

**STRUCTURAL AND FUNCTIONAL STUDIES OF ENZYMES INVOLVED
IN UNUSUAL SUGAR BIOSYNTHESIS IN PATHOGENIC BACTERIA**

by

Haley A. Brown

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN – MADISON

2018

Date of final oral examination: 8/10/2018

The dissertation is approved by the following members of the Final Oral Committee:

Hazel M. Holden, Professor, Biochemistry

Ivan Rayment, Professor, Biochemistry

Frank M. Raushel, Professor, Chemistry, Biochemistry and Biophysics, Texas A&M University

Michael G. Thomas, Professor, Bacteriology

Aaron A. Hoskins, Associate Professor, Biochemistry

Alessandro Senes, Associate Professor, Biochemistry

Abstract

Carbohydrates are the most abundant biomolecules on the planet and function in cellular adhesion, cell-cell interactions, fertilization, and the immune response. Besides other well-known roles such as energy storage and utilization, they are critical for the efficacy of a number of antibiotic and antitumor agents produced by Gram-positive bacteria. Gram-negative bacteria, conversely, synthesize cell surface polysaccharides that promote pathogenicity. In the case of the prokaryotic carbohydrates, the component monosaccharides are often deoxygenated and decorated with a variety of chemical groups at one or more positions to afford building blocks with altered chemical and physical properties.

Whereas many Gram-negative cell wall components are comprised of modified sugars, the greatest diversity is observed in those that interact with the environment, including the lipopolysaccharide O-antigens and capsular polysaccharides. The heterogenous nature of these surface structures is responsible for serotyping and contributes to evasion of host defenses in pathogens. Indeed, adaptive protection upon infection with one serotype does not lead to immunity from other serotypes. One notably pathogenic bacterium, *Mycobacterium tuberculosis*, is adept at evading host recognition and persisting, in part due to extensive carbohydrate systems in its highly fortified cell wall. Many of the unique sugars discussed here require extensive modification by a number of different enzymes. Precursors are appended to a nucleoside monophosphate and can subsequently be deoxygenated at C-2', C-3', C-4', and/or C-6' of the pyranose ring. Replacement of hydroxyl groups at these positions with a myriad of chemical substituents from hydrogens to *N*-acetyl groups to amino acids leads to an extraordinary repertoire of prokaryotic unusual sugars.

Chapter 2 details a structural analysis of glucose-1-phosphate thymidyltransferase, or RmlA, from *M. tuberculosis*. This enzyme activates glucose-1-phosphate with thymidine monophosphate

(dTMP) to produce dTDP-glucose and is therefore implicated in the synthesis of a variety of sugars destined for bacterial natural products and O-antigens. However, RmlA from *M. tuberculosis* is an important antibiotic target because of its role in L-rhamnose production, a cell wall component necessary for viability. The three-dimensional model determined here reveals for the first time the manner in which these enzymes coordinate their magnesium bound sugar products and will serve as a scaffold for structure-based drug design efforts against *M. tuberculosis*.

The focus of Chapter 3 is the identification of a complete set of enzymes for the production of 4-formamido-4,6-dideoxy-D-glucose (Qui4NFo) in *M. tuberculosis*. Specifically, the activity of a sugar 4-aminotransferase was absent from the potential biosynthesis pathway. Chapter 3 documents the characterization of Rv3402c, which carries out this required reaction. The results presented should stimulate further studies into the cellular function of Qui4NFo, when it is produced, and its potential role in *M. tuberculosis* pathogenicity.

Chapter 4 consists of a biochemical characterization of the sugar epimerase WbcA, an enzyme required for the production of a rare deoxysugar in the lipopolysaccharide O-antigen of *Yersinia enterocolitica* O:8, namely 6-deoxy-D-gulose. The enzyme was predicted to utilize a mechanism distinct from other C-3' monoepimerases. Based on amino acid sequence alignments, a tyrosine positioned at the end of a β -strand that typically functions as a catalytic acid was thought to be replaced with a cysteine. Instead, structural and kinetic data presented in Chapter 4 indicate that a tyrosine from a different β -strand takes on this role.

Presented in Chapter 5 are initial data towards the functional annotation of two enzymes, Cj1417 and Cj1418, implicated in the biosynthesis of an unprecedented O-methyl phosphoramidate (MeOPN) modification on the capsular polysaccharides of *Campylobacter jejuni*. Whereas their combined activities were hypothesized to produce phosphoramidate in part

due to the glutamine-dependent activity of Cj1417, this work demonstrates that Cj1418 actually catalyzes a glutamine-dependent reaction. The results in Chapter 5 led to the unlikely discovery that Cj1418 is an L-glutamine kinase, detailed in Appendix 1. The Cj1417 data also enabled the characterization of it and a third enzyme, Cj1416, which appends the Cj1418 L-glutamine-phosphate product to cytidine monophosphate to yield cytidine diphospho-L-glutamine. Cj1417 subsequently hydrolyzes this molecule to generate cytidine diphosphoramidate, documented in Appendix 2.

Chapter 6 encompasses a preliminary structural analysis of Cj1415, which catalyzes the phosphorylation of the Cj1417 cytidine diphosphoramidate product. The three-dimensional model described herein is the first of an enzyme necessary for the bioproduction of MeOPN in *C. jejuni*. The position of the Cj1415 active site was predicted based upon the enzyme's similarity to the adenosine 5'-phosphosulfate kinases, and the data from this chapter will be invaluable for further high-resolution structural analyses of the enzyme.

Chapter 7 describes an outreach program in the Holden laboratory called Project CRYSTAL (Colleagues Researching with Young Scientists: Teaching and Learning) that brings middle school students and graduate or undergraduate mentors together in the laboratory. An overview of the structure and goals of the program, as well as changes implemented to make it accessible to laboratories at other universities, is provided.

Acknowledgements

To my advisor, Dr. Hazel Holden: thank you for your patience, understanding, and support, in every aspect of the work, for the last five years. It was a privilege to learn from your expertise in structural analyses, passion for teaching and mentoring, and work ethic. I am grateful I could drop in your office to discuss everything from crystal diffraction to the weekend's symphony performance. Thank you for your friendship and for showing me the importance of taking care of yourself, if not for its own sake, for the sake of doing your best work.

To the rest of my thesis committee: Dr. Ivan Rayment, thank you for always encouraging me to question my results and to "just do the experiment". I appreciate you checking on my blood sugar with group meeting carbohydrates and jams, too. To Dr. Aaron Hoskins, thank you for your example of thinking outside the box and for your continued suggestions in spite of my naive comment about nucleotides and enzymology. To Dr. Alessandro Senes, thank you for joining my committee just in time to chair my preliminary exam and for your insight as a membrane biochemist. To Dr. Michael Thomas, your expertise in microbial physiology and soluble protein expression are worth noting but so are your running skills. Intimidating doesn't begin to cover it. And finally, to Dr. Frank Raushel, thank you for travelling to Madison for my preliminary and final exams, and for everything in between, from hosting me in your laboratory to learn about sequence similarity networks to making sense of my preliminary MeOPN results. I am fortunate and grateful to have collaborated with such a gifted enzymologist.

To additional collaborators: Zane, I haven't met you (that needs to change!), but it has been an absolute joy communicating with you about the MeOPN project. Thank you for your quick responses to my questions and your work characterizing the *Campylobacter* enzymes. I learned a great deal from both you and Dr. Raushel. To Dr. Peter Tipton, thank you for your help scrutinizing

kinetics results. To Dr. Michel Gilbert, I enjoyed our communications about the *Mycobacterium* work and your speculation on the role of the *N*-formylated sugar. To Dr. Evgeny Vinogradov, your NMR analyses have been critical to a number of projects in the Holden laboratory, and I thank you for your contributions to the WbcA and Rv3402c stories. To Dr. Christine Szymanski and Mr. Cory Wenzel, thank you for introducing the Holden laboratory to the fascinating MeOPN modification, and Cory for all of your cloning efforts ahead of when I started on the project.

To my friends in the Holden and Rayment laboratories: Dr. James Thoden, I will never cease to be amazed by how much you accomplish in one day, let alone before 7AM. Your remarkable ability to multitask, plus your fierce commitment to your work and to your students, were a delight to observe and learn from. Thank you for handling my daily questions, some of which were definitely dumb, and for your patience with my mistakes from day one to today. None of my work would have been possible without your help. You also tolerated the beeps coming from my phone, and thanks to you, I now refer to the low-glucose alert as the “Skittles alarm”.

To Dr. Ari Salinger: you were tasked with being the first to deal with me when I was disappointed by a result in lab. Thank you for being understanding, emphasizing the big picture, and for your jokes. They usually helped, but when not, I’m glad I could call you out and you didn’t get mad. I am grateful for your insight from a few steps ahead of me on this journey as a scientist.

To Alex R., Garrett, Brandon, Karl, and Mike: thank you for your comradery and for checking on me once I became the lone occupant in my office. Alex R., Garrett, and Brandon, it was a joy watching you transition from being undergraduates to pursuing your doctorates. Alex R., thanks for going out for beer and running trails with me; the Raushel group is lucky to have you! Mike, since our class doesn’t know much about you, I’m especially glad to have worked alongside you and experienced your humor. To my bench-mate, Karl: I appreciate our chats immensely and that

you always knew just what to ask about, be it science, the weekend, or church “drama”.

To my friends Alex J., Ana, Greg, Karl, Markus, Megan, Stephanie, and Tyler: thank you for being my family in Madison. The laughs, frustrations, shenanigans, conversations, and dogs were numerous. I treasure the time I spent with you, and I’m grateful for what you taught me about life.

To my friends from previous schools: thank you for your long-distance encouragement. Molly, I loved talking about everything from teaching to reminiscing about New Trier with you. To Dr. James West, my mentor from Wooster: thank you for introducing me to research. I am grateful for your continued advice and friendship.

To members of CPC: thank you for taking me in and lifting me up when needed. Rehearsals were frequently the highlights of my week. To the Haslers: thank you for the dinners and spontaneous wine and cheese nights when I could vent. Dale, thank you for letting me share when I solved my first protein structure at Session and for your enthusiasm for “the crystals.”

Finally, I would be lost in all of this were it not for the encouragement of my immediate and extended families. Mom, Dad, Stuart, and Connor: thank you for your unconditional and extensive support. Mom and Dad, I know of no one more hard-working, smart, reassuring, kind, and generous than you two, and I still want to grow up to be just like you. Stuart and Connor, thank you for being my biggest fans and putting up with all my games, recitals, and concerts before now. Stuart, I’m so glad you met Verity, and Verity, I’m grateful for memories made and those to come with my new sister. Aunt Jane, quite simply, you’re the best! To Grandma, the Brumitts and Scotts, Ella, Cullen, Brennan, and Maeve: I realize the perks I’ve had as the oldest cousin. Thank you for indulging them and for your encouragement. To Grandpa, Mémé, and Pop: I often reminisced about you while in lab, wishing I could tell you about my studies. Thinking of you helped me keep perspective and remind me frequently that family is the most precious facet of my life.

Table of Contents

Abstract	i
Acknowledgements	iv
Table of Contents	vii
List of Tables	xii
List of Figures and Schemes	xiii
Chapter 1: Introduction to the Biosynthesis of Unusual Sugars Promoting Bacterial Pathogenicity	1
1.1 Unusual Sugars.....	2
1.2 The Role of Unusual Sugars in Bacterial Pathogenicity.....	5
1.3 Biosynthesis of Unusual Sugars.....	14
1.4 Glucose-1-Phosphate Thymidyltransferases: dTDP-L-rhamnose and RmlA.....	19
1.5 4-Aminotransferases: dTDP-4-formamido-4,6-dideoxy-D-glucose and Rv3402c.....	24
1.6 3-Monoepimerases: CDP-6-deoxy-D-gulose and WbcA.....	29
1.7 Biosynthesis of the <i>O</i> -Methyl Phosphoramidate Capsular Polysaccharide Modification in <i>C. jejuni</i>	32
1.8 Project CRYSTAL.....	38
1.9 References.....	39
Chapter 2: The Structure of Glucose-1-Phosphate Thymidyltransferase from <i>Mycobacterium tuberculosis</i> Reveals the Location of an Essential Magnesium Ion in the RmlA-type enzymes	53
2.1 Abstract.....	54
2.2 Introduction.....	55
2.3 Materials and Methods.....	58
2.3.1 Cloning, expression, and purification.....	58

2.3.2 Protein crystallization and X-ray structural analyses.....	59
2.3.3 Kinetic analyses.....	63
2.4 Results and Discussion.....	65
2.4.1 Overall architecture of <i>M. tuberculosis</i> RmlA.....	65
2.4.2 Structure of the <i>M. tuberculosis</i> RmlA/dTDP-glucose complex.....	69
2.4.3 Kinetic analysis of <i>M. tuberculosis</i> RmlA.....	73
2.5 References.....	77
2.6 Acknowledgements.....	80
Chapter 3: The <i>Mycobacterium tuberculosis</i> Complex has a Pathway for the Biosynthesis of 4-Formamido-4,6-Dideoxy-D-Glucose.....	81
3.1 Abstract.....	82
3.2 Introduction.....	83
3.3 Materials and Methods.....	88
3.3.1 Cloning, expression, and purification of Rv3402c.....	88
3.3.2 Site-directed-mutagenesis.....	89
3.3.3 Enzymatic assay.....	89
3.3.4 Sugar product analysis.....	90
3.3.5 NMR spectroscopy.....	90
3.4 Results and Discussion.....	91
3.5 References.....	99
Chapter 4: Biochemical studies on WbcA, a sugar epimerase from <i>Yersinia enterocolitica</i>.....	102
4.1 Abstract.....	103
4.2 Introduction.....	104
4.3 Materials and Methods.....	109

4.3.1 Cloning of <i>wbcA</i> from <i>Yersinia enterocolitica</i> (type O:8)	109
4.3.2 Protein expression and purification.....	109
4.3.3 Crystallization and structural analysis.....	110
4.3.4 Site-directed mutagenesis.....	113
4.3.5 Kinetic analyses.....	113
4.4 Results and Discussion.....	115
4.5 References.....	121
4.6 Acknowledgement.....	124
Chapter 5: Insights into <i>O</i>-Methyl Phosphoramidate Biosynthesis in <i>Campylobacter jejuni</i>.....	125
5.1 Abstract.....	126
5.2 Introduction.....	127
5.3 Materials and Methods.....	130
5.3.1 Cloning, expression, and purification.....	130
5.3.2 Enzymatic assays.....	132
5.3.3 High-performance liquid chromatography analyses.....	133
5.4 Results and Discussion.....	134
5.4.1 Cj1418 does not function as a PPK or a PEPS.....	134
5.4.2 Cj1417 does not hydrolyze glutamine.....	139
5.4.3 Cj1417 and Cj1418 may function in tandem.....	140
5.4.4 Conclusions.....	144
5.5 References.....	145

Chapter 6: A Preliminary Structural Analysis of Cytidine Diphosphoramidate Kinase from <i>Campylobacter jejuni</i>	148
6.1 Abstract.....	149
6.2 Introduction.....	150
6.3 Materials and Methods.....	154
6.3.1 Protein expression and purification.....	154
6.3.2 Protein crystallization and X-ray structural analyses.....	155
6.3.3 Size exclusion chromatography.....	158
6.3.4 HPLC activity assays.....	158
6.4 Results and Discussion.....	160
6.4.1 Overall architecture of Cj1415.....	160
6.4.2 Potential Cj1415 CDP-NH ₂ binding site.....	168
6.4.3 Role of Ser85.....	169
6.4.4 Potential Cj1415 ATP binding site.....	171
6.4.5 Conclusions.....	174
6.5 References.....	175
6.6 Acknowledgements.....	179
Chapter 7: Project CRYSTAL (Colleagues Researching with Young Scientists: Teaching and Learning)	180
7.1 Goals of Project CRYSTAL.....	181
7.2 Project CRYSTAL Program Structure.....	182
7.3 Project CRYSTAL Mentors.....	185
7.4 References.....	186

Appendix 1: Discovery of a Glutamine Kinase Required for the Biosynthesis of the O-Methyl Phosphoramidate Modifications Found in the Capsular Polysaccharides of <i>Campylobacter jejuni</i>	187
A1.1 Abstract.....	188
A1.2 Introduction.....	189
A1.3 Materials and Methods.....	194
A1.3.1 Protein expression and purification.....	194
A1.3.2 Synthesis of L-glutamine phosphate (VII)	194
A1.4 Results and Discussion.....	202
A1.5 References.....	212
Appendix 2: Biosynthesis of Nucleoside Diphosphoramidates in <i>Campylobacter jejuni</i>	215
A2.1 Abstract.....	216
A2.2 Introduction.....	217
A2.3 Materials and Methods.....	222
A2.3.1 Gene expression and enzyme purification.....	222
A2.3.2 Enzyme assays and determination of kinetic constants.....	222
A2.3.3 Chemical synthesis.....	223
A2.4 Results and Discussion.....	224
A2.5 References.....	235
A2.6 Acknowledgements.....	237

List of Tables

Table 2.1	RmlA X-ray Data Collection Statistics.....	61
Table 2.2	RmlA Refinement Statistics.....	62
Table 2.3	RmlA Kinetic Parameters.....	64
Table 3.1	NMR Spectroscopic data of dTDP-4-Amino-4,6-Dideoxyglucose (dTDP-Qui4N) Produced Using Rv3402c.....	95
Table 3.2	Summary of the Occurrence of dTDP-4-Formamido-4,6-Dideoxy-D-glucose Biosynthesis Genes in <i>Mycobacterium</i> spp.....	96
Table 4.1	WbcA X-ray Data Collection Statistics and Model Refinement Statistics.....	112
Table 4.2	WbcA Kinetic Parameters.....	114
Table 5.1	Cj1417 and Cj1418 Rates Determined via Spectrophotometric Analysis.....	137
Table 6.1	Cj1415 X-ray Data Collection Statistics and Model Refinement Statistics.....	159

