway, PglC (blue) and PglA (red), are depicted independently in the lipid bilayer, each releasing its product after the transferase reaction. In this model, truncated glycans, such as the PglC product, could be translocated across the bilayer by a flippase (brown) and transferred to protein prematurely. *Right*, In this model, the five glycosyltransferases act in a sequential manner to generate the Und-PP-glycan without release of intermediate products.

One benefit of functional synergism in the second model is that the Und-PP-glycan intermediates would be directly shuttled from enzyme to enzyme, decreasing the likelihood of premature encounters with flippase enzymes. In addition, it is known that PglB, the oligosaccharyltransferase in *C. jejuni*, can recognize and transfer all of the truncated forms of the

glycan *in vitro*, however truncated glycans are not observed *in vivo*. In fact, if unfinished glycans are prematurely translocated and transferred to protein, this may have negative physiological effects. For instance, *C. jejuni* strains that lack PglH glycosyltransferase activity and cannot produce the full glycan structure, also display diminished ability to invade host cells.[9-11] This suggests that transfer of intermediate glycans to glycoproteins could weaken the virulence of the bacteria, implying that the pathway represents an antibiotic target.

A further implication of the sequential model of glycan biosynthesis is that undecaprenyl-phosphate may not act simply as a passive glycan transporter. Recent studies have identified protein-binding motifs that are involved in specific interactions with polyisoprenyl moieties and have been termed polyisoprenol recognition sequences (PIRS).[3, 4] These studies suggested that a single polyprenol could interact with four or five different PIRS peptide sequences, which are typically located in transmembrane domains. This implies that undecaprenyl-phosphate could play an active role in complex formation by interacting with the integral membrane proteins in the pathway and may be involved in diverse functions such as enzyme recruitment, stabilization of protein-protein interactions, and substrate shuttling through the pathway. Because polyprenyl-phosphates may have such important roles in glycan assembly, it is essential to consider their impact when examining protein-protein interactions within polyprenol-dependent pathways.

In order to address the function of undecaprenyl-phosphate in *C. jejuni* N-linked glycosylation in the context of Nanodiscs, the membrane-bound polyprenol must be incorporated into the Nanodisc bilayer. This chapter begins by presenting preliminary experiments that demonstrate incorporation of Und-PP-diNAcBac into Nanodiscs and activity of the enzymes, PglC and PglA, in Nanodiscs (Figure 7-3). However, the activity of PglA is greatly reduced in these preliminary experiments, because the PglA preparation was added to the assay in an

aggregated state due to the necessity to remove detergent from the preparation, which would otherwise have destabilized the Nanodisc. Following these initial experiments, this chapter presents two innovations to the functional assay that have improved the activity results. The first is the utilization of radiolabeled undecaprenyl-[^{33}P]phosphate, which allows for quantification of the PglC substrate. Secondly, co-incorporation of PglC, PglA, and undecaprenyl-[^{33}P]phosphate into Nanodiscs avoids the use of aggregate PglA in Nanodisc activity assays and results in improved activity. Importantly, this chapter provides the first, clear evidence that PglC and PglA are functionally reconstituted in Nanodiscs and establishes the framework for future functional studies.

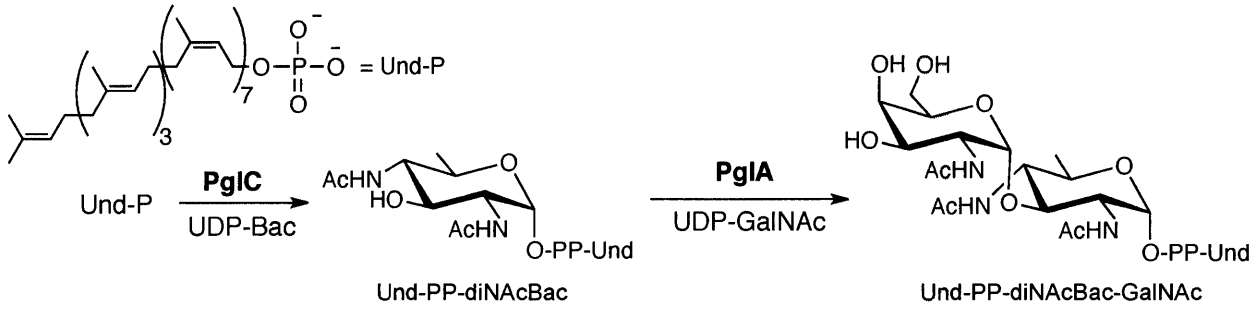


Figure 7-3. Phosphoglycosyl- and glycosyltransferase reactions carried out by PglC and PglA, respectively.

Results and discussion

Preliminary studies on the incorporation of undecaprenyl-linked substrates

All prior Nanodisc studies demonstrating functional reconstitution of integral membrane proteins have focused on enzymes that utilize soluble small molecule substrates.[12, 13] In contrast,

the glycosyltransferases involved in N-linked glycosylation require substrates that include a membrane-bound polyprenyl-phosphate carrier for glycan assembly. Thus, at the outset of our studies, it was unclear if it was feasible to reliably incorporate the polyprenol carrier into the phospholipid bilayer structure of a Nanodisc. However, in principle, it should be feasible to incorporate the polyprenyl-phosphate into phospholipids in the same way that they are incorporated in native membrane bilayers.

To detect incorporation of the undecaprenyl-linked substrate into Nanodiscs, it was necessary to prepare undecaprenyl-diphosphate (Und-PP) diNAcBac in a tritium-labeled form as described in Chapter 3. Und-PP-diNAcBac was initially chosen, since it is relatively easy to biosynthesize and is both the product of PglC and the substrate for PglA. Preparation of Und-PP-[^{3}H]diNAcBac was performed by reacting UDP-4-amino with [^{3}H]acetyl coenzyme A (AcCoA) and undecaprenyl phosphate (Und-P) in the presence of PglB(*Ng*) (Figure 7-4). This bifunctional enzyme first transfers an acetyl group from AcCoA to form UDP-[^{3}H]diNAcBac and then transfers phospho-[^{3}H]diNAcBac to form the tritiated product (Figure 7-4). The radiolabeled product can be purified by HPLC as previously described in Chapter 3.

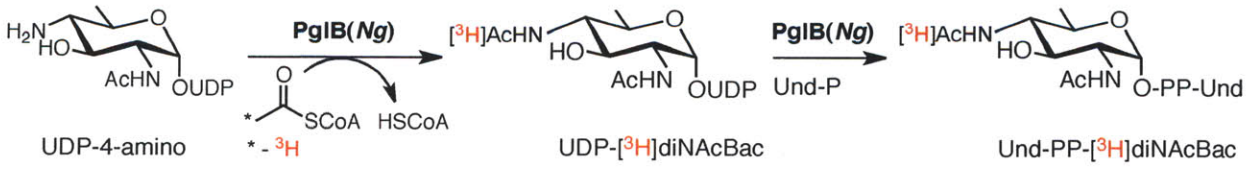


Figure 7-4. PglB(*Ng*) produces Und-PP-diNAcBac.
It is a useful chemoenzymatic tool for the one-step generation of the tritiated undecaprenyl product from the UDP-4-amino, [^{3}H]acetyl CoA, and Und-P substrates.

Initially, "unloaded" Nanodiscs were prepared without the integral membrane protein PglC to determine if undecaprenol derivatives would incorporate into discs in the absence of

integral membrane proteins (Figure 7-5). The Nanodiscs were synthesized with 0.5 mole percent Und-PP-[^{3}H]diNAcBac substituted for the native *E. coli* lipids. As previously described, when preparing PglC-containing Nanodiscs, the His_6-tag on the N-terminus of PglC was used to purify the discs (Figure 5-10). In the preparation of empty discs, a His_6-tagged version of MSPD1E3 allowed for Ni-NTA purification. Following detergent removal with Biobeads and Nanodiscs formation, the discs were purified with Ni-NTA chromatography, removing excess phospholipids and unincorporated undecaprenyl-linked substrate (Figure 7-5). The radioactivity present in the flow-through, wash, and elution fractions was quantified to determine the amount of Und-PP-[^{3}H]diNAcBac incorporation into Nanodiscs (Figure 7-6). In this experiment, most (86%) of the polyprenyl substrate was associated with MSP after Ni-NTA affinity chromatography suggesting that the substrate incorporated efficiently into Nanodiscs. At the concentrations used in this experiment, these results imply that there are one or two undecaprenyl-linked substrates per disc.

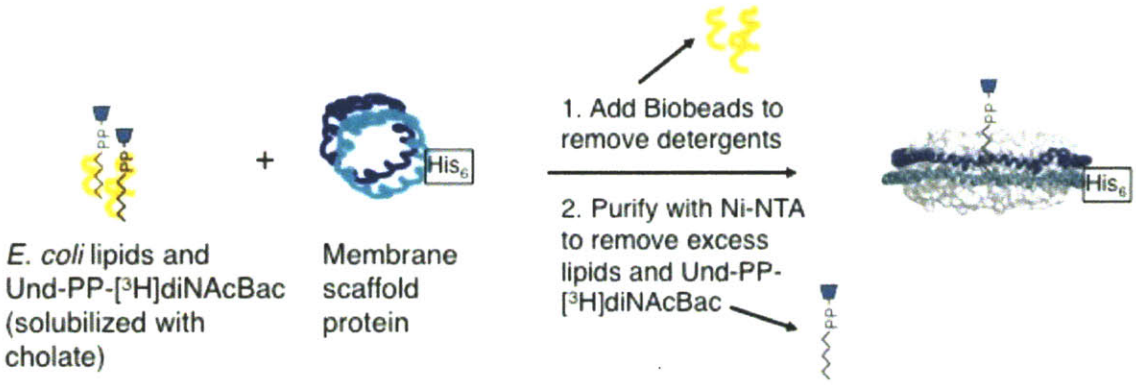


Figure 7-5. Preparation of "unloaded" Nanodiscs containing radiolabeled Und-PP-glycans.

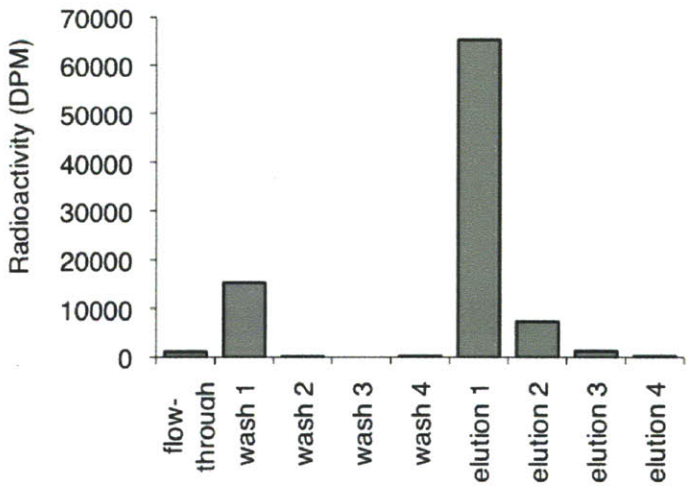


Figure 7-6. Incorporation of Und-PP-[^{3}H]diNAcBac into "unloaded" Nanodiscs.
The Nanodiscs were purified using Ni-NTA chromatography and the radioactivity in the fractions was determined by liquid scintillation counting. The specific activity of the Und-PP-[^{3}H]diNAcBac is 45 μCi/μmol such that 10,000 DPM is equivalent to 0.1 nmoles of tritiated compound.

To further probe the potential of Nanodiscs for our future experiments, it was important to incorporate both an enzyme and the undecaprenyl-linked substrate into the same disc. PglC-containing Nanodiscs were again prepared with 0.5 mole percent of the phospholipids replaced by Und-PP-[^{3}H]-diNAcBac. In PglC-containing preparations, the Nanodiscs were purified utilizing the unique His_6-tag on PglC to isolate the discs in which PglC was incorporated. Samples of the flow-through, wash, and elution fractions were measured for radioactive content (Figure 7-7). Under these conditions, approximately 56% of the Und-PP-[^{3}H]-diNAcBac was incorporated into the PglC-containing Nanodiscs. It is important to note that a MSP:PglC ratio of 4:1 was utilized in this experiment, instead of the typical 10:1, in order to increase the number of discs that would contain both PglC and Und-PP-[^{3}H]-diNAcBac.

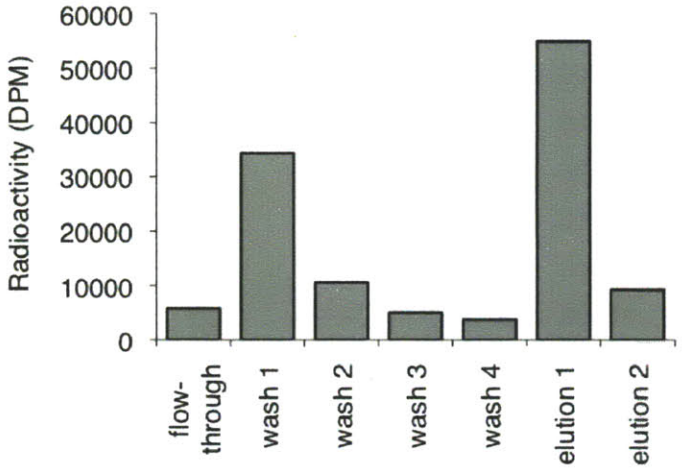


Figure 7-7. Incorporation of both Und-PP-[^{3}H]diNAcBac and PglC into Nanodiscs.
The specific activity of the Und-PP-[^{3}H]diNAcBac is 45 μCi/μmol such that 10,000 DPM is equivalent to 0.1 nmoles of tritiated compound.

Preliminary activity assays of the PglA glycosyltransferase

The experiments described above verified that Und-PP-diNAcBac could be efficiently incorporated into Nanodiscs in the absence and presence of PglC. Und-PP-diNAcBac is the product of the PglC reaction and the substrate for PglA, the next glycosyltransferase in the *C. jejuni* pathway. It is known that PglA readily reacts with Und-PP-diNAcBac in a detergent-micelle based assay,[1] and thus it was important to next evaluate whether PglA would react with Und-PP-diNAcBac integrated in "unloaded" or PglC-containing Nanodiscs.

PglA forms an aggregate when purified in the absence of detergent as determined by size exclusion chromatography, however, glycosyltransferase activity is still observable in detergent-based assays. Because addition of detergent-solubilized PglA would disassemble the Nanodiscs, the unsolubilized PglA aggregate was used in the following preliminary assays, in order to establish if a small amount of the protein would be available for incorporation into the Nanodiscs. PglA is also practical in this respect since it transfers GalNAc from UDP-GalNAc,

which can be obtained in a tritiated form and thus provides a useful label for monitoring enzyme activity. To perform the assay, Nanodiscs, with and without PglC, were incubated in the presence of PglA and UDP-GalNAc. The activity was measured using an organic extraction based assay monitoring transfer of radiolabeled [^{3}H]GalNAc from the aqueous soluble UDP-donor to the organic soluble Und-PP-diNAcBac (Figure 7-8).

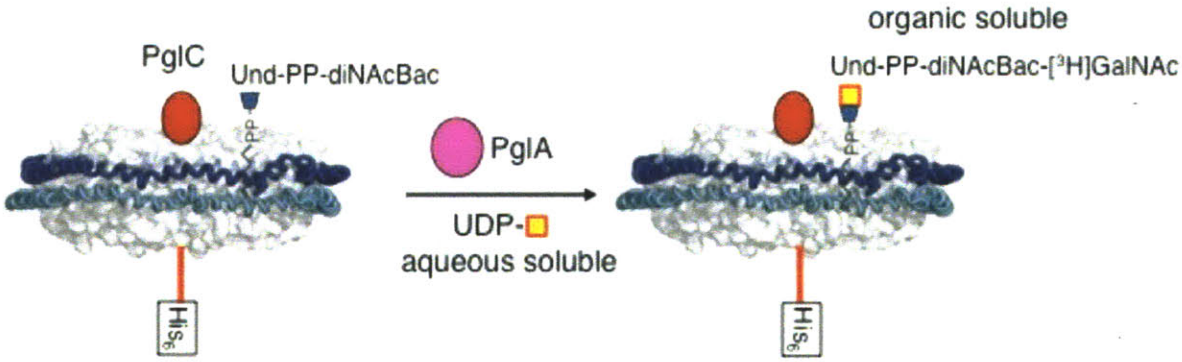


Figure 7-8. Activity assay of PglA in Nanodiscs.

As a comparison, a detergent-micelle based assay was performed with equivalent amounts of substrates and enzymes. All three reactions showed PglA glycosyltransferase activity, however, the detergent-micelle based assay reached completion within minutes, whereas the two Nanodiscs-based reactions were slower and took hours to reach similar levels of production conversion (Figure 7-9). Furthermore, PglA activity was roughly equivalent in the two Nanodisc samples, implying that the presence or absence of PglC did not affect activity. In the absence of PglA, no [^{3}H]GalNAc transfer was observed (data not shown).