List of Figures and Schemes

Figure 1.1	Chemical structures of calicheamicin and erythromycin.....	4
Figure 1.2	Surface polysaccharides of Gram-negative bacteria.....	8
Figure 1.3	Architecture of the mycobacterial cell wall.....	13
Figure 1.4	Enzyme activities utilized in deoxysugars biosynthesis.....	18
Figure 1.5	Common features of sugar nucleotidyltransferases.....	22
Figure 1.6	Comparison of RmlA and UGPase structures.....	23
Figure 1.7	Architecture of the DesI monomer.....	27
Scheme 1.1	Catalytic mechanism of a 4-aminotransferase.....	28
Figure 1.8	C-3' and C-5' epimerization reactions.....	31
Figure 1.9	Selected P-N bond containing molecules.....	34
Figure 1.10	Possible MeOPN production pathway in <i>C. jejuni</i>	37
Scheme 2.1	Pathway for the biosynthesis of dTDP-rhamnose.....	57
Figure 2.1	Overall structure of <i>M. tuberculosis</i> RmlA.....	66
Figure 2.2	Structure of the <i>M. tuberculosis</i> RmlA subunit.....	68
Figure 2.3	Close-up stereo views of the active site with bound dTDP-glucose.....	70
Figure 2.4	Comparison of the experimentally determined RmlA model presented here with that based on the homology modeling previously described.....	72
Figure 2.5	Plots of initial velocities versus substrate concentrations.....	74
Figure 2.6	Superposition of the ligands in the RmlA/dTTP and the RmlA/dTDP-glucose complexes.....	75
Scheme 3.1	Possible pathway for the <i>in vivo</i> production of dTDP-4-formamido-4,6-dideoxy-D-glucose.....	86
Figure 3.1	Alignment of Rv3402c with DesI.....	87
Figure 3.2	HPLC assay of Rv3402c activity and mass spectrum of the Rv3402c product....	92

Scheme 4.1	Predicted pathway for the production of CDP-6-deoxy-D-gulose.....	106
Scheme 4.2	Possible catalytic mechanism for a C-3' monoepimerase.....	107
Figure 4.1	Amino acid sequence alignment of WbcA and other cupin family sugar epimerases.....	108
Figure 4.2	The structure of WbcA.....	116
Figure 4.3	Superposition of the active sites for WbcA and ChmJ.....	118
Figure 4.4	Plots of initial velocities versus substrate concentration.....	120
Scheme 5.1	Structures of MeOPN modified sugars and predicted enzyme activities required for MeOPN biosynthesis.....	129
Scheme 5.2	Possible Cj1418 activity based on similarity to PPDK and PEPS.....	135
Figure 5.1	Cj1418 ATP turnover after 1 hour in the presence and absence of pyruvate or phosphate.....	138
Figure 5.2	ATP turnover after 1 hour in the presence of Cj1417 and Cj1418.....	142
Figure 5.3	ATP turnover after 1 hour in the presence or absence of Cj1417 or Cj1418.....	143
Scheme 6.1	The structures of MeOPN modified sugars and the pathway for phosphoramidate production in <i>C. jejuni</i>	152
Scheme 6.2	Comparison of Cj1415 and APSK activities.....	153
Figure 6.1	Size exclusion chromatography of Cj1415.....	161
Figure 6.2	HPLC assay of wild-type and VAR5 Cj1415 activity.....	162
Figure 6.3	Overall structure of Cj1415.....	165
Figure 6.4	Cj1415 monomer and dimer comparisons to APSKs.....	167
Figure 6.5	Superposition of the potential Cj1415 CDP-NH ₂ binding site with the APS binding site from APSK.....	170
Figure 6.6	Superposition of the potential Cj1415 ATP binding site with the ATP binding site from APSK.....	173
Scheme A1.1	Molecules relevant to the characterization of Cj1417 and Cj1418.....	192

Scheme A1.2	Predicted functions of Cj1417 and Cj1418.....	193
Scheme A1.3	Synthetic transformations for the synthesis of L-glutamine phosphate.....	197
Figure A1.1	¹ H NMR spectrum of the free acid form of L-glutamine phosphate in D ₂ O.....	198
Figure A1.2	³¹ P NMR spectrum of the free acid form of L-glutamine phosphate in D ₂ O.....	199
Figure A1.3	³¹ P NMR spectrum of the sodium salt of L-glutamine phosphate in D ₂ O.....	200
Figure A1.4	ESI negative ion mass spectrum of the sodium salt of L-glutamine phosphate prepared chemically.....	201
Figure A1.5	Anion exchange chromatograms of Cj1418 reaction products.....	203
Figure A1.6	³¹ P NMR spectra of Cj1418 reaction products.....	205
Figure A1.7	Mass spectrum of Cj1418 reaction mixture.....	207
Figure A1.8	ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1418, ATP, and L-glutamine.....	208
Scheme A1.4	Experimentally determined function of Cj1418 and possible catalytic functions of Cj1417 and Cj1416.....	211
Scheme A2.1	Structures of the <i>O</i> -methyl phosphoramidate modifications to the CPS in <i>C. jejuni</i>	218
Scheme A2.2	Reaction catalyzed by Cj1418 and the predicted functions of Cj1417 and Cj1416.....	221
Figure A2.1	Anion exchange chromatograms of nucleotide standards and enzyme-catalyzed reaction products.....	225
Figure A2.2	³¹ P NMR spectra of nucleotide standards and enzyme-catalyzed reaction products.....	227
Figure A2.3	ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1416, Cj1417, MgCTP, and L-glutamine phosphate.....	228
Figure A2.4	³¹ P NMR spectra of nucleotide standards and enzyme-catalyzed reaction products.....	231
Figure A2.5	ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1416, MgCTP, and L-glutamine phosphate.....	232

Chapter 1:

Introduction to the Biosynthesis of Unusual Sugars Promoting Bacterial Pathogenicity

1.1 Unusual Sugars

Carbohydrates constitute the most abundant biomolecules on the planet and play critical roles in all kingdoms of life. They are well known for their functions in energy storage and utilization in the form of glucose, glycogen, and starch, but are also key components of structural molecules such as cellulose and chitin (1). Furthermore, the different ribose sugars incorporated into ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) have profound effects. Indeed, DNA is more stable than RNA due the absence of a hydroxyl group at C-2' of the ribose sugar and thus makes DNA an ideal information storage molecule (2).

Besides contributions to the storage, transfer, and use of energy and genetic information, carbohydrates aid in cellular adhesion, cell-cell interactions, fertilization, and the immune response (3-6). An example of latter is the ABO blood-typing system which is based on sugar-containing antigens on the surface of erythrocytes and other cells. Those with type A blood produce A antigen containing *N*-acetylgalactosamine (GalNAc) at the distal end of the sugar core. Exchanging GalNAc with galactose (Gal) affords a B antigen structure and B-type blood, while the H antigen lacks a sugar at this position and is the basis for O-type blood (7). The seemingly simple presence or absence of one sugar or the substitution of a hydroxyl with an *N*-acetyl group in the A antigen has a profound impact on clinical decisions during blood transfusions and organ transplants (8).

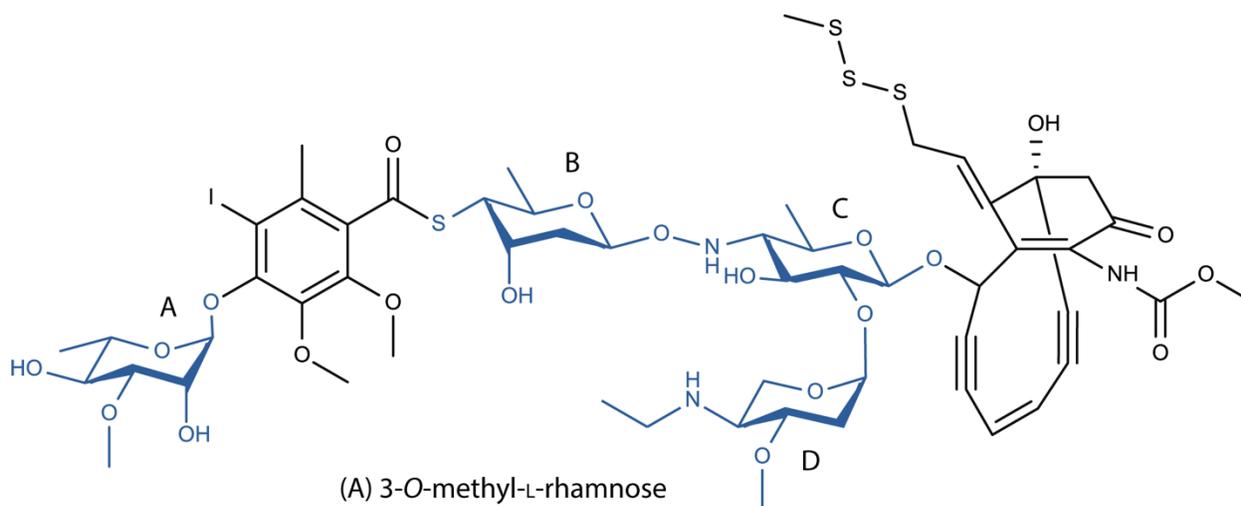
Though the ABO antigens harbor modified (relative to glucose) sugars like L-fucose (Fuc) and GalNAc, eukaryotic polysaccharide complexity is achieved by linking monosaccharides together in unique patterns rather than incorporation of unusual building blocks (3, 6, 9). Indeed, although half of the mammalian proteome is predicted to be glycosylated, the glycans are comprised of just ten monosaccharides (10). Yet, over 7,000 structures have been characterized to date due to variations in glycosidic bonds, length, branching, and substitutions (10, 11).

Bacteria, on the other hand, synthesize extraordinarily complex sugars (3, 4, 12-14). Many of these unique structures are critical for the biological activities of natural product antibiotic and antitumor agents produced by species of *Saccharopolyspora* and *Streptomyces* (3, 4, 6). One of the most potent classes of natural products is the enediyne molecules of which calicheamicin is the best characterized (Fig. 1.1) (4). Calicheamicin binds the minor groove of DNA and induces scission of the backbone. Notably, removal of the unusual sugars appended to the aglycone core leads to poor binding affinity and specificity, as well as impaired DNA cleavage (4). Due to its cytotoxicity, clinical applications of calicheamicin are limited to targeted therapies such as antibody-drug conjugates in the treatment of lymphoblastic leukemias (15).

A more medically relevant example of the importance of unusual sugars in antibiotic efficacy is that of erythromycin, a macrolide antibiotic that inhibits peptide chain elongation in the ribosome and is used to treat Gram-positive, mycoplasma, and certain Gram-negative infections (4). Erythromycin A is composed of a 14-member aglycone ring core and two deoxysugars, L-cladinose and D-desosamine (Fig. 1.1). Importantly, alteration of D-desosamine or removal of either one of the sugars diminishes or completely abrogates biological activity (4, 6).

Many natural products containing unusual sugars are produced by Gram-positive bacteria, but complex carbohydrates are critical components of bacterial surface structures in general (16-20). They are found in the *S*-layers and capsular polysaccharides (CPS) of Gram-positive and Gram-negative bacteria, the Gram-negative lipopolysaccharide (LPS) and lipooligosaccharide (LOS), and the *Mycobacterium* cell wall (16-20). Because they interact extensively with the microbial environment, many of these molecules contribute to pathogenicity in the context of a host and for the purposes of this thesis, will be discussed further (16-20).

Calicheamicin



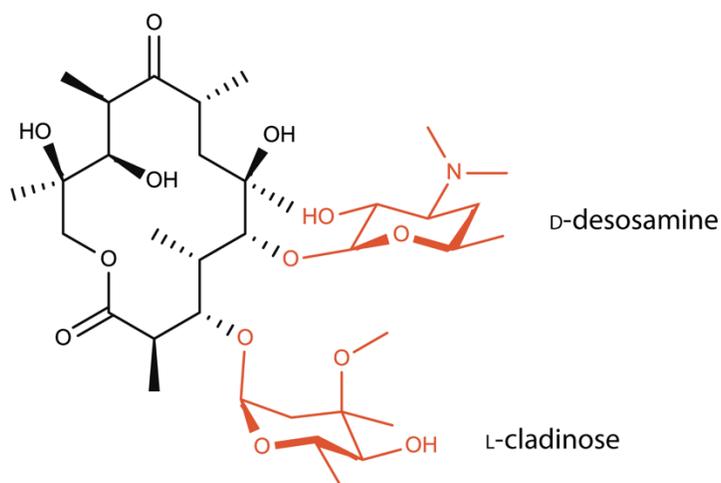
(A) 3-O-methyl-L-rhamnose

(B) 4-deoxy-4-thio-D-digitoxose

(C) 4-deoxy-4-hydroxylamino-D-quinovose

(D) 4-ethylaminodideoxy-L-pentose

Erythromycin



D-desosamine

L-cladinose

Figure 1.1 Chemical structures of calicheamicin and erythromycin. Note the four deoxysugars on the enediyne antibiotic, calicheamicin, as well as the two unusual sugars on the macrolide antibiotic erythromycin.

1.2 The Role of Unusual Sugars in Bacterial Pathogenicity

Bacteria proliferate in diverse environments on earth, however persistence in a host and induction of disease are hallmarks of pathogens. They are often compared in terms of their virulence or “the relative capacity of a pathogen to overcome body defenses” (21). An overview of bacterial pathogenicity will be provided, but the bulk of this discussion details the role of unusual sugars produced by Gram-negative and *Mycobacteria* species in immune evasion.

Bacteria co-exist with healthy humans by residing on or in a number of organs, most notably, the gastrointestinal tract (22). Commensal microbes can invade normally sterile tissue when a breakdown in host defenses occurs and are termed opportunistic pathogens (23). Non-opportunistic pathogens, however, cloak themselves from the immune response, adhere and colonize, and can invade further and replicate despite host processes to mitigate this (23, 24). Successful pathogens must overcome antimicrobial peptides, secreted antibodies, and reactive oxygen species characteristic of inflammation (24). Furthermore, some bacteria produce toxins that cause disease symptoms whereas in other cases, symptoms are manifestations of the immune and inflammatory responses (23).

Upon exposure to foreign bacteria, the host innate immune response is engaged to clear the organism (25). Pattern recognition receptors (PRRs) recognize common bacterial cell wall constituents including an essential peptidoglycan (PG) that maintains cell shape and turgor pressure (Fig. 1.2) (25-28). PG is usually comprised of a repeating *N*-acetylglucosamine (GlcNAc) -*N*-acetylmuramic acid (MurNAc) disaccharide cross-linked via short peptides (28). Pathogens can modify these sugars via *N*-deacetylation, *O*-acetylation, and *N*-glycosylation to increase resistance to lysozyme and antibiotics (25). Furthermore, *N*-glycolylmuramic acid incorporation is thought to increase PG hydrogen bonding propensity and strength in species of *Mycobacterium* (29).

In Gram-negative bacteria, the membrane bound lipid A portion of the LPS or LOS is necessary for growth and serves as the primary endotoxin in the context of disease (Fig 1.2) (16). Pathogens exhibit diversity in the number and location of 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) sugars attached to lipid A to alter the conformation and endotoxicity of the molecule (27). Lipid A is also subject to modification with phosphorylethanolamine (PEtN), 4-amino-4-deoxy-L-arabinose (L-Ara4N), and palmitate to promote resistance to certain antibiotics and cationic antibacterial peptides (CAMPs), thereby contributing to pathogenesis (30).

Broadly speaking, the structural diversity of bacterial surface components increases the farther they are from the cell (31). An evolutionary explanation stems from two observations: 1) Due to the requirement of lipid A/Kdo for membrane integrity, few modifications can be made to these molecules while maintaining membrane fidelity. 2) Bacteria are subject to selective environmental pressure, so it stands to reason that surface exposed substituents like the LPS O-antigens, the LOS, and capsules exhibit remarkable variety both within and across species (31).

Indeed, Kdo and the 7-carbon sugar *L-glycero*- β -D-*manno*-heptose, or Hep, which make up the LPS inner core, are variably modified with phosphate, pyrophosphorylethanolamine (PPEtN), and phosphorylcholine (PCho) to block the action of antimicrobial peptides (Fig 1.2) (16, 32). Strains of *Salmonella enterica* that cannot phosphorylate the inner core, but retain a complete LPS, are avirulent in a mouse model of infection, while the same variant of *Pseudomonas aeruginosa* is not viable (33, 34). This example highlights the importance of sugar modification in certain species, and that identical structures can possess different functions depending on the bacteria.

In full-length LPS molecules, an outer core bridges the inner core and O-antigen (Fig. 1.2) (16). The outer core is typically comprised of simpler sugar residues modified non-stoichiometrically. Negatively charged galacturonic acid (GalA) is found in strains with reduced

or no phosphorylation to maintain stability (32). Interestingly, certain core structures in *Escherichia coli* and *Klebsiella pneumoniae* are highly represented in clinical isolates (16). In *K. pneumoniae*, swapping core 2, which is prevalent in virulent strains, with core 1 leads to attenuation in a mouse model of infection, even while K-antigen (O-antigen like capsule) expression is maintained, implicating core structure in virulence, be it direct or indirect (35).

Mucosal pathogens do not synthesize O-antigen but instead produce a LOS with a small sugar unit attached to the inner core (36). The variable LOS structures are an example of phase variation used by bacteria in a *single* population to produce heterogeneity (37). The strategy can be thought of as “bet hedging” that an immune response clears certain cell surface phenotypes, allowing others to disseminate (38). The LOS is also one manifestation of molecular mimicry in which pathogens avoid recognition by producing similar structures to those found in a host (39). The LOS and non-stoichiometric modifications of lipid A core components are illustrations of a property of successful pathogens: heterogeneous surface polysaccharides promote immune evasion (16, 37).