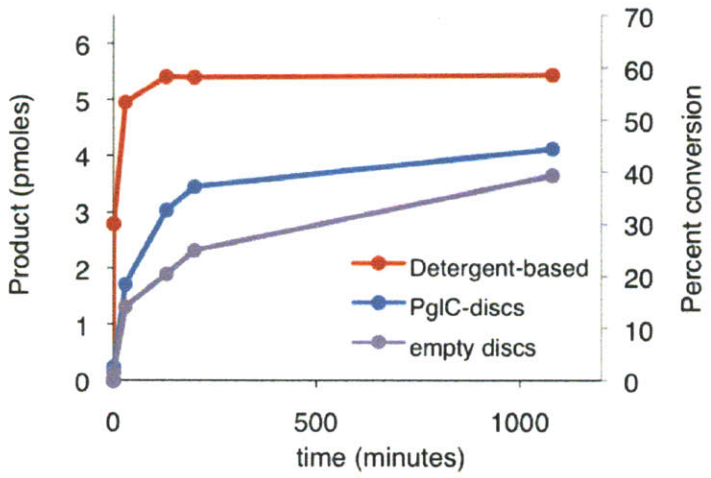


Figure 7-9. PglA activity in detergent-based micelles versus Nanodiscs.
In the detergent-based assay (red), the Und-PP-diNAcBac was detergent solubilized, whereas in the disc assays (blue and purple), the Und-PP-diNAcBac was incorporated in Nanodiscs. In addition the activity of PglA in the presence of PglC-containing discs was compared to "unloaded" discs, and both showed similar amounts of activity.

The different activities of the glycosyltransferases in Nanodisc samples as compared to detergent-micelle based assays cannot be explained by differences in the enzyme and substrate concentrations or buffer conditions, and thus, can be attributed to the insertion of Und-PP-diNAcBac in Nanodiscs. Because PglA forms an aggregate in non-detergent aqueous solutions, this limits the ability of a single PglA active site to interact with the Nanodisc membrane interface and probably contributes to the diminished activity observed in the assays.

Despite these limitations, the observed activity appears to be a real and localized in the Nanodisc. Re-purification of Nanodiscs after reaction with PglA provides supporting evidence that PglA is transferring GalNAc to Nanodisc-bound Und-PP-diNAcBac. The PglA reaction was carried out as described above and after overnight incubation, organic-aqueous extraction of a small aliquot was used to determine that approximately 50% of the tritiated GalNAc was transferred to undecaprenyl-linked substrate. The remainder of the Nanodisc sample was then re-purified using Ni-NTA resin. Excess UDP-[^{3}H]GalNAc was washed away from the resin, the

PglC-containing Nanodiscs were eluted, and the amount of [^{3}H]GalNAc incorporation was detected by scintillation counting (Figure 7-10). Quantification of conversion by this method determined that 32% of the radioactive GalNAc sugar had transferred to Nanodiscs. This implies that ~60% of the PglA product produced in the reaction is still associated with PglC-Nanodiscs. While the conversion percents are somewhat imprecise, this may be caused by poor retention of the Nanodisc samples to Ni-NTA resin in the presence of the aggregate PglA enzyme. Nonetheless, this result further demonstrates that the PglA product can be formed while the Und-PP-diNAcBac is integrated in the phospholipid bilayer of a Nanodisc.

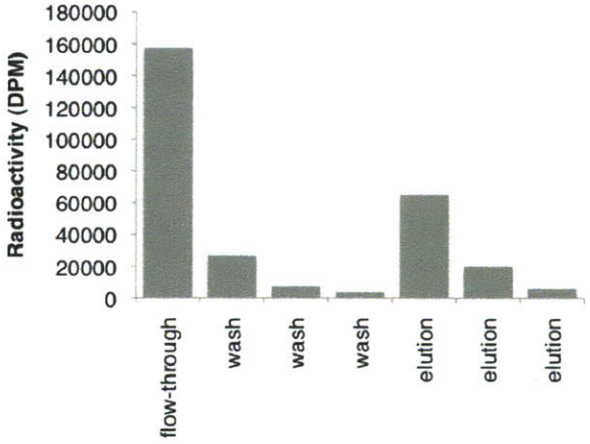


Figure 7-10. Ni-NTA purification of Nanodiscs reacted with PglA.
Incorporation of tritium-labeled GalNAc into disc-bound Und-PP-diNAcBac as indicated by radioactivity associated with the Nanodiscs, which are present in the elution fractions. Tritium-labeled UDP-[^{3}H]GalNAc with a specific activity of 20 Ci/mmol was used to track the glycosyltransferase reaction. In the chart, 20,000 DPM represents 0.5 pmoles of product.

Preliminary examination of PglC function in Nanodiscs

Further functional studies of N-linked protein glycosylation in Nanodiscs required the functional reconstitution of the integral membrane protein PglC within a Nanodisc. Although the PglA assay described above was not ideal due to the low activity and aggregated state of PglA, it did provide a straightforward read-out for measuring PglC activity. Assays described above

demonstrated that PglA showed activity in the presence of Nanodiscs containing PglC and Und-PP-diNAcBac (Figure 7-8), which validates this approach to assaying PglC activity. The two substrates required for the PglC reaction are undecaprenyl-phosphate (Und-P) and UDP-diNAcBac. At the time of these preliminary studies, neither of these compounds was available in a radiolabeled form. The experiments described in the next section address this issue, but for the initial PglC activity assay described here, PglC-Nanodiscs were prepared with 0.5 mole percent of Und-P substituted for the *E. coli* lipids. The percent of Und-P incorporation could not be quantified, because the Und-P was unlabeled. However, an assay for PglC activity would determine if any Und-P was present and more importantly, if PglC was functionally reconstituted in Nanodiscs and able to carry out the phospho-glycosyltransferase reaction.

Nanodiscs containing both PglC and Und-P were incubated with UDP-diNAcBac, tritiated UDP-[^{3}H]GalNAc, and PglA (Figure 7-11). The results of this assay, shown in Figure 7-12, demonstrate that PglC was functionally reconstituted in Nanodisc and is able to facilitate transfer of phospho-diNAcBac to Und-P to provide the product for PglA activity. The negative control, which omitted UDP-diNAcBac, verified that PglA was not active in the absence of the PglC product. This was an exciting and promising result and suggested that integration of glycosyltransferases into Nanodiscs is a viable method for the study of enzyme activity in the N-linked protein glycosylation pathway.

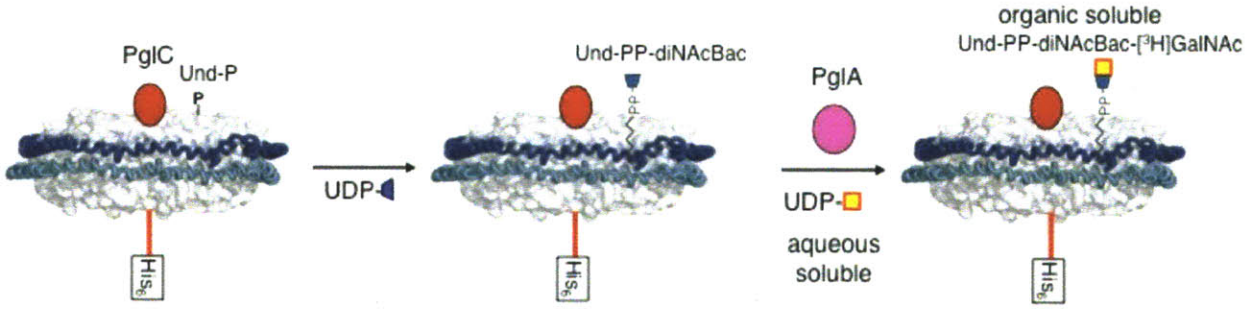


Figure 7-11. Coupled enzyme assay of PglC in Nanodiscs.

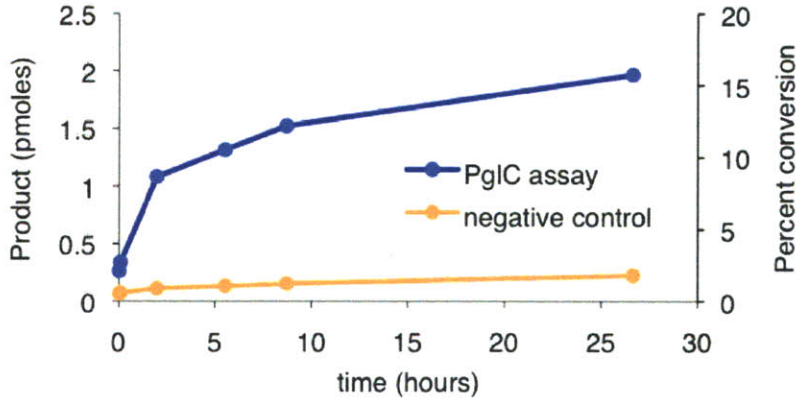


Figure 7-12. Preliminary assay of PglC in Nanodiscs coupled to PglA activity.
In the negative control, the PglC substrate, UDP-diNAcBac, was omitted from the reaction. For the conversion amounts shown in the graph, it was assumed that Und-P incorporated at levels similar to Und-PP-diNAcBac.

Nanodiscs prepared with Und-^{33}P

The preliminary experiments described above were performed at the outset of our Nanodiscs studies and convinced us that activity studies were feasible. More recent work has focused on improving these studies through the generation of quantifiable, differentially labeled Und-P. The *S. mutans* kinase described in Chapter 2, provided a valuable reagent for the chemoenzymatic synthesis of Und-[^{33}P]P from undecaprenol and γ-[^{33}P]ATP. The half-life of ^{33}P

is 25.3 days, almost twice as long as the half-life of ^{32}P (14.3 days), and thus the ^{33}P isotope was chosen as being more convenient for the following experiments. Und-[^{33}P]P expands the Nanodisc assay capabilities by providing a method for quantifying the amount of undecaprenyl-phosphate incorporation. In addition, Und-[^{33}P]P provides an orthogonal radiolabel that can be used in concert with the tritium-based assays to allow for precise tracking of the Und-PP-linked substrates and products of the glycosyltransferase reactions.

The results in Chapter 5 and 6 suggested that co-incorporation of PglA and PglC into Nanodiscs is a viable method for co-localization of these proteins into the model membrane system provided by the Nanodiscs. Inclusion of monodispersed, detergent-solubilized PglA in the Nanodisc preparation mixture provided an alternative to the use of aggregate PglA as described in the preliminary assays. Assaying PglA in a native form was preferable and would provide functional results that better represented how enzyme activity occurs at the membrane interface. Nanodiscs were prepared with co-incorporation of PglC, PglA, and Und-[^{33}P]P with ratios of MSP:PglC:PglA of 10:1:1 to ensure that each disc would contain at most one copy of each enzyme (Figure 7-14). Nanodiscs were purified by the His_6-tagged PglC protein as usual and 7% of Und-[^{33}P]P was incorporated into PglC-containing Nanodiscs (Figure 7-14). Given the fact that the PglC-containing Nanodiscs comprise only 20% of all of the discs formed, this value suggests that at least 35% of Und-[^{33}P]P was incorporated into Nanodiscs overall. As shown in Figure 7-13, most discs would be expected to contain one or two copies of undecaprenyl-phosphate. In addition, SDS-PAGE analysis confirmed the presence of MSP, PglC, and PglA (Figure 7-14).

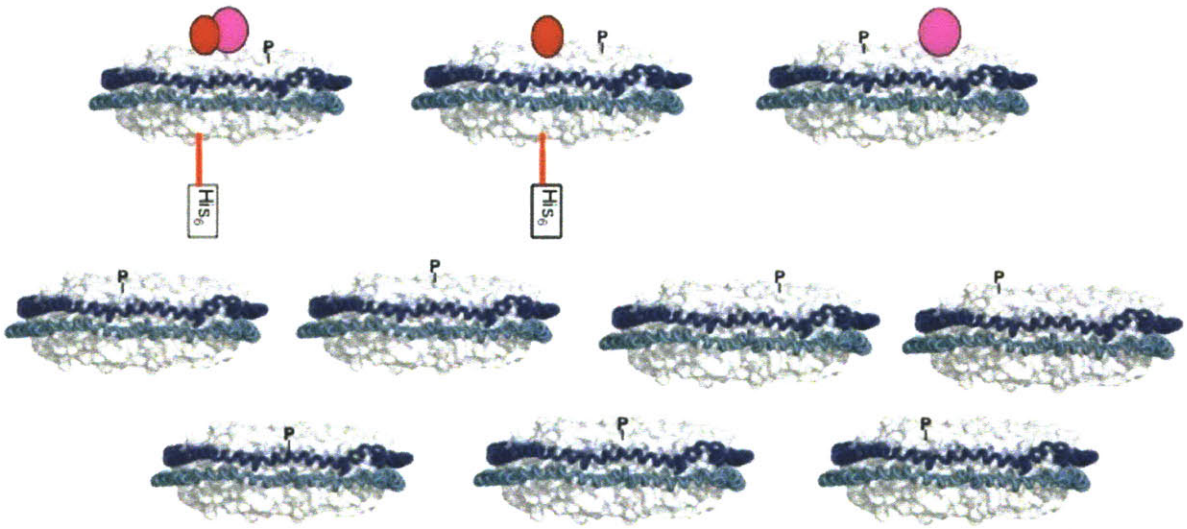


Figure 7-13. A schematic of a Nanodisc preparation with PglC, PglA, and Und-[^{33}P]P.
In a preparation mixture of MSP:PglC:PglA of 10:1:1, theoretically, 20% of the discs will contain PglC and 20% will contain PglA. As demonstrated in Chapter 5, PglA co-localizes more frequently with PglC-containing discs. Und-[^{33}P]P would be expected to incorporate into all of the discs.

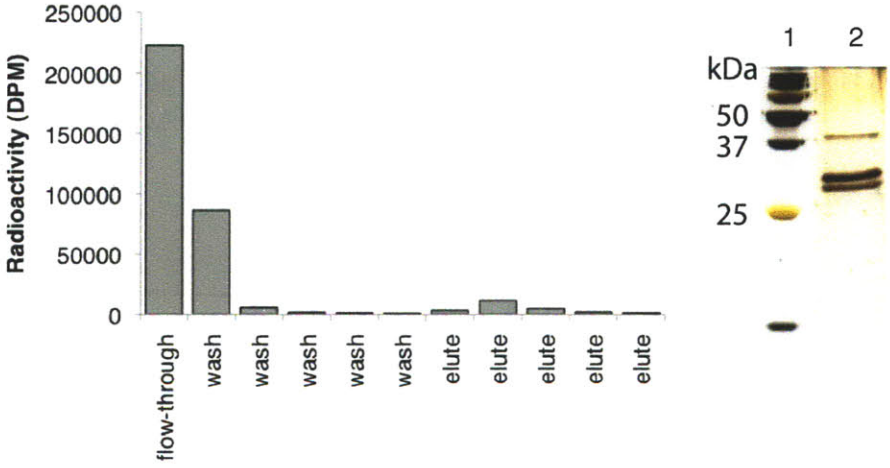


Figure 7-14. Analysis of Nanodiscs containing PglC, PglA, and Und-[^{33}P]P.
Left, Incorporation of Und-[^{33}P]P into PglC and PglA-containing Nanodiscs. The presence of radioactivity in the elution fractions suggests that Und-P has been incorporated into Nanodiscs. The specific activity of the Und-[^{33}P]P is 90 μCi/μmol; 50,000 DPM is equivalent to 0.25 nmoles. *Right*, SDS-PAGE representation with silver staining. Lane 1, molecular weight standards; lane 2, elution fraction containing Nanodiscs with PglC and PglA incorporated.

The Nanodiscs containing PglC, PglA and Und-[^{33}P]P were assayed by adding the two soluble sugar substrates, UDP-diNAcBac and UDP-[^{3}H]GalNAc. Aliquots from the reaction

were quenched at several time-points and organic-aqueous extractions were performed to assess transfer of the tritium-labeled sugar to the organic soluble Und-[^{33}P]P (Figure 7-15). The assay data in Figure 7-16 shows linear production of the dual-labeled Und-[^{33}P]PP-diNAcBac-[^{3}H]GalNAc product. This evidence suggests that glycosyltransferase activity can be monitored in a Nanodisc on the timescale of the assay, which is fundamental for further explorations of the pathway activity.