A prominent feature of some Gram-negative bacteria is the capsular polysaccharide or CPS (Fig. 1.2) (17). Capsules are virulence markers in both Gram-negative and Gram-positive bacteria as encapsulated strains are more pathogenic than non-encapsulated ones (17, 40-42). Their functions are species dependent but include the prevention of desiccation, adherence, and resistance to non-specific and specific host immunity (17, 42). Capsules are comprised of 2-10 sugars in repeating units, and relative to the unique sugars of the O-antigen, use simpler building blocks (Fig 1.2). Though, components do include rare furanoses, deoxysugars, and heptoses, and variation results from modification with *N*- and *O*-acetyl groups, amino acids, glycerol, and phosphate (42). The large number of serotypes in certain species contributes to immune evasion as infection with one type does not lead to adaptive immune protection from another (17, 43).

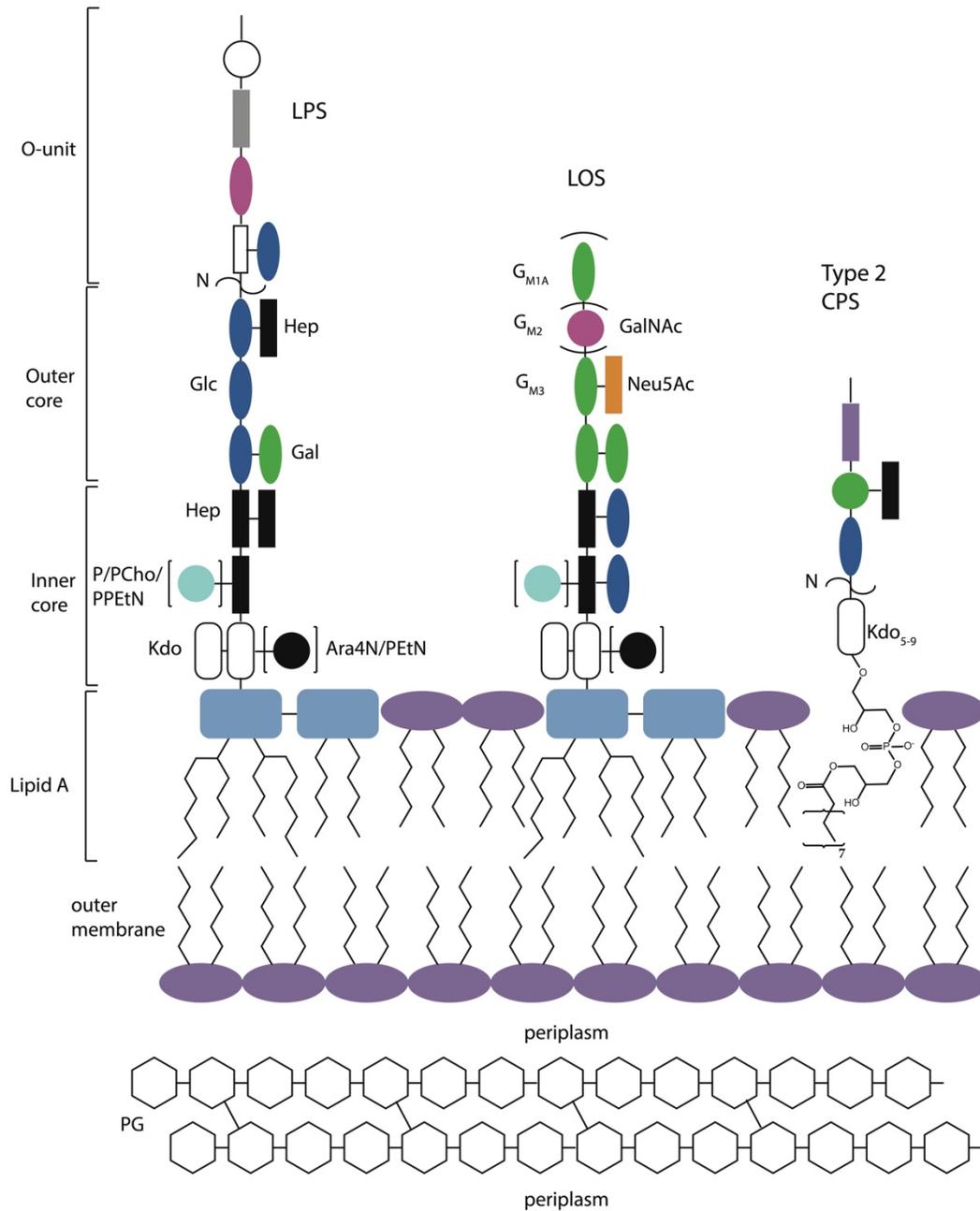


Figure 1.2 Surface polysaccharides of Gram-negative bacteria. Presented is a generic LPS molecule, adapted from Ref (16); the LOS of *Campylobacter jejuni* 11168/HS:2; a generic type 2 CPS. Kdo: 3-deoxy-D-manno-oct-2-ulsonic acid, Hep: L-glycero- β -D-manno-heptose, Gal: galactose, Glc: glucose, P: phosphate, PCho: phosphorylcholine, PPeTn: pyrophosphoryl-ethanolamine, PEtN: phosphorylethanolamine, Ara4N: 4-amino-4-deoxy-L-arabinose, Neu5Ac: N-acetylneuraminic acid, GalNAc: N-acetylgalactosamine, PG: peptidoglycan. Modifications in brackets are non-stoichiometric. Sugars in parentheses are phase variable in *C. jejuni* strain 11168 and contribute to mimicry of host gangliosides G_{M2} and G_{M1A}, whereas a mutant GalNAc transferase in HS:2 leads to G_{M3} mimicry. Proteins and the inner membrane are omitted for clarity.

Certain capsules expressed by *E. coli*, *Streptococcus pneumoniae*, and *Neisseria meningitidis* are prevalent in strains that cause disease, suggesting that the structure of the capsule, not merely its presence, contributes to pathogenicity (42). Additionally, O-acetylation of the capsule has varying immunomodulatory effects in *N. meningitidis*, *Salmonella typhi*, and species of *Burkholderia*, highlighting the importance of non-carbohydrate modifications (44-46).

Many Gram-negative bacteria append an O-antigen to the lipid A core sugars to produce full-length LPS (16). The O-polysaccharide is comprised of repeating units of two to eight sugars, protrudes farthest from the membrane, and is therefore the most structurally diverse component of the LPS (18). Over 100 monosaccharides and 40 non-carbohydrate substituents are known, making the O-antigen one of the most prominent reservoirs of unusual sugars in nature (Fig. 1.2) (18). Most are hexoses and include the 2(3)(6)-deoxyhexoses, 2(3)(4),6-dideoxyhexoses, 2,3-dideoxyhexoses, and a rare trideoxyhexose (14, 18, 47). The O-antigen is the basis for serotyping, masks cell wall antigens from the alternative complement pathway, inhibits clearance via phagocytosis and plays a role in membrane integrity, invasion, adherence, colonization, spreading, and expression of other virulence factors (16, 18, 48-52). Importantly, protective immunity against one serotype does not defend against infection by a different type (16, 18, 53, 54). In some instances, this is mediated by a non-carbohydrate group as lactic acid in *Providencia alcalifaciens* O32 produces antiserum that reacts poorly with the equivalent O-antigen lacking lactic acid in O29 (55). Additionally, O-antigen heterogeneity in a *single bacteria* arises due to non-stoichiometric glycosylation, acetylation, methylation, amidation, or phosphorylation (56). Importantly, “smooth” strains with full length O-antigen tend to be more pathogenic than “rough” strains with one that is truncated or absent, making the structure a potent virulence marker (57-59). Indeed, deletion of O-antigen biosynthesis or processing genes in *Brucella abortus* and enterohemorrhagic

E. coli attenuates pathogenicity and increases susceptibility to host antimicrobial factors (60, 61).

Substantial O-antigen heterogeneity does little to rationalize why it is composed of unusual sugars in the first place, especially because the O-antigen is typically not necessary for viability outside the host but rather enhances virulence. The contribution of individual sugars to pathogenicity is difficult to study because disruption of biosynthesis genes typically leads to a rough phenotype. However, in *Salmonella*, serogroups expressing the dideoxysugars paratose, abequose, and tyvelose, as opposed to common sugars, tend to be more pathogenic (57). Indeed, strains containing abequose or tyvelose are taken up by macrophages at reduced rates compared to those with Glc, GlcNAc, and mannose, correlating with the known virulence of each strain (62). Similarly, exchanging tyvelose biosynthesis genes in *Salmonella typhi* O9 with those for abequose production increases resistance to complement, with an otherwise equivalent O-antigen, implicating the dideoxyhexoses in immune recognition (63).

A basis for the observed lack of protective immunity across serogroups was revealed by the structure of a mouse antibody in the presence of an O-antigen fragment. The common sugars, galactose and mannose, are exposed to solvent whereas the abequose is tightly associated with the antigen binding pocket (64). Similarly, serum antibodies can distinguish capsules differing by one glycosidic linkage (1-3 vs. 1-4) or sugar configuration (*Galp* vs. *Galf*) and the presence or absence of an acetyl group at equivalent positions (65, 66). From these examples, it appears that individual components contribute to immune evasion and pathogenicity, not just the presence of an O-antigen or capsule. As such, characterization of enzymes involved in the biosynthesis of unusual sugars in these surface structures is warranted in order to inhibit them and potentially control pathogenesis.

With respect to unique polysaccharides, one particularly virulent species deserves attention, namely *Mycobacterium tuberculosis* (Mtb). Mtb is considered Gram-positive but due to an

extensive outer membrane of mycolic acids, the organism resembles Gram-negative bacteria. The cell wall is critical for viability and virulence and is a formidable barrier to hydrophilic antibiotics (20, 29, 67, 68). A simplified depiction of the Mtb cell wall is provided in Figure 1.3, and structures containing unusual sugars will be considered in the context of traits that promote pathogenicity, including the ability to modulate the immune system and persist in macrophages (20, 29, 67, 69).

Like capsules previously described, the Mtb α -glucan contributes to evasion of innate and acquired immunity, but also promotes persistence of the organism (70-72). Trehalose mono or dimycolate (TMM/TDM), di- and poly-acyltrehaloses (DAT/PAT), and the sulfated glycolipid, SL-1, are built on trehalose scaffolds, a non-reducing disaccharide of glucose in an α 1-1 glycosidic arrangement. SL-1 has been assigned varied and controversial roles (20). Nevertheless, buildup of SL-1 precursors attenuates virulence, and it is implicated in propionate stress modulation as a result of using host cholesterol for carbon upon invasion of macrophages (73-75). Similarly, the role of DAT and PAT in infection is unclear despite immune modulation effects *in vitro*, possibly because all acyltrehaloses seem to promote reduced binding to immune cells, making it difficult to decipher the role of each molecule (20, 76). Still, the acyltrehaloses may contribute to phagosome arrest in macrophages, propionate stress tolerance, and serve as immune system antagonists to evade recognition (73, 74, 77, 78). Furthermore, TDM is a potent virulence factor involved in immune modulation and protection from killing by macrophages and antibiotics (20, 67, 79).

Mtb incorporates *O*-methylated deoxysugars into cell wall or secreted components, namely the phenolic glycolipid PGL-tb and *para*-hydroxybenzoic acid (*p*-HBAD) (80). While both are implicated in abrogating the production of pro-inflammatory cytokines, an effect dependent on the presence of the deoxysugars, the cellular roles of PGL and *p*-HBAD are undefined (81).

The Mtb arabinogalactan (AG), comprised of galactofuranose and arabinofuranose, is required

for viability and forms a substantial portion of the cell-wall, serving as a barrier to hydrophilic compounds (29, 79). The AG is attached to the peptidoglycan by an essential L-rhamnose-GlcNAc linker and is subject to modification by succinyl or galactosamine (GalN) residues (68, 82). Both are thought to promote AG rigidity, but GalN is associated with reduced immune cell activation and increased secretion of anti-inflammatory cytokines (68, 83).

Phosphatidyl-*myo*-inositol mannosides (PIMs), lipomannan (LM), and lipoarabinomannan (LAM) are deposited noncovalently in the inner and mycolic membranes of all species of *Mycobacteria* (68). While all three maintain cell wall structure, integrity, and permeability, the LM and LAM are involved in host-organism interactions, immune modulation, and are structurally heterogeneous (68, 79, 84). Mannose-capped LAM, found in pathogenic species like Mtb and *Mycobacterium leprae*, binds receptors on a number of cell types with anti-inflammatory effects, impairs phagosome maturation, inhibits apoptosis, and negatively regulates reactive oxygen species in macrophages (68, 84). Mtb can append 5-deoxy-5-methylthio-xylofuranose (MTX) to mannose-capped LAM that alters cytokine production and may promote intracellular survival (85). Many of the capped LAM effects have only been observed *in vitro* due to their overlapping functions and the difficulty in generating isogenic Mtb mutants (69, 84). Nonetheless, the complex and diverse LAM structures are antigenic (86). From the few cases detailed here, it is clear that Mtb takes advantage of unique polysaccharides to promote pathogenicity, albeit using slightly different mechanisms than the Gram-negative bacteria previously described (84).

Characterization of enzymes from pathogenic bacteria required for the production and modification of unusual sugars is critical because many of them may serve as targets for antibiotic development. An overview of these biosynthetic pathways is provided below, specifically in the context of the O-antigen, CPS, and Mtb polysaccharides.

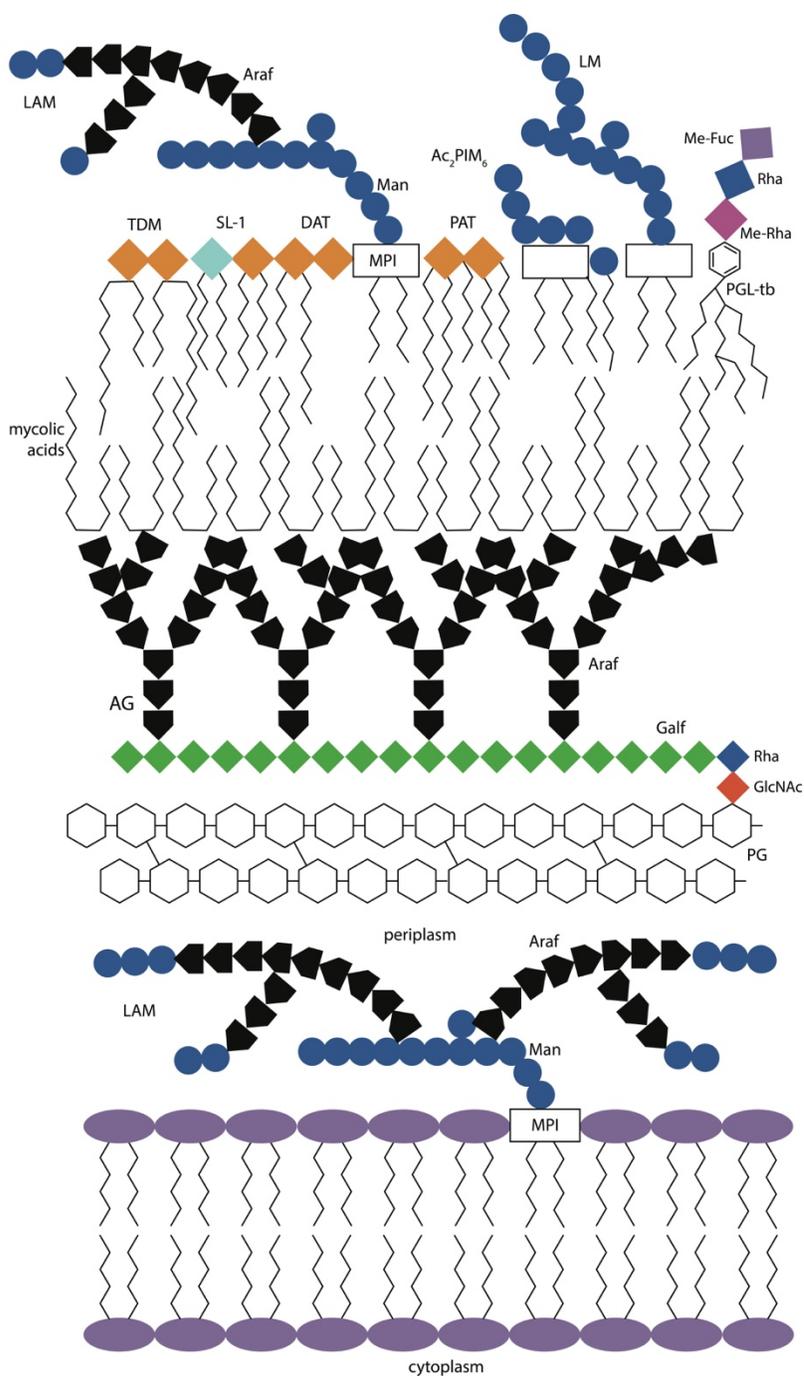


Figure 1.3 Architecture of the mycobacterial cell wall. Man: mannose, Araf: arabinofuranose, LAM: lipoarabinomannan, LM: lipomannan, MPI: mannosyl phosphate inositol, PG: peptidoglycan, GlcNAc: *N*-acetylglucosamine, Rha: L-rhamnose, Galf: galactofuranose, AG: arabinogalactan, TDM: trehalose dimycolate, SL-1: sulfo-glycolipid-1, DAT: diacyltrehaloses, PAT: polyacyltrehaloses, Ac₂PIM₆ – diacylated phosphatidyl-*myo*-inositol (6) mannoside, PGL-tb: phenolic glycolipid tuberculosis, Me-Rha: *O*-methyl L-rhamnose, Me-Fuc: *O*-methyl fucose. The proteins and α -glucan capsule are omitted for clarity, and the mycolic acid layer is denser than depicted here.