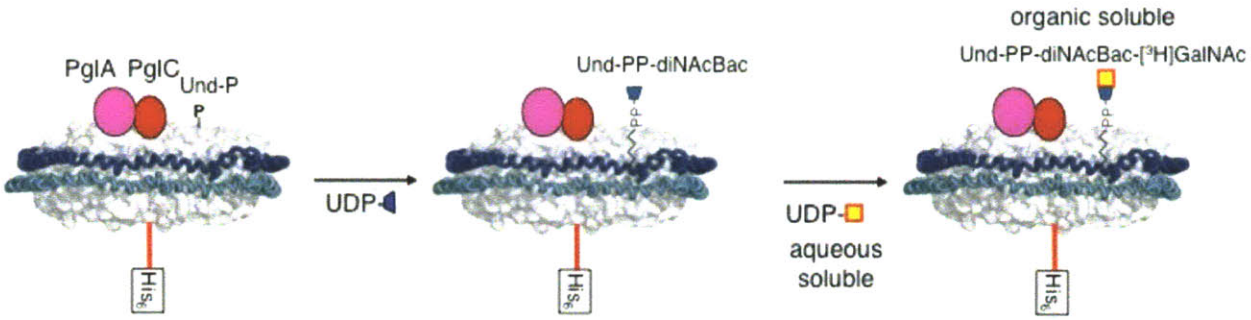


Figure 7-15. Assay of Nanodiscs pre-loaded with PglC, PglA, and Und-[^{33}P]P.
Activity is measured by incorporation of radiolabeled [^{3}H]-GalNAc into Und-PP-diNAcBac integrated into Nanodiscs.

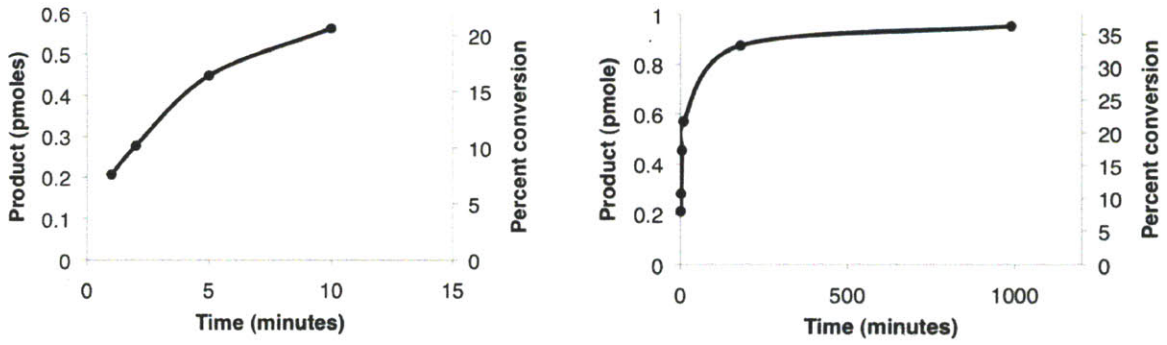


Figure 7-16. Activity of Nanodiscs containing PglC, PglA, and Und-P.
Left, The enzyme activity during the first ten minutes of the reaction. *Right*, The same reaction is shown on a longer timescale.

To further verify the Nanodisc-bound activity, another informative experiment was performed to examine if the glycosyltransferase activity was dependent on Nanodisc

concentration. If each Nanodisc is a stable entity containing one copy of PglC, one copy of PglA, and Und-P, then the activity should not depend on disc concentration, since the local concentration of the enzymes is constant. An equivalent molar amount of Nanodiscs was assayed at three different dilutions, while the concentrations of the soluble UDP-sugars were kept constant. As observed in Figure 7-17, the reaction rates for the three dilutions are nearly identical, which is consistent with the localization of PglC, PglA and Und-P in the same Nanodisc. This result supports the proposal that PglC and PglA are functionally reconstituted in a stable manner at the Nanodisc membrane interface.

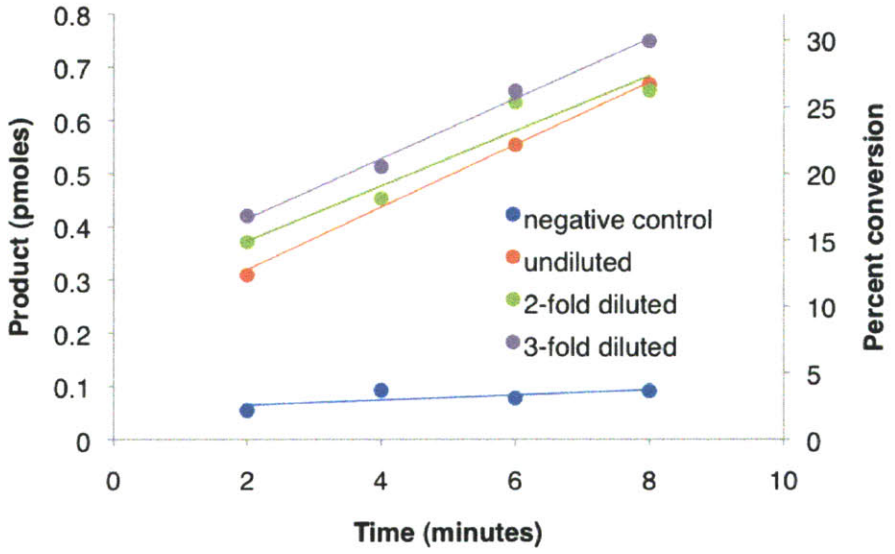


Figure 7-17. Assay of Nanodiscs containing PglC, PglA and Und-P.
The Nanodiscs were assayed at three different concentrations, but the concentration of the soluble sugar was kept constant. The activity rates for the three reactions are roughly equivalent suggesting that the PglC, PglA, and Und-P are co-localized in a Nanodisc.

Conclusions

To validate the Nanodisc model membrane system as a tool for the study of N-linked protein glycosylation in a membrane environment, undecaprenyl-based substrate incorporation

was required. The studies in this chapter provide early evidence that both undecaprenyl-phosphate and Und-PP-diNAcBac can be integrated into the lipid bilayer of Nanodiscs in the presence and absence of the integral membrane protein PglC. Preliminary assays, in which aggregated PglA was incubated with Nanodiscs, showed markedly reduced activity relative to detergent-based assays. The aggregated state of the PglA likely limits the number of PglA proteins able to incorporate into Nanodiscs, and this may be a major factor contributing to the diminished activity in the initial experiments. More recent activity studies have focused on improving assay methodology by co-incorporating PglC, PglA, and Und-[^{33}P]P into Nanodiscs, and thus avoiding the use of aggregated PglA. Activity assays of the discs containing PglC, PglA, and Und-[^{33}P]P result in more rapid conversion relative to the previous results, and generation of a dual-labeled Und-[^{33}P]PP-diNAcBac-[^{3}H]GalNAc product. In addition, identical samples of Nanodiscs assayed at different concentrations have similar rates of activity, which further validates that PglA and PglC are uniquely localized in the same Nanodisc. These experiments provide the basis for future studies of how the *C. jejuni* glycosyltransferases interact with undecaprenyl-phosphate-linked substrates. In addition, it is demonstrated that Nanodiscs provide a promising platform for biophysical explorations of how polyprenyl-phosphate linked substrates affect lipid bilayers in the presence and absence of integral membrane proteins.

Acknowledgements

I gratefully acknowledge my colleague, Marcie Jaffee, who critically commented on this chapter in a most helpful manner.

Experimental Protocols

Materials

Und-[^{33}P]P was biosynthesized on a 50 nmole scale as follows. A reaction was prepared with 3% DMSO, 1% Triton X-100, 30 mM Tris-Acetate (pH 8.0), 50 mM $MgCl_2$, and 50 µL of *S. mutans* kinase CEF in a 100 µL total volume. The reaction was quenched into organic solvent and purified using a normal phase Varian Microsorb HPLC column and 4:1 CH_3Cl:MeOH (solvent A) with 10:10:3 CH_3Cl:MeOH:2M ammonium acetate. A flow rate of 1 mL/min was used and the gradient was as follows: 0 to 3 min, 0% B; 3 to 5 min, 0 to 22% B; 5 to 45 min, 22 to 30 % B; 45 to 50 min, 30 to 100% B; 50 to 55 min, 100% B. The Und-P elutes at 21-23 minutes in this gradient. Und-PP-[^{3}H]diNAcBac was also biosynthesized on a 50 nmole scale as depicted in Figure 7-4 and purified using the same solvents with the normal phase-HPLC gradient described in Chapter 3 and the following reference.[14] Following HPLC purification, aliquots containing 2-5 nmoles were prepared, dried in the SpeediVac, and stored at -80 °C.