1.3 Biosynthesis of Unusual Sugars

Unusual sugars destined for the O-antigen or capsule are synthesized as individual nucleotide-linked precursors in the cytoplasm. In the first step, a nucleotidyltransferase activates either glucose-1-phosphate (Glc-1-P) or mannose-1-phosphate (Man-1-P) with a nucleoside monophosphate (NMP) (5, 14, 47, 87, 88). In the case of deoxygenated sugars, a nicotinamide adenine dinucleotide (NAD⁺)-dependent 4,6-dehydratase subsequently oxidizes C-4' to a keto-moiety and reduces C-6' via replacement of the C-6' hydroxyl group with a hydrogen (89-92). The overall reaction is redox neutral, but newly acidic protons at C-3' and C-5' support the activities of downstream enzymes (Fig. 1.4). The reactions depicted in Figure 1.4 are a subset of the remarkable chemical transformations available to nucleotide-linked sugars. Indeed, the structural enzymology of unusual sugar biosynthetic enzymes is a lens through which to study the chemical transformations necessary to support life as these enzymes use a variety of cofactors during catalysis (3, 9, 92).

Some sugar modifications are derived from unique co-substrates, such as 3-hydroxybutanoyl-coenzyme A (CoA) and 3-hydroxy-3-methylbutyryl-CoA in the case of the *N*-acyltransferases FdhC and AntD, respectively (93, 94). Other substituents, such as phosphoenolpyruvate used by MurA in peptidoglycan MurNAc production or the amino acids found in the O-antigens of *Providencia*, are derived from primary metabolism (95, 96). Finally, some unusual chemical groups, such as the *O*-methyl phosphoramidate found in the capsules of *Campylobacter jejuni*, require extensive enzymatic machinery for production (97).

Sedoheptulose-7-phosphate is the precursor for heptoses incorporated into both the capsule and lipid A core. Those marked for the CPS are activated with guanosine diphosphate (GDP), whereas those incorporated into lipid A are activated with adenosine diphosphate (ADP) (16, 98).

The sugar biosynthesis pathways employ nucleotide-linked substrates to target them to different pathways, for regulatory purposes, and because the NDP is as an ideal leaving group in glycosyltransferase reactions that transfer the modified sugar to its final destination (99, 100).

Indeed, most deoxysugars appended to the aglycone scaffolds of natural products are synthesized about thymidine diphosphate-linked precursors (dTDP), however, some cell wall sugars are constructed around this nucleotide as well. Activation with cytidine diphosphate (CDP) is utilized almost exclusively for the production of 3,6-dideoxysugars found in the O-antigens of species of *Yersinia* and *Salmonella*. Uridine diphosphate (UDP)-linked monosaccharides function in primary metabolism, peptidoglycan synthesis, and production of cell wall polysaccharides. GDP-linked precursors are most often mannose derived and used as building blocks in the biosynthesis of both prokaryotic and eukaryotic cell surface glycoconjugates. Finally, the ADP-linked sugars are involved in construction of the lipid A core and glycogen metabolism in bacteria (3, 57, 99). The nucleotidyltransferases and glycosyltransferases are often subject to engineering experiments for their re-design and application in the *in vitro* synthesis of compounds with novel biological activity, an effort called glycorandomization (3, 12).

Once the nucleotide-linked sugar precursors are made, O-antigen polymerization proceeds via one of two main routes. In the Wzy-dependent pathway, which appears confined to heteropolysaccharides that are often branched, a single O-unit is synthesized by glycosyltransferases on an undecaprenol-pyrophosphate (Und-PP) anchored in the cytoplasmic leaflet of the inner membrane. The O-unit is transferred to the periplasmic leaflet by the Wzx flippase, and O-unit polymerization is catalyzed by Wzy (16). The second prevalent pathway, dependent on an adenosine triphosphate (ATP)-binding cassette (ABC)-transporter, is responsible for the production of non-branched homopolymers. The entire O-antigen is synthesized on a single Und-

PP on the cytosolic leaflet of the inner membrane, and the ABC-transporter transfers the O-antigen to the periplasmic leaflet. In both systems, WaaL subsequently ligates the O-polysaccharide to the lipid A core (16). Finally, the LptA-G trans-envelope complex is responsible for the extraordinary biophysical journey the LPS molecule must take from the periplasm to the extracellular leaflet of the outer membrane (101).

CPS operons fall into four different groups, but for the sake of relevance to this work, only the group 2 ABC-transporter-dependent pathway will be highlighted (17, 88, 102). The initiating process for chain synthesis was only recently unraveled and employs a membrane bound, Kdo capped (lyso)-phosphatidylglycerol (a glycerophospholipid with one acyl chain) or phosphatidylglycerol instead of Und-PP (103-105). Cytosolic glycotransferases produce the repeating capsule after initiation. KpsM and KpsT, the ABC-transporter components, translocate the completed capsule to the periplasm where it is transferred to KpsD and KpsE for deposition into the outer leaflet of the outer membrane (106). Finally, in both O-antigen and capsule synthesis, some modifications like *O*-acetylations can occur after polysaccharide polymerization (16, 17).

Mtb synthesizes a number of unusual sugar containing structures, but those promoting pathogenesis, rather than ones required for viability, will be highlighted here. PIM, LAM, and LM biosynthesis all initiate in the cytosol with the transfer of mannose from GDP-mannose to the inositol ring of a membrane-bound phosphatidyl-inositol (PI). Addition of mannose and acyl chains leads to mono or di-acylated PI with four mannose residues, the branch point in PIM, LAM, and LM synthesis (68, 79). The manner in which this intermediate is transported to the periplasm is unknown. Because subsequent steps are carried out in the periplasm, the mannose donor for these reactions is a polyprenyl-phosphate-based mannose, not GDP-mannose (107). Three mannosyl-transferases, MptA-C, synthesize the mannan core and introduce branch points.

Arabinofuranose is introduced via a number of transferases and mannose capping is catalyzed by CapA and MptC (68, 79). The MTX cap is made and modified on a nucleotide linked scaffold in the cytosol, although the identity of the substrate for translocation to the periplasm is unknown. MtxT, however, transfers the MTX sugar to the mannoside cap (108).

Trehalose, required for acyltrehalose production, is synthesized primarily through the OtsAB pathway using UDP-glucose and glucose-6-phosphate precursors (109). Though mycolic acid and trehalose biosynthesis both occur in the cytoplasm, the mechanism(s) governing TMM formation from each is currently unknown. However, once made, TMM is transported to the periplasm by a flippase, MmpL3 (110). TDM formation from TMM and transfer of mycolates to the arabinogalactan is accomplished by Antigens 85A/B/C (20).

PGL-tb is comprised of three deoxysugars, two of which are *O*-methylated, appended to a phenol derivatized with phthiocerol dimycocerosate (PDIM) (Fig. 1.3) (80). dTDP-L-rhamnose, which is also required for cell wall biosynthesis, and GDP-fucose serve as precursors (111). Three dedicated glycosyltransferases attach the sugars onto the phenol-PDIM core in the correct order, and four *S*-adenosylmethionine (SAM)-dependent *O*-methyltransferases modify the hydroxyl groups in L-rhamnose and fucose (81). PGL-tb synthesis occurs in the cytosol, however the process of translocation to the outermost layer of the Mtb cell wall is not well understood (20).

The following sections will preview the structural and functional characteristics of enzymes involved in O-antigen and Mtb deoxysugar biosynthesis as well as the production of capsule modifications in *C. jejuni*. The unifying motivation for the study of these enzymes is that they may serve as novel antibiotic targets in pathogenic bacteria given their roles in immune evasion and/or persistence.

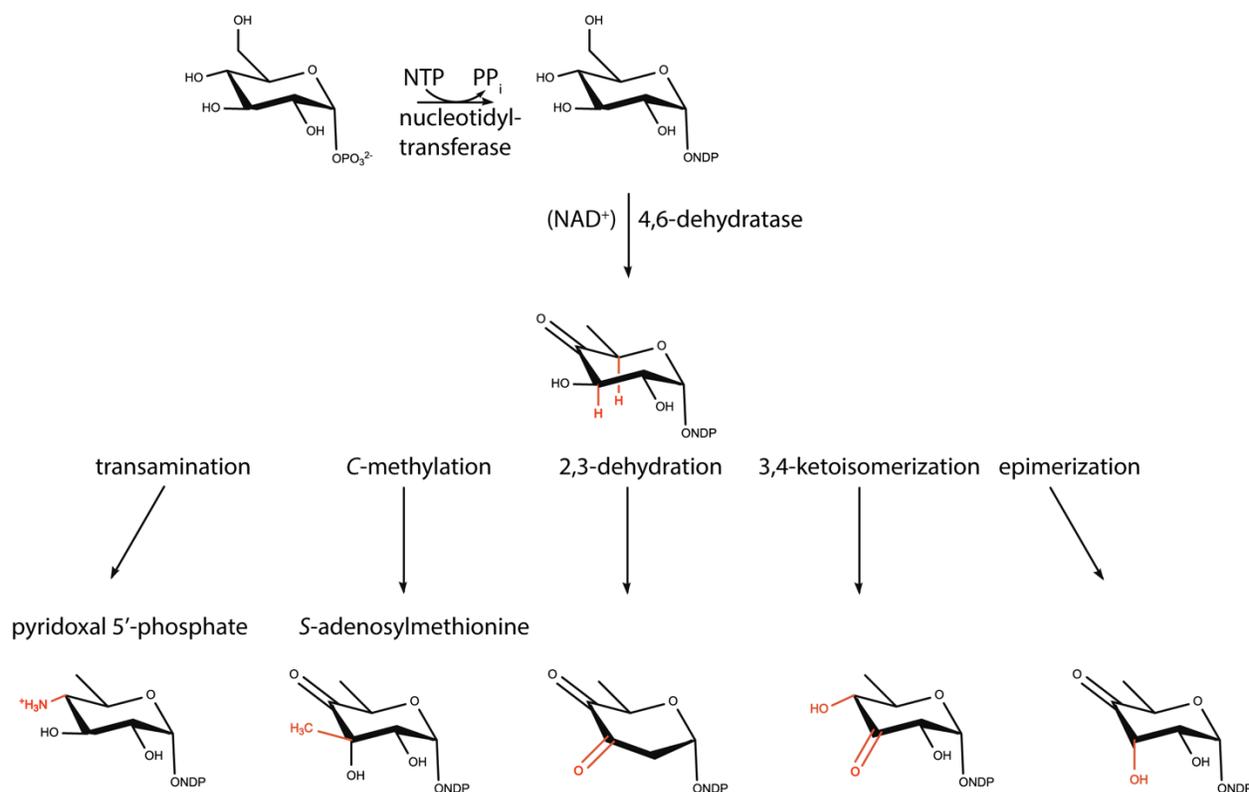


Figure 1.4 Enzyme activities utilized in deoxysugar biosynthesis. In the first step, mannose-1-phosphate or glucose-1-phosphate (pictured) is activated with a nucleoside monophosphate. The NDP-glucose is then dehydrated to yield a 4-ketosugar with newly acidic protons at C-3' and C-5', highlighted in red. This activation supports downstream chemical reactions, a subset of which are shown and highlighted in red. Some of the enzymes require cofactors like pyridoxal 5'-phosphate (PLP)-dependent aminotransferases and S-adenosylmethionine (SAM)-dependent methyltransferases.

1.4 Glucose-1-Phosphate Thymidyltransferases: dTDP-L-rhamnose and RmlA

Chapter 2 documents a structural characterization of glucose-1-phosphate thymidyltransferase, or RmlA, from Mtb. While RmlA from this organism is necessary for viability, these enzymes are found in a variety of pathogenic bacteria that make L-rhamnose (112). Given its roles in L-rhamnose biosynthesis and the production of a variety of deoxysugars found on natural products, the lack of structural data supporting the positions of active site Mg^{2+} ions is striking (3). An introduction to the sugar nucleotidyltransferases is provided below.

As outlined in section 1.3, deoxysugar biosynthesis pathways initiate with the activation of Glc-1-P or Man-1-P with an NMP, from a nucleoside triphosphate (NTP), to yield NDP-glucose or NDP-mannose and pyrophosphate (14, 87). Such reactions are catalyzed by hexose-1-phosphate nucleotidyltransferases but are often referred to in the literature as NDP-hexose pyrophosphorylases due to the reversibility of the reaction *in vitro* [Fig 1.5 (c)] (99). Sugars are synthesized as nucleotide-linked precursors because: 1) the NDP functions as an ideal leaving group in the glycosyltransferase reactions that transfer the sugar onto its target 2) the use of different nucleotides marks sugars for disparate pathways and 3) it enables differential regulation (99, 100). As such, nucleotidyltransferases using the triphosphates of thymidine, adenosine, uridine, guanosine, and cytidine (dTTP, ATP, UTP, GTP, and CTP) have all been identified, and most are regulated via feedback inhibition (100, 113-117).

The sugar nucleotidyltransferases possess remarkably similar N-termini as this represents the NTP binding site (Fig. 1.5). The flexible (G)GXG(T)R(L)X_{6,9}(P)K loop is a defining feature of these enzymes despite the fact that they adopt a variety of oligomerization states; they contain additional domains like the left-handed β -helix (LBH) motif in the ATP and GTP enzymes; some are bi-functional (phosphomannose-isomerase/guanylyltransferase, GlmU – *N*-acetyltransferase/

uridylyltransferase) (118-126). Though nucleotidyltransferases are subject to inhibition, structural studies on the cytidylyltransferases have not determined how this is achieved despite biochemical data invoking regulation at an alternative site (116, 122, 127). Conversely, the points of control have been characterized in the adenylyltransferases and RmlA-type thymidylyltransferases (119, 128). The site in RmlA from *P. aeruginosa* displays modest specificity for dTDP-L-rhamnose, the biological inhibitor, and instead towards dTDP-containing compounds in general (119). Given that the regulatory sites are not well conserved, targeting them may be a valuable way to inhibit specific RmlAs, a strategy undertaken with the *P. aeruginosa* enzyme (129).

The available structural models for hexose-1-phosphate nucleotidyltransferases all demonstrate a hallmark seven to nine stranded mixed β -sheet flanked by α -helices, represented by RmlA in Figure 1.5 (b). The nucleotidyltransferase and LBH domains of the GTP and ATP dependent enzymes are quite similar (123, 124). Likewise, the monofunctional dTTP and UTP using enzymes (RmlA and UGPase) display significant structural homology, demonstrating ~30% sequence identity (119, 121). A comparison of the subunits indicates the only obvious differences of an extra α -helix and insertion sequence in UGPase [Fig. 1.6 (a)]. Despite the tetrameric arrangement of both enzymes defined by 222-symmetry, their dimer-dimer interfaces are distinct because the extra helix in UGPase precludes the formation of the characteristic RmlA 4-helix bundle. As such, the dimer-dimer interface in UGPase is on the opposite face [Fig. 1.6 (b) and (c)] (119, 121). Finally, Glc-1-P cytidylyltransferase is distinct from RmlA and UGPase, functioning as a hexamer with an active site comprised of residues from two subunits. A glutamate and lysine that bind the NTP β -phosphorous and hexose in the product are donated by the second subunit, whereas the active site is fully enclosed in one monomer in RmlA and UGPase (119, 121, 122, 127).

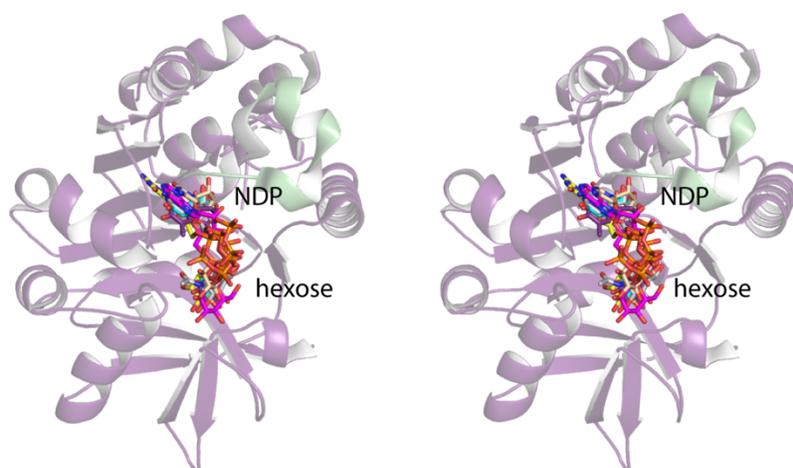
All sugar nucleotidyltransferases proceed, or are predicted to proceed, via a sequential Bi-Bi mechanism in which the NTP binds first (119, 122, 123, 130-132). The arginine residue in the conserved N-terminal loop pulls and strains the β - γ pyrophosphate leaving group on the NTP to make room for the hexose-1-P to attack the α -phosphate in an S_N2 type reaction (118-120, 122, 123, 128). Arginine and lysine residues, along with the required Mg^{2+} ions, stabilize the negative charge on the NTP and hexose-1-P, prime the hexose-1-P for nucleophilic attack on the NTP α -phosphate, and stabilize the transition state [Fig. 1.5 (c)] (118-120, 125, 127, 130, 133).

The Mg^{2+} ions are absolutely required for nucleotidyltransferase activity (100, 130). The Mg^{2+} ligated by the α and β phosphoryl oxygens and di-aspartate, or aspartate/asparagine pair in GlmU, has been observed in three-dimensional models of the NDP-hexose complex and/or NTP substrate complex in the dTTP, UTP, CTP, and GTP enzymes (120-123, 125, 127, 133). On the other hand, the cytidyltransferase enzyme is the only nucleotidyltransferase with structural data supporting the positions of both Mg^{2+} ions. The second Mg^{2+} appears to be positioned near the α,β -bridging oxygen and the γ -phosphoryl oxygen, supporting predictions that Mg^{2+} /pyrophosphate functions as a leaving group during catalysis (122). However, this mechanism is only hypothesized in the RmlA-type thymidyltransferases, as the second Mg^{2+} ion and the Mg^{2+} /dTDP-glucose complexes have never been observed (118, 119).

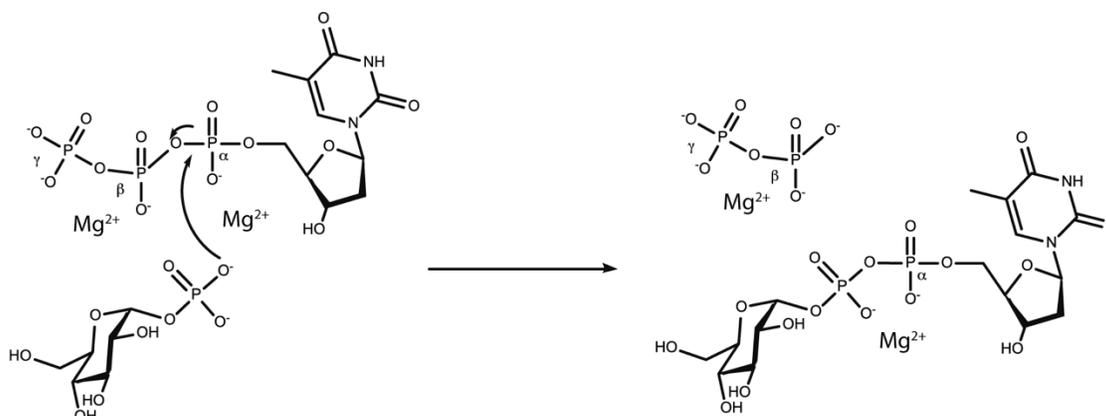
Given the role of the enzyme in the production of dTDP-linked deoxysugars deposited on natural products, the O-antigens, and capsules, such structural information is critical for drug design efforts against RmlA as well as re-design of the enzyme to generate novel natural products with desired biological activities (3, 4, 12, 14, 112). Furthermore, characterization of the RmlA from Mtb is especially critical because it is necessary for viability of the organism (82, 134).

<i>T. maritima</i> , GTP	-----MKALILAGGSGERFWPLSTPETPKQFLKLF	30
<i>P. aeruginosa</i> , dTTP	-----MKRKGIIILAGGSGTRLHPAT-LAISKQLLPVY	31
<i>C. glutamicum</i> , UTP	-----MSLPIDEHVNAVKTVVVPAAGLGTRFLPAT-KTVPKELLPVV	41
<i>S. pneumoniae</i> , GlnU	-----MSNFAIILAAGKGRMK--SDL--PKVLHKVA	28
<i>S. typhi</i> , CTP	-----MASKAVILAGGLGTRLSEET-IVKPKPMVEIG	31
<i>S. tuberosum</i> , ATP	MAVSDSQNSQTCLDPDASRSVLGIILGGGAGTRLYPLT-KKRAKPAVPLG	49

(a)

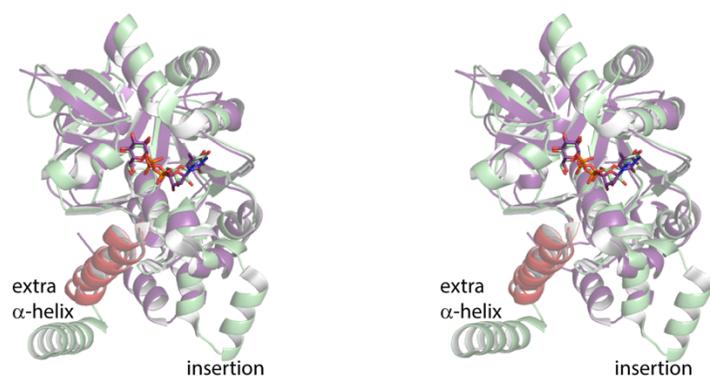


(b)

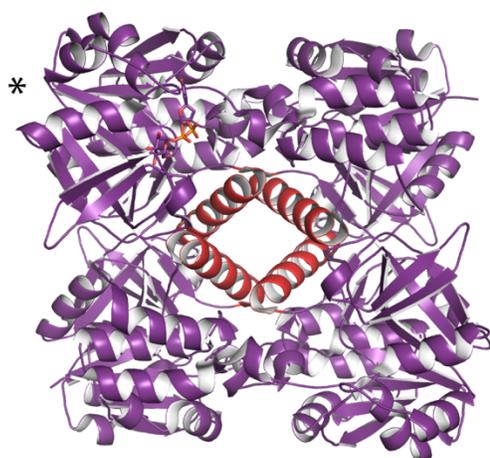


(c)

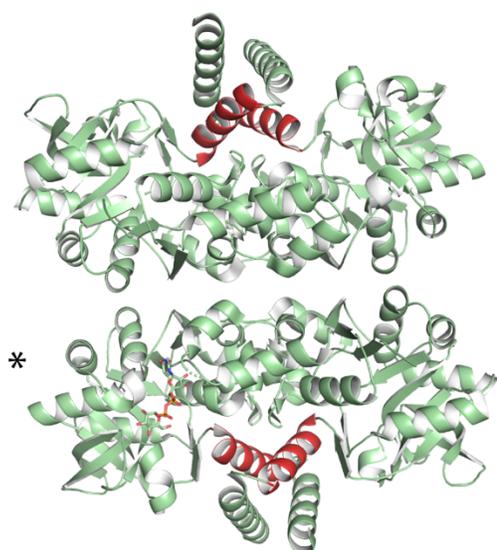
Figure 1.5 Common features of sugar nucleotidyltransferases. In part (a), primary sequence alignments of representative nucleotidyltransferases were constructed with Clustal Omega to highlight the (G)GXG(T)R(L)X₆₋₉(P)K loop, boxed in red. In (b), the nucleotidyltransferase domain of the enzymes from (a) were aligned and their NDP-hexoses are included in the context of the RmlA monomer, shown in stereo, with the flexible loop highlighted in green. dTDP-glucose is shown in purple, UDP-glucose in teal, ADP-glucose in magenta, GDP-mannose in yellow, CDP-glucose in wheat, and UDP-*N*-acetylglucosamine in wheat with PDB IDs 1G1L, 2PA4, 1YP4, 2X5Z, 1TZF, and 1G97 (119, 121, 123, 125, 127, 128). Finally, the proposed catalytic mechanism for the nucleotidyltransferases is provided in (c) using dTTP. This figure and Figures 1.6 and 1.7 were generated using PyMol (135).



(a)



(b)



(c)

Figure 1.6 Comparison of RmlA and UGPase structures. Depicted in (a) is a superposition of the RmlA and UGPase monomers shown in stereo. RmlA is highlighted in purple while UGPase is shown in green. The dTDP-glucose and UDP-glucose products are shown as sticks in the same color scheme. The α -helix implicated in the RmlA tetramer interface and the equivalent α -helix from UGPase are colored in red. Finally, the extra α -helix and insertion in UGPase are labeled. The RmlA tetramer is depicted in (b) while the UGPase tetramer is depicted in (c). The subunit used to align the tetramers is that containing the NDP-glucose and marked with an asterisk. Furthermore, the α -helices highlighted in red are the same as those highlighted in (a) to demonstrate the different oligomeric assembly between the enzymes.

1.5 4-Aminotransferases: dTDP-4-formamido-4,6-dideoxy-D-glucose and Rv3402c

Chapter 3 describes the purification and functional characterization of a predicted sugar aminotransferase, Rv3402c, from Mtb implicated in the biosynthesis of an unusual *N*-formylated sugar. Whereas these sugars are typically found in the O-antigens of Gram-negative bacteria, there are no reports in the literature of Mtb producing such a sugar (136). This section previews *N*-formylated sugars and the properties of sugar aminotransferases.

N-formylated sugars are found in the O-antigens of extremely pathogenic Gram-negative bacteria such as *Francisella tularensis*, *C. jejuni*, *Providencia alcalifaciens*, *S. enterica*, and species of *Brucella* (136). The accepted pathway for their production requires four or five steps, namely the attachment of glucose-1-phosphate to a nucleoside monophosphate, dehydration via a 4,6-dehydratase, 3,4-ketoisomerization in the case of the 3-formamido-3,6-dideoxy-D-pyranoses, transamination, and a formyltransfer reaction (136). Strikingly, structural and biochemical studies of a predicted sugar *N*-formyltransferase, Rv3404c, from Mtb recently revealed that the organism makes an enzyme capable of producing dTDP-4-formamido-4,6-dideoxy-D-glucose (137). However, there are no publications suggesting that Mtb utilizes this sugar at any point in its life cycle, begging the question as to whether or not Mtb contains all the enzymes necessary to make this formylated deoxysugar.

Mtb encodes a functional glucose-1-phosphate thymidyltransferase and 4,6-dehydratase to produce L-rhamnose, so the only enzyme required to complete the pathway is a 4-aminotransferase (138). Sugar aminotransferases have been extensively studied in the context of natural product and O-antigen unusual sugar biosynthesis (3, 4, 14). The aminosugar products of these enzymes are often subject to further modification by formylation, acylation, and methylation (3, 14, 18). Notably, deletion of the aminotransferase GDP-perosamine synthase, required for the production

of the O-antigen sugar perosamine, in *E. coli* O157:H7, reduces lethality in insect and mouse models of infection (61). Similarly, inactivation of a sugar *N*-formyltransferase in *B. abortus* attenuates pathogenicity of the organism (60).

Sugar aminotransferases belong to the pyridoxal 5'-phosphate (PLP) dependent Type 1 aspartate aminotransferase (AAT) superfamily and function as dimers (47, 139). The first structure of a sugar aminotransferase determined was that of ArnB, which is involved in the synthesis of L-Ara4N used to modify lipid A. The structural and kinetic analysis of ArnB placed sugar aminotransferases in the AAT superfamily, identified the lysine involved in internal aldimine formation, and led to glutamate being the proposed amino group donor (Scheme 1.1) (140). PseC, a 4-aminotransferase, was the first solved in the presence of its nucleotide-linked sugar product, whereas DesI, also a 4-aminotransferase, was the first crystallographic analysis to trap the external aldimine (Scheme 1.1) (141, 142).

Following these studies, the structures of a number of sugar aminotransferase were solved that led to speculation as to how these enzymes discern carbons 3 or 4 for amino group transfer and how they achieve stereochemical specificity. Coordinates for the following aminotransferases were superimposed and their external aldimine or PLP/NDP-hexoses are shown as sticks situated in the DesI active site in Figure 1.7: DesI (C-4' equatorial), PseC (C-4' axial), WlaRG (C-3' equatorial), WecE (C-4' axial), QdtB (C-3' equatorial), GDP-perosamine synthase (C-4' equatorial), and PglE (C-4' equatorial) (141-147). The characteristic AAT fold contains a 7-stranded mixed β -sheet flanked on each side by 3 α -helices. The position of the PLP cofactors is demonstrably invariant, however, the hexose, pyrophosphate, and nucleotide ring adopt different conformations, the most disparate being WecE (Fig. 1.7). Relative to C-4', C-3' recognition in QdtB and WlaRG is achieved via altered pyrophosphate and nucleotide binding, which positions

the hexose ring for catalysis at C-3'. Both can enzymes accept either C-4' substrate (axial or equatorial) due to a lack of interactions with the hydroxyl group at this position (143, 145). Notably, no aminotransferase catalyzing the formation of a C-3' axial sugar has ever been structurally characterized.

With respect to the synthesis of C-4' sugars, stereochemical specificity is dictated by the position of the hexose; compared with PseC (axial), the DesI (equatorial) sugar is flipped 180° in the active site due to a rotation of the pyrophosphoryl group (142). Conversely, WecE (axial), seems to control stereochemistry by positioning the nucleotide base and ribose such that the sugar is flipped relative to DesI (144). Intriguingly, PglE (equatorial) binds its nucleotide base differently than all other C-4' modifying enzymes yet manages to orient the hexose analogous to DesI (147).

Sugar aminotransferases utilize a ping-pong mechanism comprised of two half reactions (92, 139, 140). The accepted catalytic mechanism for a generic 4-aminotransferase is provided in Scheme 1.1. First, glutamate reacts with the internal aldimine to displace the PLP cofactor, yielding pyridoxamine 5'-phosphate (PMP) and α -ketoglutarate, which leaves the active site. Subsequently, the primary amine group on PMP attacks the 4-ketosugar, forming an external aldimine species that is resolved via the conserved lysine residue to form the internal aldimine and the NDP-4-amino-4,6-dideoxy-D-glucose (NDP-Qui4N) product (92).

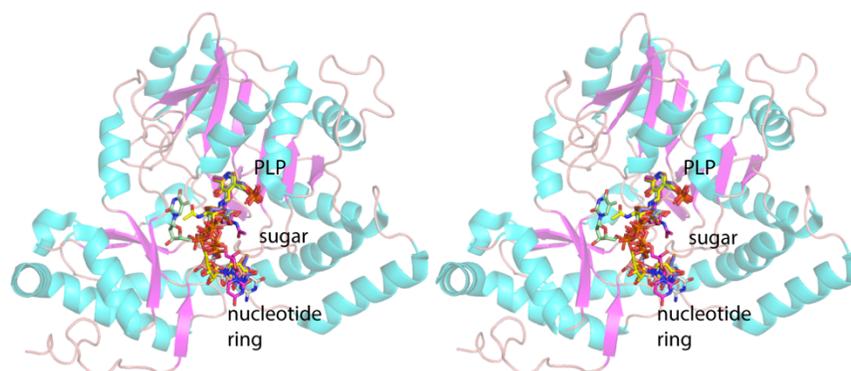
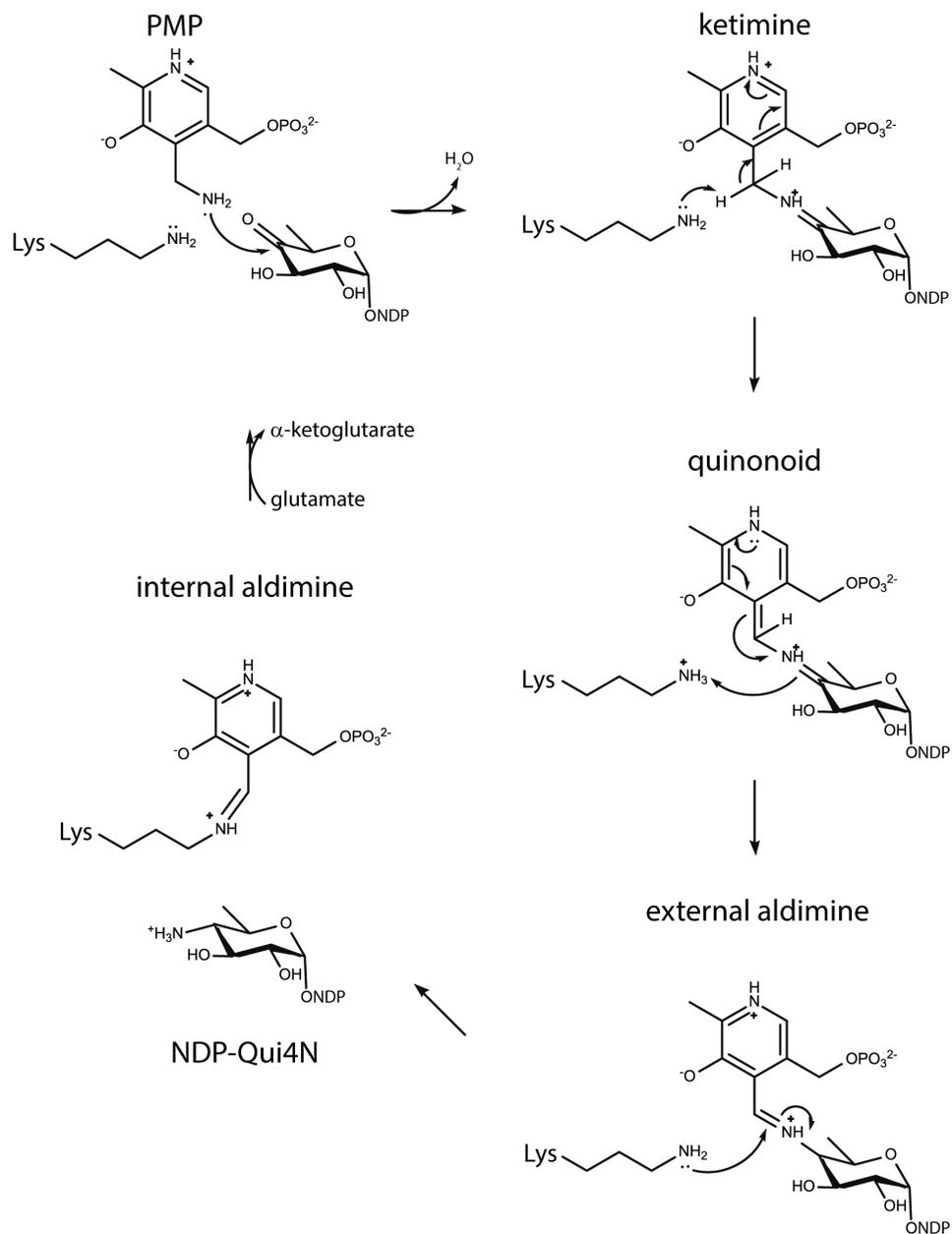


Figure 1.7 Architecture of the DesI monomer. The following aminotransferases solved in the presence of a nucleotide-linked sugar were aligned to highlight the conservation of the PLP binding site and the greater flexibility in the nucleotide-linked sugar binding sites: DesI (salmon), PglE (magenta), PseC (yellow), WecE (green), GDP-perosamine synthase (grey), WlaRG (slate), and QdtB (orange), PDB IDs: 2PO3, 4ZTC, 2FNU, 4ZAH, 3DR4, 5U23, and 3FRK (141-147). The sugar nucleotides are positioned with reference to DesI, whose α -helices are colored in blue, β -strands in magenta, and loops in wheat.