Incorporation of undecaprenyl-linked substrates into Nanodiscs

Nanodiscs were prepared as described in Chapter 5 with the exception that 0.5 mole percent of Und-[^{33}P]P or Und-PP-[^{3}H]diNAcBac was incorporated in the preparation mixture. A typical reaction included 2 nmoles of the undecaprenyl-linked product in a 300 µL preparation mixture. Cholate detergent from 100 mM cholate-containing buffer is normally added during Nanodisc preparations. When preparing Nanodiscs containing undecaprenyl-linked substrates, the cholate buffer was first added to the dried undecaprenyl-linked substrate and the tube was vigorously vortexed and sonicated to solubilize the polyprenol. Then the other Nanodisc components were added as appropriate, including lipids, MSP, PglC and PglA, and Nanodisc

preparation continued as described in Chapter 5. Purification of the discs by Ni-NTA affinity chromatography proceeded as described in Chapter 5. Prior to assaying the activity of the Nanodisc samples, the elution fractions containing incorporated substrate were pooled and dialyzed overnight into PBS (pH 7.0) at room temperature.

For preparations containing radiolabeled substrates, the Nanodiscs were analyzed by liquid scintillation counting. To determine the level of incorporation, the radioactivity present in 10% of the flow-through, wash, and elution samples was quantified. These data were then used to calculate the values depicted in Figures 7-5 and 7-11.

For the preliminary experiments, in which Und-PP-[^{3}H]diNAcBac was incorporated into the Nanodiscs with PglC, the Nanodisc preparation mixture contained a MSP:PglC ratio of 4;1, instead of 10:1. This increased the probability that a single disc will contain more than one copy of PglC (Table 5-4), but it also increased the yield of PglC-containing Nanodiscs, making it easier to detect co-incorporation of PglC and Und-PP-diNAcBac. For the later experiments in which PglC, PglA, and Und-P were co-incorporated, the MSP:PglC:PglA ratio used was 10:1:1 to ensure that each copy only contained one copy of PglC and one copy of PglA, which was important for functional analysis studies.

Activity analysis of Nanodisc samples

For the activity assays, the amount of radiolabeled undecaprenyl-linked substrate incorporated into the Nanodiscs was used to quantify the amount of discs added to the assay. For the preliminary PglA assays, Nanodiscs, containing 0.4 μM Und-PP-[^{3}H]diNAcBac, were incubated with 1 μM PglA, 0.1 μM UDP-[^{3}H]GalNAc, and 50 mM $MgCl_2$ with an assay volume of 100 μL. PglA was purified in the absence of detergent as described previously.[1] The assay

setup was identical for the preliminary PglC-PglA coupled assay, except that 2.5 μM UDP-diNAcBac was also added to the reaction mixture.

For the later assays, PglA was incorporated into the Nanodiscs as described in Chapter 5. For these assays, Nanodiscs, containing 0.05 μM Und-[^{33}P]P, were incubated with 40 μM UDP-diNAcBac, 0.1 μM UDP-[^{3}H]GalNAc, and 40 mM $MgCl_2$ in a 100 μL total volume. For the dilution assays, Nanodiscs were assayed at 0.05 μM Und-[^{33}P]P, 0.025 μM Und-[^{33}P]P, or 0.017 μM Und-[^{33}P]P, while the concentrations of the other substrates remained constant (40 μM UDP-diNAcBac, 0.1 μM UDP-[^{3}H]GalNAc, and 40 mM $MgCl_2$).

For the quenched time-point assays, aliquots of 12 μL were taken at various timepoints and quenched into 1 mL of 2:1 chloroform:methanol. The organic layer was washed three times with 400 μL of an aqueous solvent, containing 240 mL of methanol, 15 mL of chloroform, and 235 mL of water with 1.83 g of potassium chloride. The organic layer was dried and re-suspended in 200 μL Solvable. Scintillation fluid (Opti-Fluor, Perkin-Elmer or Ecolite, MP Biomedicals) was added to the organic and aqueous layers, respectively, and the radioactivity present in the samples was determined by scintillation counting.

Alternatively, the reacted Nanodiscs were isolated by Ni-NTA chromatography to determine if the radiolabel was incorporated into the disc. After incubation, the reaction mixture was incubated with 15-20 μL of Ni-NTA resin for 1 hour. The resin was washed 3-5 times with 500 μL of PBS (pH 7.0) supplemented with 30 mM imidazole. The Nanodiscs were eluted in three fractions of 250 μL with PBS (pH 7.0) containing 300 mM imidazole. A portion (one-tenth) of each flow-through, wash, and elution samples was mixed with 200 μL DMSO and 5 mL of Ecolite and the radioactivity present was determined by scintillation counting.

References

1. Glover, K. J.; Weerapana, E.; Imperiali, B., *In vitro* assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. U.S.A.* **2005,** 102, (40), 14255-9.
2. Swiezewska, E.; Sasak, W.; Mankowski, T.; Jankowski, W.; Vogtman, T.; Krajewska, I.; Hertel, J.; Skoczylas, E.; Chojnacki, T., The search for plant polyprenols. *Acta Biochim. Pol.* **1994,** 41, (3), 221-60.
3. Zhou, G. P.; Troy, F. A., 2nd, NMR study of the preferred membrane orientation of polyisoprenols (dolichol) and the impact of their complex with polyisoprenyl recognition sequence peptides on membrane structure. *Glycobiology* **2005,** 15, (4), 347-59.
4. Zhou, G. P.; Troy, F. A., 2nd, Characterization by NMR and molecular modeling of the binding of polyisoprenols and polyisoprenyl recognition sequence peptides: 3D structure of the complexes reveals sites of specific interactions. *Glycobiology* **2003,** 13, (2), 51-71.
5. Murgolo, N. J.; Patel, A.; Stivala, S. S.; Wong, T. K., The conformation of dolichol. *Biochemistry* **1989,** 28, (1), 253-60.
6. Knudsen, M. J.; Troy, F. A., Nuclear magnetic resonance studies of polyisoprenols in model membranes. *Chem Phys Lipids* **1989,** 51, (3-4), 205-12.
7. de Ropp, J. S.; Troy, F. A., 2H NMR investigation of the organization and dynamics of polyisoprenols in membranes. *J Biol Chem* **1985,** 260, (29), 15669-74.
8. McCloskey, M. A.; Troy, F. A., Paramagnetic isoprenoid carrier lipids. 2. Dispersion and dynamics in lipid membranes. *Biochemistry* **1980,** 19, (10), 2061-6.
9. Kelly, J.; Jarrell, H.; Millar, L.; Tessier, L.; Fiori, L. M.; Lau, P. C.; Allan, B.; Szymanski, C. M., Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.* **2006,** 188, (7), 2427-34.
10. Larsen, J. C.; Szymanski, C.; Guerry, P., N-linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. *J Bacteriol* **2004,** 186, (19), 6508-14.
11. Karlyshev, A. V.; Everest, P.; Linton, D.; Cawthraw, S.; Newell, D. G.; Wren, B. W., The *Campylobacter jejuni* general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **2004,** 150, (Pt 6), 1957-64.
12. Ritchie, T. K.; Grinkova, Y. V.; Bayburt, T. H.; Denisov, I. G.; Zolnerciks, J. K.; Atkins, W. M.; Sligar, S. G., Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **2009,** 464, 211-31.
13. Nath, A.; Atkins, W. M.; Sligar, S. G., Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins. *Biochemistry* **2007,** 46, (8), 2059-69.
14. Troutman, J. M.; Imperiali, B., *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* **2009,** 48, (12), 2807-16.