Scheme 1.1 Catalytic mechanism of a 4-aminotransferase.

1.6 3-Monoepimerases: CDP-6-deoxy-D-gulose and WbcA

Chapter 4 encompasses the characterization of an enzyme involved in O-polysaccharide biosynthesis, namely that of WbcA. This enzyme is a predicted C-3' monoepimerase required for the production of CDP-6-deoxy-D-gulose, a rare deoxysugar found in the O-antigen of *Yersinia enterocolitica* O:8 (148). Primary sequence alignments suggested that the active site of WbcA differs substantially from other sugar monoepimerases, prompting a structural investigation of the enzyme. Background regarding the sugar epimerases is provided below.

Two of the most well characterized enzymes in this class include UDP-galactose-4-epimerase (GalE) and RmlC, involved in the Leloir pathway of galactose metabolism and L-rhamnose production, respectively (149, 150). GalE does not function on a deoxysugar but catalyzes epimerization about C-4' and belongs to the NAD⁺-dependent short-chain dehydrogenase/reductase (SDR) superfamily whereas RmlC catalyzes epimerization about C-3' and C-5' and is a member of the cupin superfamily (151, 152). The GalE mechanism invokes hydride transfer and a keto-intermediate while RmlC works on a ketosugar substrate and proceeds via a proton abstraction/elimination mechanism (153-156). In GDP-mannose derived sugar synthesis, the SDR enzyme GDP-fucose synthase catalyzes a combination 3,5-diepimerization and 4-ketoreduction; thus its NADPH is a co-substrate, unlike GalE which uses NAD⁺ catalytically (157). GDP-mannose-3,5-epimerase mimics the overall diepimerization reaction of RmlC but is an SDR enzyme that more closely resembles how GalE utilizes NAD⁺ (158).

GalE has been studied extensively in part due to its role in certain presentations of galactosemia (149). Conversely, L-rhamnose is a common component of bacterial O-antigens and an essential component of the Mtb cell wall, thus RmlC has been extensively characterized for its potential as an antibiotic target (14, 138). Some O-antigen and natural product sugars only require

monoepimerization about C-3' or C-5' for production (Fig. 1.8) (3, 13, 14). The prokaryotic 3,(5)-monoepimerases all appear to belong to the cupin superfamily of enzymes which harbor a characteristic “small barrel”, or in Latin, a “*cupa*”, the namesake of the superfamily (3, 159). Cupin enzymes demonstrate broad biochemical activity; in addition to their role in unusual sugar biosynthesis in prokaryotes, they function as transcription factors in mammals and storage proteins in plant seeds (159).

The sugar-modifying cupins work on a variety of substrates. For example, the enzyme AscE catalyzes epimerization about C-5' in the production of CDP-ascarylose, a 3,6-dideoxysugar destined for the O-antigens in species of *Yersinia* (5, 160). Similarly, EvaD catalyzes epimerization about C-5' in the biosynthesis dTDP-epivancosamine during vancomycin production (161). C-3' monoepimerases such as ChmJ and NovW are required for the bioproduction of the natural products chalcomycin and novobiocin, respectively (162-164). A structural and kinetic analysis of ChmJ confirmed that a conserved histidine base and tyrosine acid required for RmlC and EvaD activity are required for C-3' epimerization (161, 164, 165). However, the ChmJ study questioned the notion that the position of the tyrosine is responsible for discerning between C-3' and/or C-5' epimerization and rather that its orientation is simply a result of ligand binding (164). Thus, X-ray structures alone cannot be used to assign function in the case of the cupin sugar epimerases.

Nonetheless, because of their role in O-antigen sugar production, as well as lack of human homologues, these enzymes are suitable targets for the development of new antibiotics. Indeed, RmlC deletion in Mtb is lethal and leads to attenuated virulence and infection initiation in *P. aeruginosa* and *Streptococcus mutans*, respectively (166-168).

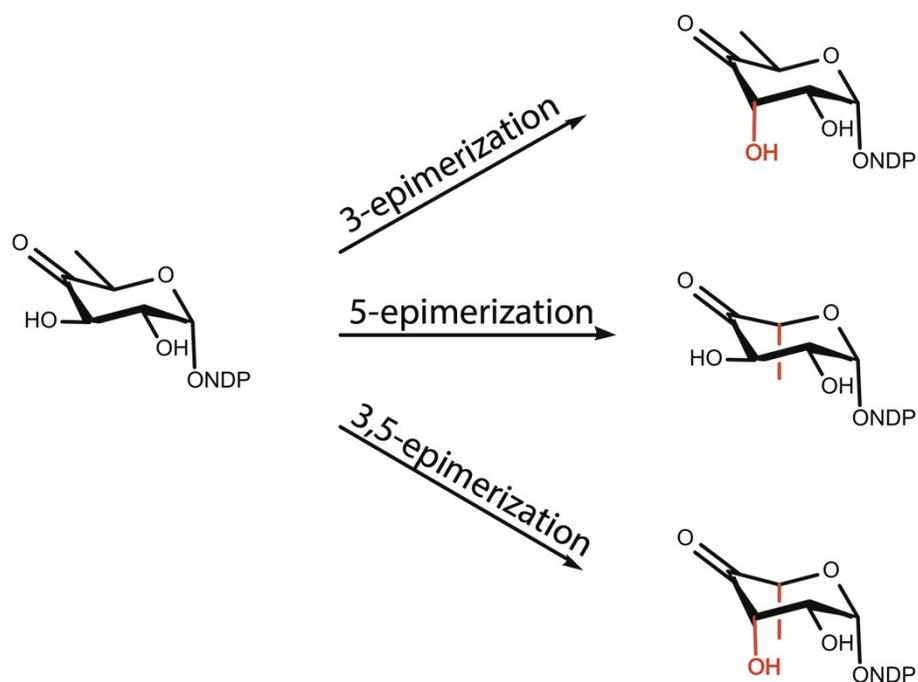


Figure 1.8 C-3' and C-5' epimerization reactions. While the enzymes in the text work on structurally more complex sugars than those depicted here, this figure is meant to distinguish the different types of epimerases. Some invert the stereochemistry about one carbon, either the hydroxyl group on C-3' or the methyl group on C-5'. Others invert the stereochemistry around both atoms.

1.7 Biosynthesis of the *O*-Methyl Phosphoramidate Capsular Polysaccharide Modification in *C. jejuni*

Chapters 5 and 6, as well as Appendices 1 and 2, describe functional and structural characterizations of enzymes involved in the bioproduction of a novel *O*-methyl phosphoramidate (MeOPN) moiety found on the capsular polysaccharides of *C. jejuni*. The modification is found nowhere else in nature, suggesting that unique enzymes may be required for its synthesis (169). Furthermore, because MeOPN presence promotes *C. jejuni* infectivity, these potentially unique enzymes may be attractive targets for drug design (170). Background information on the biological roles of MeOPN and a potential pathway for its production are provided below.

C. jejuni is one of the most common causes of gastroenteritis worldwide, with contaminated poultry being a frequent source (171). In most diagnoses of the autoimmune disorder, Guillain-Barré Syndrome (GBS), the antecedent illness is a *C. jejuni* infection due LOS mimicry of host gangliosides (172, 173). The phase-variable LOS contributes to pathogenicity and immune avoidance, however, the extensive glycosylation systems beyond the LOS enhance the organism's pathogenicity (98, 174). Indeed, prokaryotic *O*- and *N*-linked glycosylation were first discovered in the organism (175, 176). A fourth system for CPS production, the major serotype determinant, is responsible for invasion of human cells and disease in animal models (177, 178).

The first *C. jejuni* genome sequenced led to the proposal that the high molecular weight polysaccharide used for serotyping was a capsule, not the assumed lipopolysaccharide (178, 179). Like other capsule producing species, rough strains of *C. jejuni* tend to be less virulent (177, 180-182). Furthermore, serogroups, as well subpopulations thereof, are heterogeneous due to phase variation of the CPS and its modifications (169, 179, 183-185). Unusual non-carbohydrate moieties decorating the capsule include 2-amino-2-deoxyglycerol, methyl groups, *N*-ethanolamine, and the unprecedented MeOPN (169, 177). Remarkably, 75% of *Campylobacters*

attach MeOPN to sugars in the CPS (97). The modification serves as a bacteriophage receptor and is responsible for serum resistance, colonization of certain animals, and may be a source of antigenicity in conjugate vaccine trials (170, 177, 182, 186, 187). Two genes predicted to transfer MeOPN to its target sugars are phase-variable, suggesting that fine-tuned attachment of this modification is critical (177, 179, 183, 185). Indeed, in strain 81-176, MeOPN at one position enhances resistance to serum killing whereas its presence at other positions promotes sensitivity to complement (184). The unique chemical group also has immunomodulatory effects, and therefore represents a remarkable example of how extensive adaptation and heterogeneity promote pathogen fitness (98, 177, 187, 188).

Furthermore, the incidence of naturally occurring phosphoramidates is quite rare, thus the enzyme chemistry implicated in P-N bond formation in MeOPN may be harnessed to synthesize novel, biologically active compounds. Two well-known examples of P-N bonds include the phosphagens, used in the rapid restoration of ATP levels in muscle and brain, and the sensor histidine kinases, which are two-component regulatory systems in bacteria, fungi, archaea, and plants (189, 190). Furthermore, DNA ligase requires an adenylated lysine intermediate during catalysis, an example of a transient P-N bond (191). Less well-known cases are depicted in Figure 1.9 and include a number of bacterially derived natural products as well as synthetic compounds like the insecticide Fenamiphos, the GTP mimic guanosine 5'-[β,γ -imido] triphosphate (GMP-PNP) employed frequently in crystallography experiments, the morpholino oligomer Eteplirsen used to treat Duchenne Muscular Dystrophy, and Stampidine which is in clinical trials as a less toxic human immunodeficiency virus (HIV) nucleoside reverse transcriptase inhibitor (192-201).

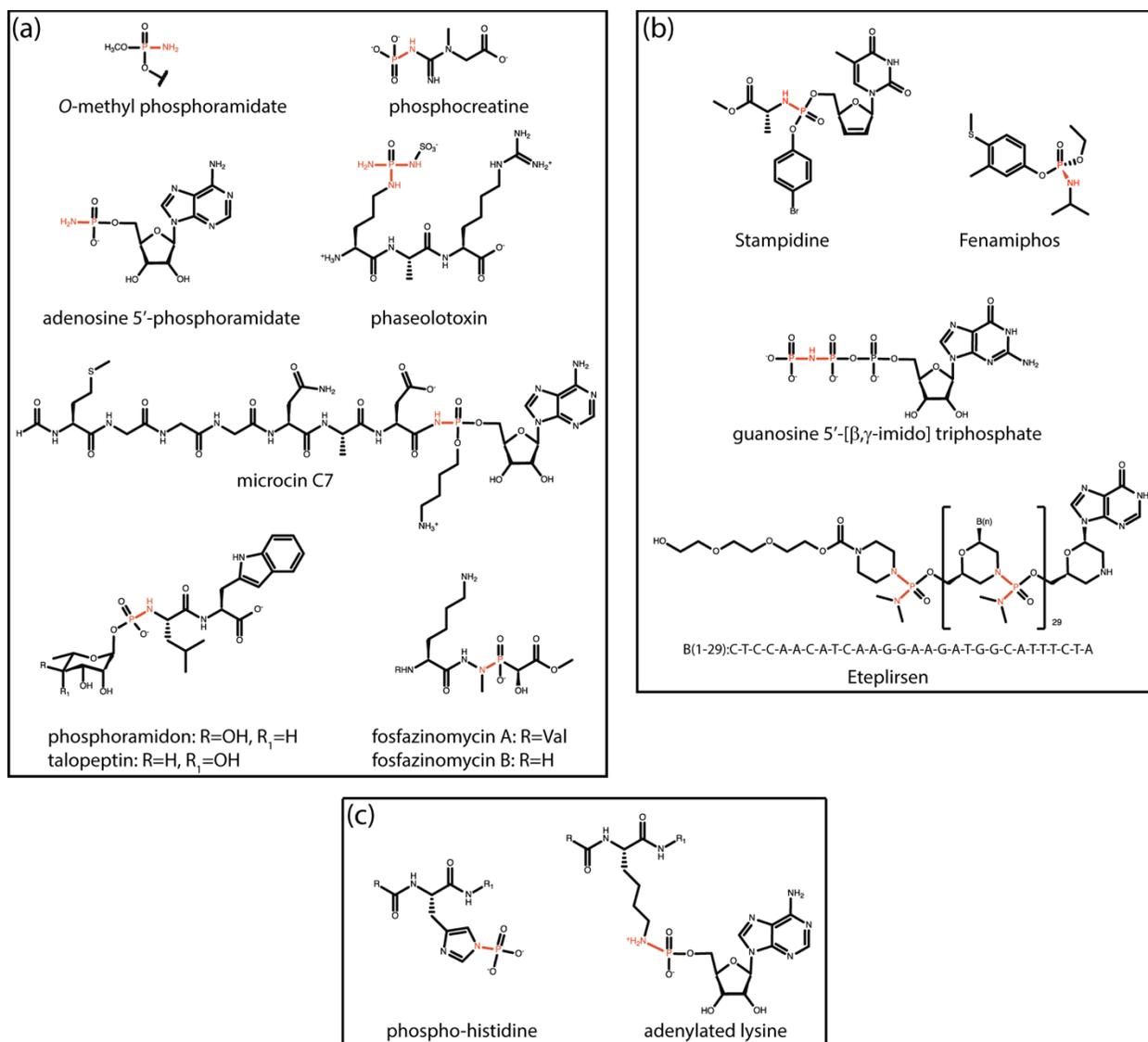


Figure 1.9 Selected P-N bond containing molecules. Naturally occurring compounds are shown in part (a). Panel (b) contains the structures of synthetic examples detailed in the text. P-N bonds are also implicated in the mechanisms of certain enzymes. Examples of histidine and lysine are shown in (c). The P-N bond is colored in red in each panel.

A functional and structural understanding of MeOPN biosynthetic enzymes is merited for two reasons: 1) its potential utility in the production of novel phosphoramidate compounds; 2) targeting the pathway to control *C. jejuni* infectivity. Based on amino acid sequence alignments, a prior study demonstrating MeOPN incorporation of ^{15}N from labeled NH_4Cl in the growth media, and knockout strains that no longer make MeOPN, the enzyme activities shown in Figure 1.10 were proposed for the synthesis of the unique modification (97, 202). Cj1418 is predicted to catalyze the formation of phosphoenolpyruvate (PEP) because of its similarity to the PEP synthetases and pyruvate, phosphate dikinases (203, 204). Cj1417 is thought to produce phosphoramidate due to its annotation as a type 1 glutamine-dependent amidotransferase (205). The phosphoramidate might then be appended to a NMP by the predicted nucleotidyltransferase Cj1416 based on its similarity to LicC, which transfers phosphocholine to cytidine monophosphate (206). The nucleoside diphosphoramidate may subsequently be phosphorylated by Cj1415, annotated as an adenosine 5'-phosphosulfate kinase which phosphorylates the C-3' hydroxyl in the ribose ring of its substrate (207).

Cj1419 and Cj1420 are predicted SAM dependent *O*-methyltransferases, and Cj1421 and Cj1422 are the proposed MeOPN transferases that deposit the moiety onto Gal β NAc and D-*glycero*- α -L-*gluco*-hep, respectively (Fig. 1.10) (97, 169). Interestingly, all presumed MeOPN transferases have conserved N-termini, whereas the more disparate C-termini are thought to recognize the sugar being modified (177, 179, 184, 185). Rationale for the presence of two methyltransferases is unclear and complicated by single gene knockouts not resulting in loss of a MeOPN signal via nuclear magnetic resonance (NMR) spectroscopy (97). In addition to *cj1421c* and *cj1422c*, *cj1420c* is subject to phase variation, suggesting that producing variably methylated MeOPN may be advantageous (97). Finally, the substrates for the MeOPN transferases are unknown

(nucleotide-linked sugars or those already in the capsule) as are the substrates for the methyltransferases (Cj1415 product, nucleotide-linked sugar-phosphoramidate, or capsule sugar-phosphoramidate) (Fig. 1.10). A thorough understanding of the temporal and spatial requirements of these reactions necessitates an understanding of how the likely phosphoramidate substrate is synthesized in the first place. Thus, the focus of the work in this thesis is on the activities of Cj1415-Cj1418.

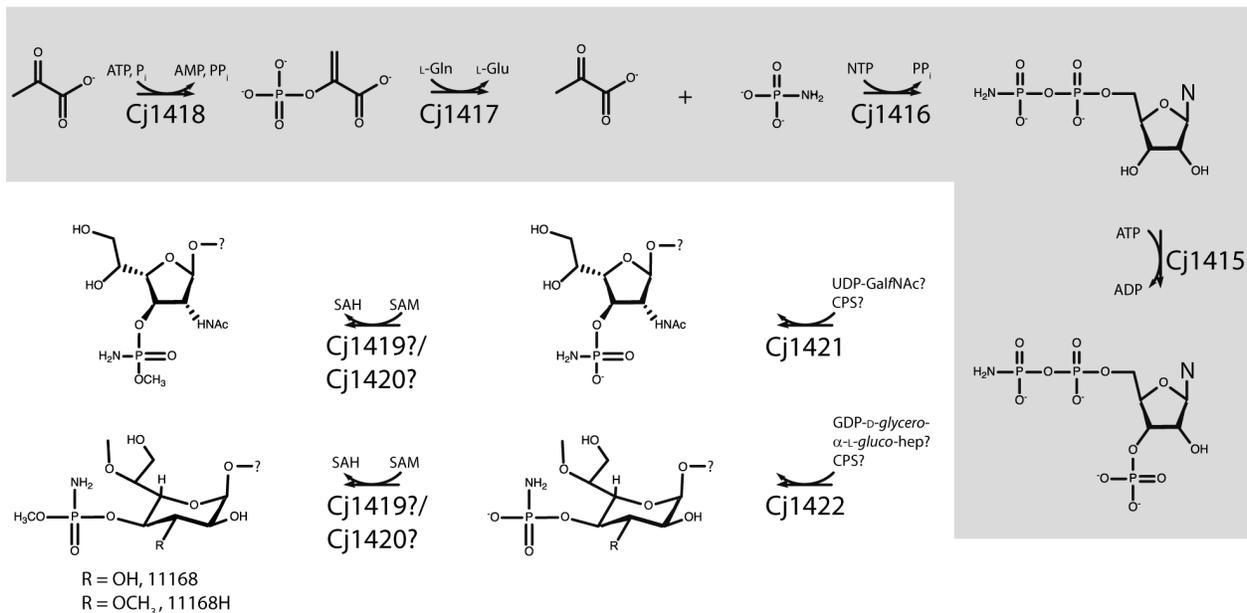


Figure 1.10 Possible MeOPN production pathway in *C. jejuni*. Putative functions of Cj1415-Cj1421 are described in the text. While enzymes responsible for UDP-GalfNAc and GDP-D-glycero- α -L-gluco-hep production are known, the exact timing of phosphoramidate and *O*-methyl group transfer is unclear (185, 208). Enzymes highlighted in grey, Cj1415-Cj1418, are the focuses of Chapter 5, Chapter 6, Appendix 1, and Appendix 2.

1.8 Project CRYSTAL

In a 2009 speech to the National Academy of Sciences, President Barack Obama declared that via implementation of new funding programs, “American students will move from the middle to the top of the pack in science and math over the next decade.” It was apt timing, then, that in 2009, the Holden Laboratory launched a partnership with local Madison middle schools, now referred to as Project CRYSTAL (Colleagues **R**esearching with **Y**oung Scientists: **T**eaching and **L**earning). The program was started in order to expose gifted students to laboratory research at a young age given that, historically, interest in science wanes as they progress through school. Indeed, if students are captivated by the topic early on, they are more likely to “seek experiences that prepare them for future encounters with science ” (209).

Six years after President Obama’s declaration, unfortunately, the United States’ standing with respect to gains in science education were lacking. The 2015 results of the Programme for International Student Assessment (PISA) ranked the US 38th and 24th out of 71 countries in math and science, respectively. Given that a significant amount of learning in the K-12 stage occurs outside the traditional classroom, Project CRYSTAL is in a unique and important position to not only foster interest in science, but to heighten student’s academic achievement.

Project CRYSTAL accomplishes this by working with four to six students every week in the laboratory and classroom, teaching them topics from molecular cloning to building a three-dimensional model of a protein. While the program was originally designed to involve middle schoolers with Holden laboratory research on deoxysugar biosynthetic enzymes, Chapter 7 will document the transition from this manner of teaching to one that can be distributed to laboratories across the country looking to positively impact young students. The benefits of this program with respect to the graduate and undergraduate mentors will also be discussed.

1.9 References

1. Voet, D., Voet, J. G., and Pratt, C. W., *Fundamentals of Biochemistry*. 3 ed. 2008: John Wiley & Sons, Inc.
2. Poole, A., Penny, D., and Sjöberg, B. M. (2000) Methyl-RNA: an evolutionary bridge between RNA and DNA? *Chem. Biol.* 7, R207-R216.
3. Thibodeaux, C. J., Melançon, C. E., and Liu, H. W. (2008) Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed. Engl.* 47, 9814-9858.
4. Weymouth-Wilson, A. C. (1997) The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* 14, 99-110.
5. Liu, H. W. and Thorson, J. S. (1994) Pathways and mechanisms in the biogenesis of novel deoxysugars by bacteria. *Annu. Rev. Microbiol.* 48, 223-256.
6. He, X. M. and Liu, H. W. (2002) Formation of unusual sugars: mechanistic studies and biosynthetic applications. *Annu. Rev. Biochem.* 71, 701-754.
7. Lowe, J. B. (1993) The blood group-specific human glycosyltransferases. *Baillière's Clin. Haematol.* 6, 465-492.
8. Yazer, M. H. (2005) What a difference 2 nucleotides make: a short review of ABO genetics. *Transfus. Med. Rev.* 19, 200-209.
9. Holden, H. M., Cook, P. D., and Thoden, J. B. (2010) Biosynthetic enzymes of unusual microbial sugars. *Curr. Opin. Struct. Biol.* 20, 543-550.
10. Cummings, R. D. (2009) The repertoire of glycan determinants in the human glycome. *Mol. Biosyst.* 5, 1087-1104.
11. Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* 13, 448-462.
12. Thibodeaux, C. J., Melançon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature.* 446, 1008-1016.
13. Singh, S., Phillips, G. N., and Thorson, J. S. (2012) The structural biology of enzymes involved in natural product glycosylation. *Nat. Prod. Rep.* 29, 1201-1237.
14. Hao, Y. and Lam, J. S., *Pathways for the biosynthesis of NDP-sugars*, in *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Y.A. Knirel and M.A. Valvano, Editors. 2011, Springer. p. 195-235.
15. Uy, N., Nadeau, M., Stahl, M., and Zeidan, A. M. (2018) Inotuzumab ozogamicin in the treatment of relapsed/refractory acute B cell lymphoblastic leukemia. *J. Blood Med.* 9, 67-74.
16. Raetz, C. R. H. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* 71, 635-700.
17. Roberts, I. S. (1996) The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annu. Rev. Microbiol.* 50, 285-315.
18. Knirel, Y. A., *Structure of O-antigens*, in *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Y.A. Knirel and M.A. Valvano, Editors. 2011, Springer. p. 41-115.
19. Fagan, R. P. and Fairweather, N. F. (2014) Biogenesis and function of bacterial S-layers. *Nat. Rev. Microbiol.* 12, 211-222.
20. Daffé, M., Crick, D. C., and Jackson, M. (2014) Genetics of capsular polysaccharide and cell envelope (glyco)lipids. *Microbiol. Spectrum.* 2, MGM2-0021-2013.

21. *Vaccines and Immunizations Glossary*. 2015; Available from: <https://www.cdc.gov/vaccines/terms/glossary.html>.
22. Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature*. *486*, 207-214.
23. Ribet, D. and Cossart, P. (2015) How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes. Infect.* *17*, 173-183.
24. McGuckin, M. A., Lindén, S. K., Sutton, P., and Florin, T. H. (2011) Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* *9*, 265-278.
25. Wolf, A. J. and Underhill, D. M. (2018) Peptidoglycan recognition by the innate immune system. *Nat. Rev. Immunol.* *18*, 243-254.
26. Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J. D. (2010) Complement: a key system for immune surveillance and homeostasis. *Nat. Immunol.* *11*, 785-797.
27. Cochet, F. and Peri, F. (2017) The role of carbohydrates in the lipopolysaccharide (LPS)/toll-like receptor 4 (TLR4) signalling. *Int. J. Mol. Sci.* *18*, E2318.
28. Vollmer, W., Blanot, D., and de Pedro, M. A. (2008) Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* *32*, 149-167.
29. Alderwick, L. J., Harrison, J., Lloyd, G. S., and Birch, H. L. (2015) The mycobacterial cell wall-peptidoglycan and arabinogalactan. *Cold Spring Harb. Perspect. Med.* *5*, a021113.
30. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) Lipid A modification systems in gram-negative bacteria. *Annu. Rev. Biochem.* *76*, 295-329.
31. Cunneen, M. M. and Reeves, P. R., *Evolution of lipopolysaccharide biosynthesis genes*, in *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Y.A. Knirel and M.A. Valvano, Editors. 2011, Springer. p. 339-370.
32. Mamat, U., Skurnik, M., and Bengoechea, J. A., *Lipopolysaccharide core oligosaccharide biosynthesis and assembly*, in *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*. 2011, Springer. p. 237-274.
33. Yethon, J. A., Gunn, J. S., Ernst, R. K., Miller, S. I., Laroche, L., Malo, D., and Whitfield, C. (2000) *Salmonella enterica* serovar Typhimurium *waaP* mutants show increased susceptibility to polymyxin and loss of virulence *in vivo*. *Infect. Immun.* *68*, 4485-4491.
34. Walsh, A. G., Matewish, M. J., Burrows, L. L., Monteiro, M. A., Perry, M. B., and Lam, J. S. (2000) Lipopolysaccharide core phosphates are required for viability and intrinsic drug resistance in *Pseudomonas aeruginosa*. *Mol. Microbiol.* *35*, 718-727.
35. Regué, M., Izquierdo, L., Fresno, S., Piqué, N., Corsaro, M. M., Naldi, T., De Castro, C., Waidelich, D., Merino, S., and Tomás, J. M. (2005) A second outer-core region in *Klebsiella pneumoniae* lipopolysaccharide. *J. Bacteriol.* *187*, 4198-4206.
36. Griffiss, J. M. and Schneider, H., *The chemistry and biology of lipooligosaccharides: the endotoxins of bacteria of the respiratory and genital mucosae*, in *Endotoxin in Health and Disease*, H. Brade, et al., Editors. 1999, Marcel Dekker. p. 179-194.
37. Apicella, M. A. and Jennings, M. P., *Phase variation of bacterial surface glycosylated molecules in immune evasion*, in *Microbial Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier. p. 837-846.
38. Avery, S. V. (2006) Microbial cell individuality and the underlying sources of heterogeneity. *Nat. Rev. Microbiol.* *4*, 577-587.
39. Moran, A. P., *Molecular mimicry of host glycosylated structures by bacteria*, in *Microbial Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier. p. 847-870.

40. O’Riordan, K. and Lee, J. C. (2004) *Staphylococcus aureus* capsular polysaccharides. Clin. Microbiol. Rev. 17, 218-234.
41. Whitfield, C. and Roberts, I. S. (1999) Structure, assembly, and regulation of expression of capsules in *Escherichia coli*. Mol. Microbiol. 31, 1307-319.
42. Cescutti, P., *Bacterial capsular polysaccharides and exopolysaccharides*, in *Microbiol Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier. p. 93-108.
43. Weintraub, A. (2003) Immunology of bacterial polysaccharide antigens. Carbohydr. Res. 338, 2539-2547.
44. Tzeng, Y. L., Thomas, J., and Stephens, D. S. (2016) Regulation of capsule in *Neisseria meningitidis*. Crit. Rev. Microbiol. 42, 759-722.
45. Szu, S. C., Li, X. R., Stone, A. L., and Robbins, J. B. (1991) Relation between structure and immunologic properties of the Vi capsular polysaccharide. Infect. Immun. 59, 4555-4561.
46. Bayliss, M., Donaldson, M. I., Nepogodiev, S. A., Pergolizzi, G., Scott, A. E., Harmer, N. J., Field, R. A., and Prior, J. L. (2017) Structural characterization of the capsular polysaccharide expressed by *Burkholderia thailandensis* strain E555:: *wbiI* (pKnock-KmR) and assessment of the significance of the 2-O-acetyl group in immune protection. Carbohydr. Res. 452, 17-24.
47. Samuel, G. and Reeves, P. (2003) Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr. Res. 338, 2503-2519.
48. Skurnik, M., Venho, R., Bengoechea, J. A., and Moriyon, I. (1999) The lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3 is required for virulence and plays a role in outer membrane integrity. Mol. Microbiol. 31, 1443-1462.
49. Bengoechea, J. A., Najdenski, H., and Skurnik, M. (2004) Lipopolysaccharide O antigen status of *Yersinia enterocolitica* O:8 is essential for virulence and absence of O antigen affects the expression of other *Yersinia* virulence factors. Mol. Microbiol. 52, 451-469.
50. Morona, R., Daniels, C., and Van Den Bosch, L. (2003) Genetic modulation of *Shigella flexneri* 2a lipopolysaccharide O antigen modal chain length reveals that it has been optimized for virulence. Microbiology. 149, 925-939.
51. O’Donoghue, E. J., Sirisaengtaskin, N., Browning, D. F., Bielsak, E., Hadis, M., Fernandez-Trillo, F., Alderiwck, L., Jabbari, S., and Krachler, A. M. (2017) Lipopolysaccharide structure impacts the entry kinetics of bacterial outer membrane vesicles into host cells. PLoS Pathog. 13, e1006760.
52. Nesper, J., Lauriano, C. M., Klose, K. E., Kapfhammer, D., Kraib, A., and Reidl, J. (2001) Characterization of *Vibrio cholerae* O1 E1 Tor *galU* and *galE* mutants: influence on lipopolysaccharide structure, colonization, and biofilm formation. Infect. Immun. 69, 435-445.
53. Kingsley, R. A. and Bäumler, A. J. (2000) Host adaptation and the emergence of infectious disease: the *Salmonella* paradigm. Mol. Microbiol. 36, 1006-1014.
54. Lindberg, A. A., Segall, T. ., Weintraub, A., and Stocker, B. A. D. (1993) Antibody response and protection against challenge in mice vaccinated intraperitoneally with a live *aroA* O4-09 hybrid *Salmonella dublin* strain. Infect. Immun. 61, 1211-1221.

55. Bushmarinov, I. S., Ovchinnikova, O. G., Kocharova, N. A., Toukach, F. V., Torzewska, A., Shashkov, A. S., Knirel, Y. A., and Rozalski, A. (2007) Structure of the O-polysaccharide and serological cross-reactivity of the lipopolysaccharide of *Providencia alcalifaciens* O32 containing *N*-acetylismuramic acid. *Carbohydr. Res.* 342, 268-273.
56. Knirel, Y. A., *O-specific polysaccharides of gram-negative bacteria*, in *Microbiol Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier. p. 57-74.
57. Lüderitz, O., Staub, A. M., and Westphal, O. (1966) Immunochemistry of O and R antigens of *Salmonella* and related *Enterobacteriaceae*. *Bacteriol. Rev.* 30, 192-255.
58. Rittig, M. G., Kaufmann, A., Robins, A., Shaw, B., Sprenger, H., Gemsa, D., Foulongne, V., Rouot, B., and Dornand, J. (2003) Smooth and rough lipopolysaccharide phenotypes of *Brucella* induce different intracellular trafficking and cytokine/chemokine release in human monocytes. *J. Leukoc. Biol.* 74, 1045-1055.
59. Darwin, A. J. and Miller, V. L. (1999) Identification of *Yersinia enterocolitica* genes affecting survival in an animal host using signature-tagged transposon mutagenesis. *Mol. Microbiol.* 32, 51-62.
60. Lacerda, T. L. S., Cardoso, P. G., Augusto De Almeida, L., da Cunha Camarg, I. L. B., Afonso, D. A. F., Trant, C. C., Macedo, G. C., Campos, E., Cravero, S. L., Salcedo, S. P., Gorvel, J. P., and Oliveira, S. C. (2010) Inactivation of formyltransferase (wbkc) gene generates a *Brucella abortus* rough strain that is annuated in macrophages and in mice. *Vaccine.* 28, 5627-5634.
61. Miyashita, A., Iyoda, S., Ishii, K., Hamamoto, H., Sekimizu, K., and Kaito, C. (2012) Lipopolysaccharide O-antigen of enterohemorrhagic *Escherichia coli* O157:H7 is required for killing both insects and mammals. *FEMS Microbiol. Lett.* 333, 59-68.
62. Liang-Takasaki, C. J., Makela, H., and Leive, L. (1982) Phagocytosis of bacteria by macrophages: changing the carbohydrate of lipopolysaccharide alters interactions with complement and macrophages. *J. Immunol.* 128, 1229-1235.
63. Hone, D. M., Harris, A. M., Lim, V., and Levin, M. M. (1994) Construction and characterization of isogenic O-antigen variants of *Salmonella typhi*. *Mol. Microbiol.* 13, 525-530.
64. Cygler, M., Rose, D. R., and Bundle, D. R. (1991) Recognition of a cell-surface oligosaccharide of pathogenic *Salmonella* by an antibody Fab fragment. *Science.* 253, 442-445.
65. Gening, M. L., Maira-Litrán, T., Kropec, A., Skurnik, D., Grout, M., Tsvetkov, Y. E., Nifantiev, N. E., and Pier, G. B. (2010) Synthetic β -(1 \rightarrow 6)-linked *N*-acetylated and nonacetylated oligoglucosamines used to produce conjugate vaccines for bacterial pathogens. *Infect. Immun.* 78, 764-772.
66. Geno, K. A., Gilbert, G. L., Song, J. Y., Skovsted, I. C., Skovsted, K. C., Klugman, K. P., Jones, C., Kondrasen, H. B., and Nahm, M. H. (2015) Pneumococcal capsules and their types: past, present, and future. *Clin. Microbiol. Rev.* 28, 871-899.
67. Forrellad, M. A., Klepp, L. I., Gioffré, A., Sabio y Garcia, J., Morbidoni, H. R., de la Paz Santangelo, M., Cataldi, A. A., and Bigi, F. (2013) Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence.* 4, 1-64.
68. Jankute, M., Cox, J. A. G., Harrison, J., and Besra, G. S. (2015) Assembly of the mycobacterial cell wall. *Annu. Rev. Microbiol.* 69, 405-423.