Appendix

Attempts towards the purification of PglC for X-ray crystallography

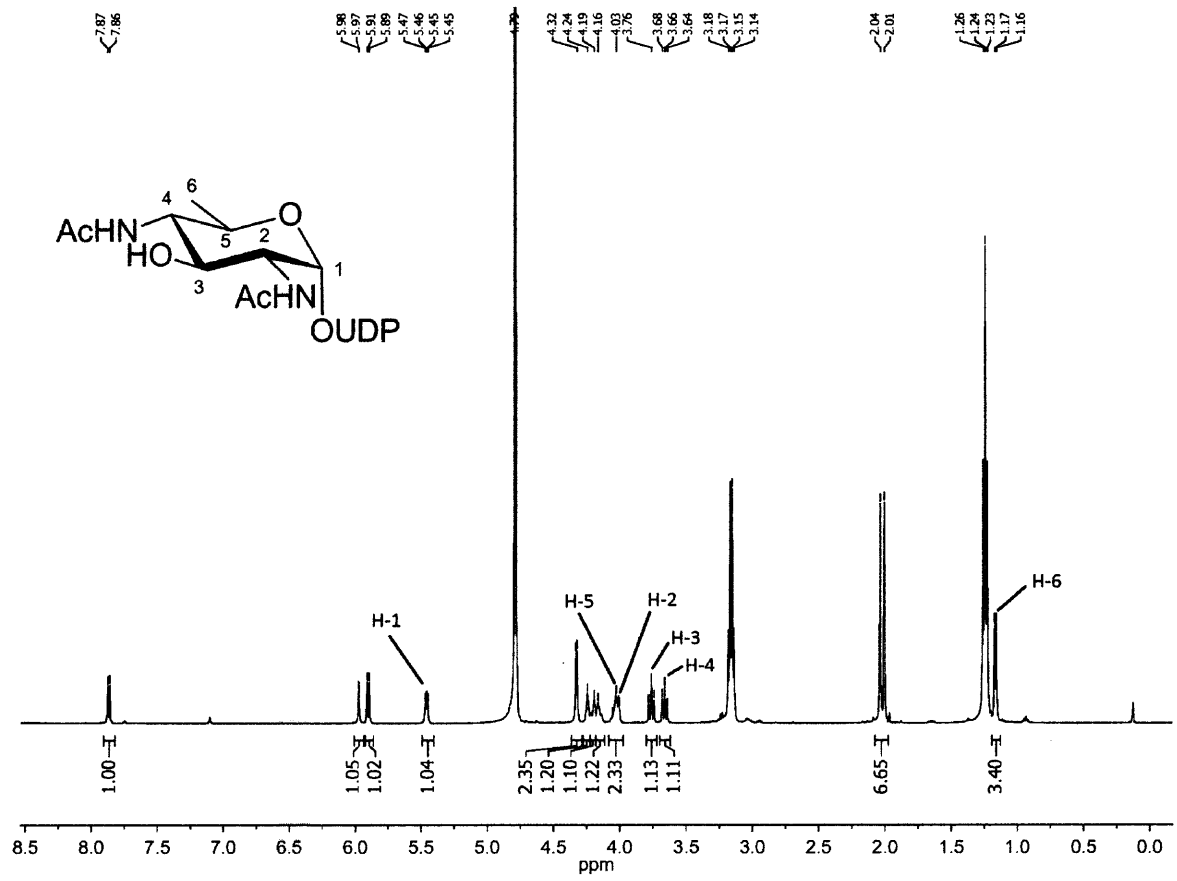
Date	Detergent (solubilization %, elution %)	[Total protein] during solubilization	Yield	Comments
10-31-08	Triton X-100 (1%, 0.1%)	3-5 mg/ml	1.5 mg/L	Works well
3-30-09	DDM (1%, 0.03%)	1-2 mg/ml	0.3 mg/L	Lysing was poor – malfuntioning sonicator
4-30-09	DDM (1%, 0.03%)	1-2 mg/ml	0.1 mg/L	" "
5-7-09	DDM (1%, 0.02%)	1-2 mg/ml	1 mg/L	None
5-25-09	LDAO (1%, 0.03%)	1-2 mg/ml	Very low	??
6-9-09	LDAO (2.5%, 0.03%)	1-2 mg/ml	Very low	Binding to Ni-NTA
6-16-09	CHAPSO (2.5%, 1%)*	3-5 mg/ml	1.5 mg/L	50% monodispersed
7-8-09	CHAPSO (2.5%, 1%)*	3-5 mg/ml	1.5 mg/L	All aggregate in SEC
7-28-09	CHAPSO (1%, 1%)*	3-5 mg/ml	1.5 mg/L	Joanne Yeh protocol with extra high salt wash[1]
8-27-09	CHAPSO (1%, 1%)*	7-9 mg/ml	1.5 mg/L	70% monodispersed by SEC (later shown to be <5% of yield after Ni-NTA)
9-14-09	CHAPSO (1%, 1%)*	3-5 mg/ml	1.5 mg/L	Did not pass through 100kDa filter
9-23-09	LDAO (0.05%, 0.05%)	3-5 mg/ml	Very low	Does not bind Ni-NTA resin
10-1-09	β-OG (1.1%, 1.1%)*	3-5 mg/ml	--	β-OG similar to CHAPSO
FLAG-PglC 11-4-09	LDAO (0.05%, 0.05%)	3-5 mg/ml	Low	Poor binding to resin
12-14-09	LDAO (0.05%, 0.05%)	3-5 mg/ml	Low	Used TALON resin – poor binding

1-5-09	DDM (0.03%), LDAO (0.05%)	3-5 mg/ml	0.1 mg/L	Poor solubilization
1-11-09**	DDM (0.03%), LDAO (0.05%)	3-5 mg/ml	Low	Used French press to lyse
1-11-09**	DDM (0.03%), LDAO (0.05%)	3-5 mg/ml	Low	Solubilized with overnight incubation
1-26-10	DDM (1%, 0.03%)	3-5 mg/ml	1.5 mg/L	Stroud Protocol with higher detergent percentage,[2] great solubilization
2-15-10**	DDM (1%) β-OG (1.4%)	3-5 mg/ml	1.5 mg/L	With and without PglD, PglC precipitated after elution

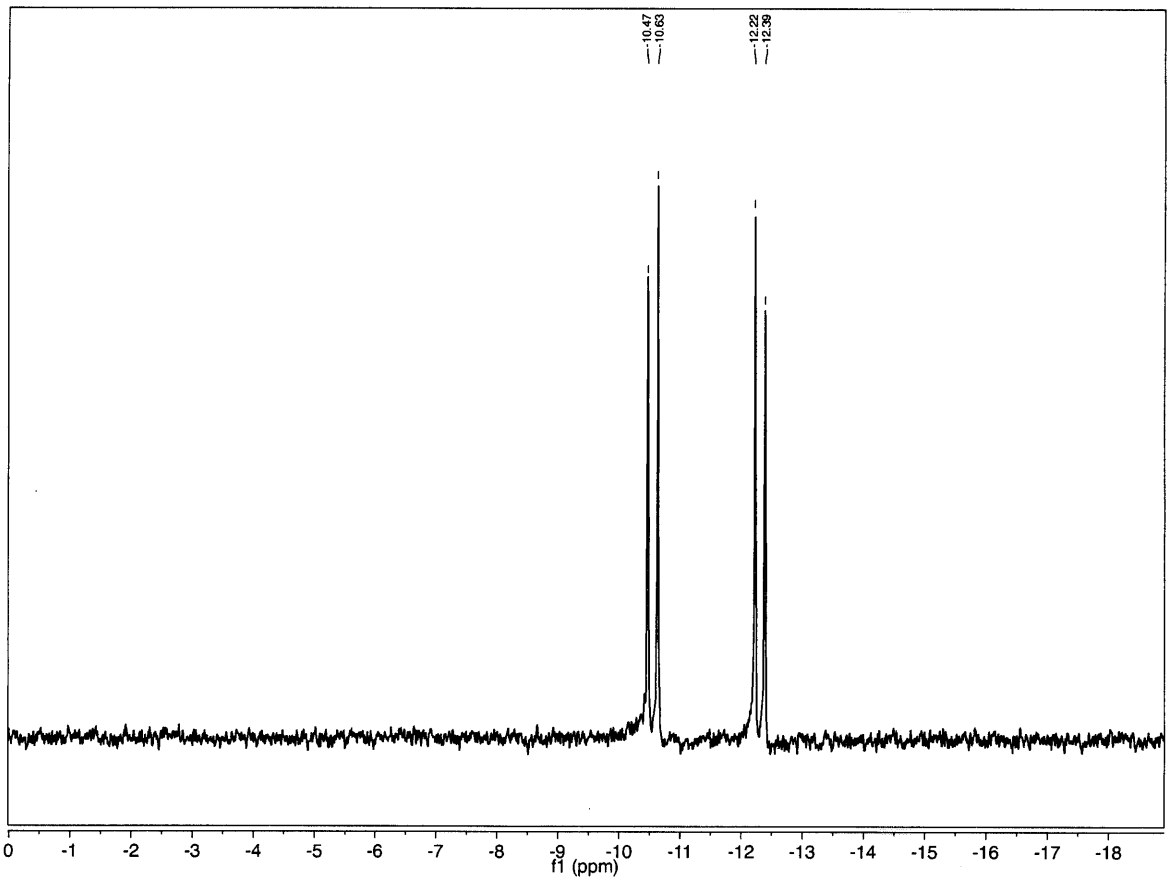
*After solubilization with CHAPSO or β -OG, the protein (after the final high speed spin) had a light yellow color, whereas solubilization with LDAO, Triton X-100, or DDM resulted in a colorless solution. In addition, SDS-PAGE analysis of PglC purified with CHAPSO or β-OG disclosed the presence of a high molecular weight band that did not enter the gel and was not present with other detergents.
** Performed with Vinita Lukose

References

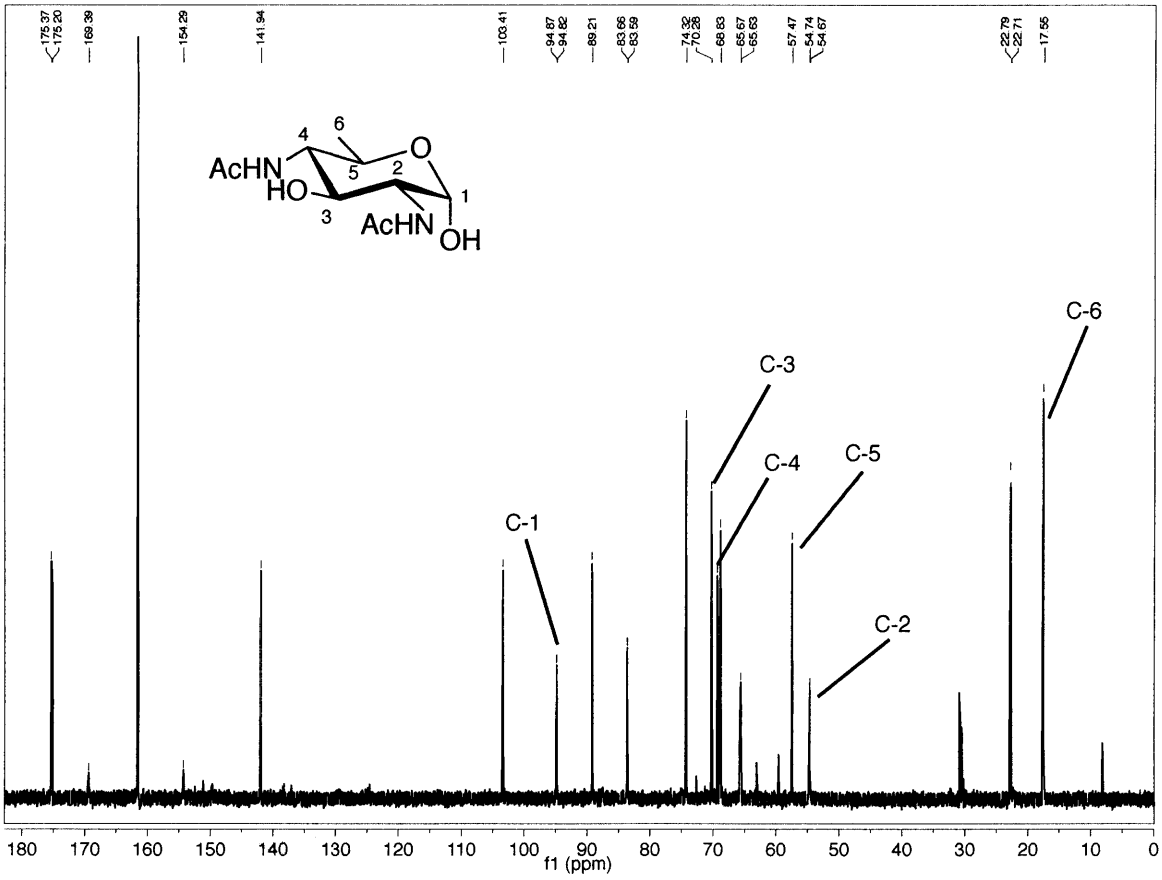
1. Yeh, J. I.; Du, S.; Tortajada, A.; Paulo, J.; Zhang, S., Peptergents: peptide detergents that improve stability and functionality of a membrane protein, glycerol-3-phosphate dehydrogenase. *Biochemistry* **2005,** 44, (51), 16912-9.
2. Newby, Z. E.; O'Connell, J. D., 3rd; Gruswitz, F.; Hays, F. A.; Harries, W. E.; Harwood, I. M.; Ho, J. D.; Lee, J. K.; Savage, D. F.; Miercke, L. J.; Stroud, R. M., A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat Protoc* **2009,** 4, (5), 619-37.



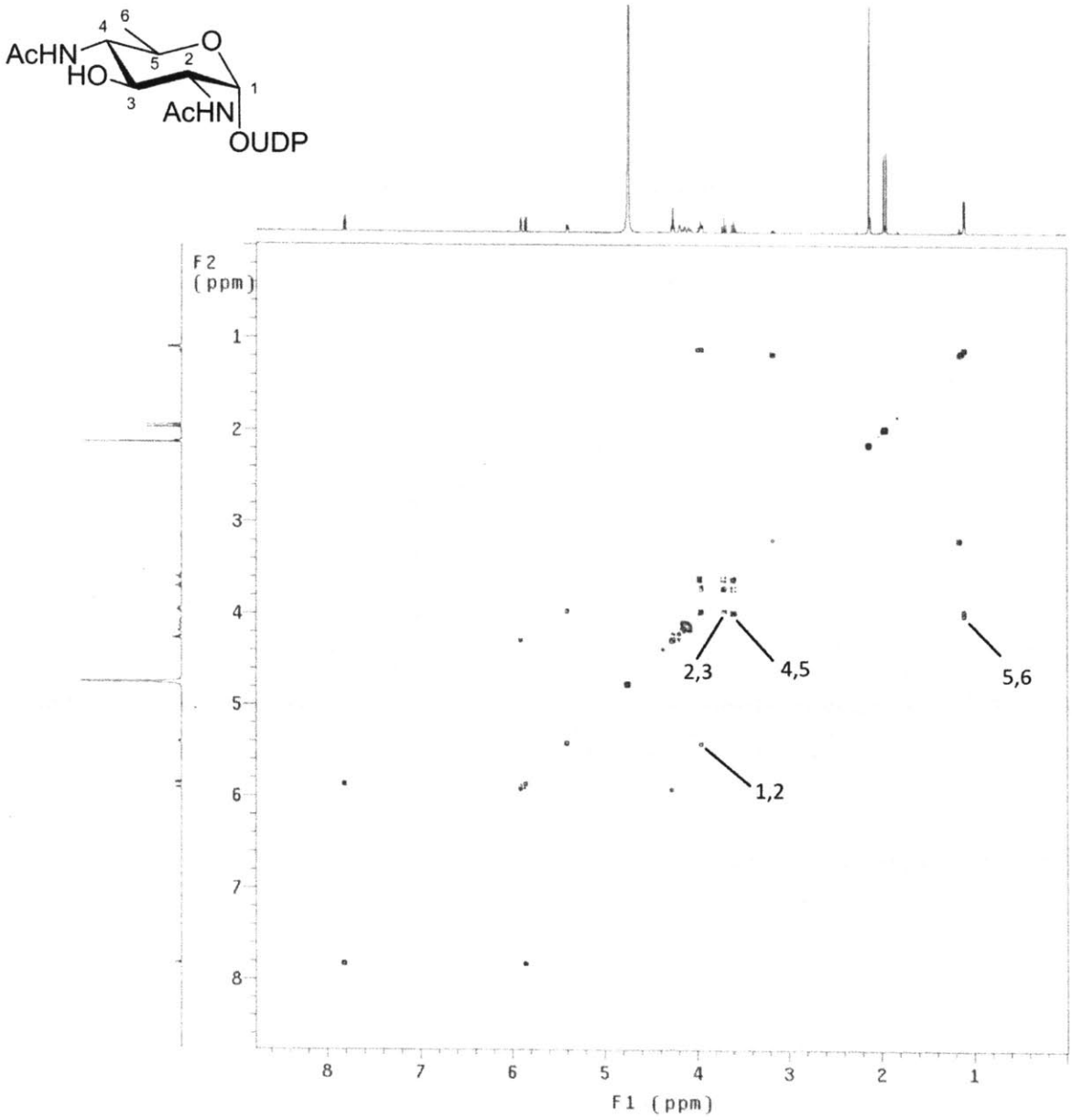
^{1}H NMR spectrum of UDP-diNAcBac.



^{31}P NMR spectrum of UDP-diNAcBac.



^{13}C NMR spectrum of UDP-diNAcBac.



^{1}H-^{1}H COSY NMR spectrum of UDP-diNAcBac.