69. Ernst, J. D. (2012) The immunological life cycle of tuberculosis. *Nat. Rev. Immunol.* *12*, 581-591.
70. Sambou, T., Dinadayala, P., Stadthagen, G., Barilone, N., Bordat, Y., Constant, P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A., Daffé, M., and Jackson, M. (2008) Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on persistence in mice. *Mol. Microbiol.* *70*, 762-774.
71. Geurtsen, J., Chedammi, S., Mesters, J., Cot, M., Driessen, N. N., Sambou, T., Kakutani, R., Ummels, R., Maaskant, J., Takata, H., Baba, O., Terashima, T., Bovin, N., Vandenbroucke-Grauls, C. M. J. E., Nigou, J., Puzo, G., Lemassu, A., Daffé, M., and Appelmek, B. J. (2009) Identification of mycobacterial α -glucan as a novel ligand for DC-SIGN: involvement of mycobacterial capsular polysaccharides in host immune modulation. *J. Immunol.* *183*, 5221-5231.
72. Cywes, C., Hoppe, H. C., Daffé, M., and Ehlers, M. R. W. (1997) Nonopsonic binding of *Mycobacterium tuberculosis* to complement receptor type 3 is mediated by capsular polysaccharides and is strain dependent. *Infect. Immun.* *65*, 4258-4266.
73. Singh, A., Crossman, D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B., and Steyn, A. J. (2009) *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog.* *5*, e1000545.
74. Lee, W., Vandervan, B. C., Fahey, R. J., and Russell, D. G. (2013) Intracellular *Mycobacterium tuberculosis* exploits host-derived fatty acids to limit metabolic stress. *J. Biol. Chem.* *288*, 6788-6800.
75. Jackson, M., Stadthagen, G., and Gicquel, B. (2007) Long-chain multiple methyl-branched fatty acid-containing lipids of *Mycobacterium tuberculosis*: biosynthesis, transport, regulation and biological activities. *Tuberculosis.* *87*, 78-86.
76. Lee, K. S., Dubey, V. S., Kolattukudy, P. E., Song, C. H., Shin, A. R., Jung, S. B., Yang, C. S., Kim, S. Y., Jo, E. K., Park, J. K., and Kim, H. J. (2007) Diacyltrehalose of *Mycobacterium tuberculosis* inhibits lipopolysaccharide- and mycobacteria-induced proinflammatory cytokine production in human monocytic cells. *FEMS Microbiol. Lett.* *267*, 121-128.
77. Blanc, L., Gilleron, M., Prandi, J., Song, O. R., Jang, M. S., Gicquel, B., Drocourt, D., Neyrolles, O., Brodin, P., Tiraby, G., Vercellone, A., and Nigou, J. (2017) *Mycobacterium tuberculosis* inhibits human innate immune responses via the production of TLR2 antagonist glycolipids. *Proc. Natl. Acad. Sci. USA.* *114*, 11205-11210.
78. Brodin, P., Poquet, Y., Levillain, F., Peguillet, I., Larrouy-Maumus, G., Gilleron, M., Ewann, F., Christophe, T., Fenistein, D., Jang, J., Jang, M.-S., Park, S. J., Rauzier, J., Carralot, J. P., Shrimpton, R., Genovesio, A., Gonzalo-Asensio, J. A., Puzo, G., Martin, C., Brosch, R., Stewart, G. R., Gicquel, B., and Neyrolles, O. (2010) High content phenotypic cell-based visual screen identifies *Mycobacterium tuberculosis* acyltrehalose-containing glycolipids involved in phagosome remodeling. *PLoS Pathog.* *6*, e1001100.
79. Abrahams, K. A. and Besra, G. S. (2016) Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology.* *15*, 1-18.
80. Daffé, M. and Lemassu, A., *Glycobiology of the mycobacterial surface: structures and biological activities of the cell envelope glycoconjugates.*, in *Glycomicrobiology*, R.J. Doyle, Editor. 2000, Kluwer Academic/Plenum. p. 225-273.

81. Guilhot, C., Chaut, C., and Daffé, M., *Biosynthesis and roles of phenolic glycolipids and related molecules in Mycobacterium tuberculosis* in *The Mycobacterial Cell Envelope*. 2008, ASM Press. p. 273-286.
82. McNeil, M., Daffé, M., and Brennan, P. J. (1990) Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J. Biol. Chem.* *265*, 18200-18206.
83. Wheat, W. H., Dhouiba, R., Angala, S. K., Larrouy-Maumus, G., Dobos, K., Nigou, J., Spencer, J. S., and Jackson, M. (2015) The presence of a galactosamine substituent on the arabinogalactan of *Mycobacterium tuberculosis* abrogates full maturation of human peripheral blood monocyte-derived dendritic cells and increases secretion of IL-10. *Tuberculosis.* *95*, 476-489.
84. Angala, S. K., Belardinelli, J. M., Huc-Claustre, E., Wheat, W. H., and Jackson, M. (2014) The cell envelope glycoconjugates of *Mycobacterium tuberculosis*. *Crit. Rev. Biochem. Mol. Biol.* *49*, 361-399.
85. Turnbull, W. B. and Stalford, S. A. (2012) Methylthioxylose - a jewel in the mycobacterial crown? *Org. Biomol. Chem.* *10*, 5698-5706.
86. Choudary, A., Patel, D., Honnen, W., Lai, Z., Prattipati, R. S., Zheng, R. B., Hsueh, Y. C., Gennaro, M. L., Lardizabal, A., Restrepo, B. I., Garcia-Viveros, M., Joe, M., Bai, Y., Shen, K., Sahloul, K., Spencer, J. S., Chatterjee, D., Broger, T., Lowary, T. L., and Pinter, A. (2018) Characterization of the antigenic heterogeneity of lipoarabinomannan, the major surface glycolipid of *Mycobacterium tuberculosis*, and complexity of antibody specificities toward this antigen. *J. Immunol.* *200*, 3053-3066.
87. Trefzer, A., Salas, J. A., and Bechthold, A. (1999) Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat. Prod. Rep.* *16*, 283-299.
88. Reid, A. N. and Szymanski, C. M., *Biosynthesis and assembly of capsular polysaccharides*, in *Microbial Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier. p. 351-374.
89. Koropatkin, N. M. and Holden, H. M. (2005) Structure of CDP-D-glucose 4,6-dehydratase from *Salmonella typhi* complexed with CDP-D-xylose. *Acta. Crystallogr. D. Biol. Crystallogr.* *61*, 365-373.
90. Vogan, E. M., Bellamacina, C., He, X., Liu, H. W., Ringe, D., and Petsko, G. A. (2004) Crystal structure at 1.8 Å resolution of CDP-D-glucose 4,6-dehydratase from *Yersinia pseudotuberculosis*. *Biochemistry* *43*, 3057-3067.
91. Allard, S. T. M., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G., Maskell, D. J., and Naismith, J. H. (2002) Toward a structural understanding of the dehydratase mechanism. *Structure.* *10*, 81-92.
92. Field, R. A. and Naismith, J. H. (2003) Structural and mechanistic basis of bacterial sugar nucleotide-modifying enzymes. *Biochemistry.* *42*, 7637-7647.
93. Salinger, A. J., Thoden, J. B., and Holden, H. M. (2016) Structural and functional investigation of FdhC from *Acinetobacter nosocomialis*: a sugar *N*-acyltransferase belonging to the GNAT superfamily. *Biochemistry.* *55*, 4509-4518.
94. Kubiak, R. L. and Holden, H. M. (2012) Structural studies of AntD: an *N*-acyltransferase involved in the biosynthesis of D-Anthrose. *Biochemistry.* *51*, 867-878.

95. Zhu, J. Y., Yang, Y., Han, H., Betzi, S., Olesen, S. H., Marilio, F., and Schönbrunn, E. (2012) Functional consequence of covalent reaction of phosphoenolpyruvate with UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase (MurA). *J. Biol. Chem.* *287*, 12657-12667.
96. Ovchinnikova, O. G., Rozalski, A., Liu, B., and Knirel, Y. A. (2013) O-antigens of bacteria of the genus *Providencia*: structure, serology, genetics, and biosynthesis. *Biochemistry (Moscow)*. *78*, 798-817.
97. McNally, D. J., Lamoureux, M. P., Karlyshev, A. V., Fiori, L. M., Li, J., Thacker, G., Coleman, R. A., Khieu, N. H., Wren, B. W., Brisson, J. R., Jarrell, H. C., and Szymanski, C. M. (2007) Commonality and biosynthesis of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* *282*, 28566-28576.
98. Guerry, P. and Szymanski, C. M. (2008) *Campylobacter* sugars sticking out. *Trends Microbiol.* *16*, 428-435.
99. Ginsburg, V. (1964) Sugar nucleotides and the synthesis of carbohydrates. *Adv. Enzymol. Relat. Subj. Biochem.* *26*, 35-88.
100. Bernstein, R. L. and Robbins, P. W. (1965) Control aspects of uridine 5'-diphosphate glucose and thymidine 5'-diphosphate glucose synthesis by microbial enzymes. *J. Biol. Chem.* *240*, 391-397.
101. Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W., and Dong, C. (2014) Structural basis for outer membrane lipopolysaccharide insertion. *Nature*. *511*, 52-56.
102. Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* *75*, 39-68.
103. Liston, S. D., Mann, E., and Whitfield, C. (2017) Glycolipid substrates for ABC transporters required for the assembly of bacterial cell-envelope and cell-surface glycoconjugates. *Biochim. Biophys. Acta.* *1862*, 1394-1403.
104. Willis, L. M., Stupak, J., Richards, M. R., Lowary, T. L., Li, J., and Whitfield, C. (2013) Conserved glycolipid termini in capsular polysaccharides synthesized by ATP-binding cassette transporter-dependent pathways in gram-negative pathogens. *Proc. Natl. Acad. Sci. USA.* *110*, 7868-7873.
105. Willis, L. M. and Whitfield, C. (2013) KpsC and KpsS are retaining 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) transferases involved in synthesis of bacterial capsules. *Proc. Natl. Acad. Sci. USA.* *110*, 20753-20758.
106. Willis, L. M. and Whitfield, C. (2013) Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr. Res.* *378*, 35-44.
107. Berg, S., Kaur, D., Jackson, M., and Brennan, P. J. (2007) The glycosyltransferases of *Mycobacterium tuberculosis* - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. *Glycobiology.* *17*, 35-56R.
108. Angala, S. K., McNeil, M. R., Shi, L., Joe, M., Pham, H., Zuberogitia, S., Nigou, J., Boot, C. M., Lowary, T. L., Gilleron, M., and Jackson, M. (2017) Biosynthesis of the methylthioxylose capping motif of lipoarabinomannan in *Mycobacterium tuberculosis*. *ACS Chem. Biol.* *12*, 682-691.
109. Murphy, H. N., Stewart, G. R., Mischenko, V. V., Apt, A. S., Harris, R., McAlister, M. S. B., Driscoll, P. C., Young, D. B., and Robertson, B. D. (2005) The OtsAB pathway is essential for trehalose biosynthesis in *Mycobacterium tuberculosis*. *J. Biol. Chem.* *280*, 14524-14529.

110. Xua, Z., Meshcheryakova, V. A., Poce, G., and Chng, S.-S. (2017) MmpL3 is the flippase for mycolic acids in mycobacteria. *Proc. Natl. Acad. Sci. USA.* *114*, 7993-7998.
111. Malaga, W., Constant, P., Euphrasie, D., Cataldi, A., Daffé, M., Reyrat, J. M., and Guilhot, C. (2008) Deciphering the genetic bases of the structural diversity of phenolic glycolipids in strains of the *Mycobacterium tuberculosis* complex. *J. Biol. Chem.* *283*, 15177-15184.
112. Giraud, M. F. and Naismith, J. H. (2000) The rhamnose pathway. *Curr. Opin. Struct. Biol.* *10*, 687-696.
113. Ballicora, M. A., Iglesias, A. A., and Preiss, J. (2003) ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* *67*, 213-225.
114. Melo, A. and Glaser, L. (1965) The nucleotide specificity and feedback control of thymidine diphosphate D-glucose pyrophosphorylase. *J. Biol. Chem.* *240*, 398-405.
115. Munch-Petersen, A., Kalckar, H. M., Cutolo, E., and Smith, E. E. B. (1953) Uridyl transferases and the formation of uridine triphosphate; enzymic production of uridine triphosphate: uridine diphosphoglucose pyrophosphorolysis. *Nature.* *172*, 1036-1037.
116. Rubenstein, P. A. and Strominger, J. L. (1974) Enzymatic synthesis of cytidine diphosphate 3,6-dideoxyhexoses: IX. purification and properties of the cytidine diphosphate-D-glucose pyrophosphorylase from *Pasteurella pseudotuberculosis*, type V. *J. Biol. Chem.* *249*, 3789-3796.
117. Kornfeld, R. H. and Ginsburg, V. (1966) Control of synthesis of guanosine 5'-diphosphate D-mannose and guanosine 5'-diphosphate L-fucose in bacteria. *Biochim. Biophys. Acta.* *117*, 79-87.
118. Barton, W. A., Lesniak, J., Biggins, J. B., Jeffrey, P. D., Jiang, J., Rajashankar, K. R., Thorson, J. S., and Nikolov, D. B. (2001) Structure, mechanism and engineering of a nucleotidyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* *8*, 545-551.
119. Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.* *19*, 6652-6663.
120. Sivaraman, J., Sauvé, V., Matte, A., and Cygler, M. (2002) Crystal structure of *Escherichia coli* glucose-1-phosphate thymidyltransferase (RffH) complexed with dTTP and Mg²⁺. *J. Biol. Chem.* *277*, 44214-44219.
121. Thoden, J. B. and Holden, H. M. (2007) Active site geometry of glucose-1-phosphate uridyltransferase. *Protein Sci.* *16*, 1379-1388.
122. Koropatkin, N. M., Cleland, W. W., and Holden, H. M. (2005) Kinetic and structural analysis of α -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *280*, 10774-10780.
123. Pelissier, M. C., Lesley, S. A., Kuhn, P., and Bourne, Y. (2010) Structural insights into the catalytic mechanism of bacterial guanosine-diphospho-D-mannose pyrophosphorylase and its regulation by divalent ions. *J. Biol. Chem.* *285*, 27468-27476.
124. Comino, N., Cifuentes, J. O., Marina, A., Orrantia, A., Eguskiza, A., and Guerin, M. E. (2017) Mechanistic insights into the allosteric regulation of bacterial ADP-glucose pyrophosphorylases. *J. Biol. Chem.* *292*, 6255-6268.

125. Kostrewa, D., D'Arcy, A., Takacs, B., and Kamber, M. (2001) Crystal structures of *Streptococcus pneumoniae* N-acetylglucosamine-1-phosphate uridylyltransferase, GlmU, in apo form at 2.33 Å resolution and in complex with UDP-N-acetylglucosamine and Mg²⁺ at 1.96 Å resolution. *J. Mol. Biol.* *305*, 279-289.
126. Wu, B., Zhang, Y., Zheng, R., Guo, C., and Wang, P. G. (2002) Bifunctional phosphomannose isomerase/GDP-D-mannose pyrophosphorylase is the point of control for GDP-D-mannose biosynthesis in *Helicobacter pylori*. *FEBS Lett.* *519*, 87-92.
127. Koropatkin, N. M. and Holden, H. M. (2004) Molecular structure of α-D-glucose-1-phosphate cytidylyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *279*, 44023-44029.
128. Jin, X., Ballicora, M. A., Preiss, J., and Geiger, J. H. (2005) Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO J.* *24*, 694-704.
129. Alphey, M. S., Pirrie, L., Torrie, L. S., Boulkeroua, W. A., Gardiner, M., Sarkar, A., Maringer, M., Oehlmann, W., Brenk, R., Scherman, M. S., McNeil, M., Rejzek, M., Field, R. A., Singh, M., Gray, D., Westwood, N. J., and Naismith, J. H. (2013) Allosteric competitive inhibitors of glucose-1-phosphate thymidylyltransferase (RmlA) from *Pseudomonas aeruginosa*. *ACS Chem. Biol.* *8*, 387-396.
130. Zuccotti, S., Zanardi, D., Rosano, C., Sturla, L., Tonetti, M., and Bolognesi, M. (2001) Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidylyltransferase. *J. Mol. Biol.* *313*, 831-843.
131. Paule, M. R. and Preiss, J. (1971) Biosynthesis of bacterial glycogen X: the kinetic mechanism of adenosine diphosphoglucose pyrophosphorylase from *Rhodospirillum rubrum*. *J. Biol. Chem.* *246*, 4602-4609.
132. Hossain, S. A., Tanizawa, K., Kazuta, Y., and Fukui, T. (1994) Overproduction and characterization of recombinant UDP-glucose pyrophosphorylase from *Escherichia coli* K-12. *J. Biochem.* *115*, 965-972.
133. Olsen, L. R. and Roderick, S. L. (2001) Structure of the *Escherichia coli* GlmU pyrophosphorylase and acetyltransferase active sites. *Biochemistry.* *40*, 1913-1921.
134. Qu, H., Xin, Y., Dong, X., and Ma, Y. (2007) An *rmlA* gene encoding D-glucose-1-phosphate thymidylyltransferase is essential for mycobacterial growth. *FEMS Microbiol. Lett.* *275*, 237-243.
135. DeLano, W. L., The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. 2002.
136. Holden, H. M., Thoden, J. B., and Gilbert, M. (2016) Enzymes required for the biosynthesis of N-formylated sugars. *Curr. Opin. Struct. Biol.* *41*, 1-9.
137. Dunsirn, M. M., Thoden, J. B., and Holden, H. M. (2017) Biochemical investigation of Rv3404c from *Mycobacterium tuberculosis*. *Biochemistry.* *56*, 3818-3825.
138. Ma, Y., Stern, R. J., Scherman, M. S., Vissa, V. D., Yan, W., Jones, V. C., Zhang, F., Franzblau, S. G., Lewis, W. H., and McNeil, M. R. (2001) Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob. Agents. Chemother.* *45*, 1407-1416.
139. Toney, M. D. (2014) Aspartate aminotransferase: an old dog teaches new tricks. *Arch. Biochem. Biophys.* *544*, 119-127.

140. Noland, B. W., Newman, J. M., Hendle, J., Badger, J., Christopher, J. A., Tresser, J., Buchanan, M. D., Wright, T. A., Rutter, M. E., Sanderson, W. E., Müller-Dieckmann, H. J., Gajiwala, K. S., and Buchanan, S. G. (2002) Structural studies of *Salmonella typhimurium* ArnB (PmrH) aminotransferase: a 4-amino-4-deoxy-L-arabinose lipopolysaccharide-modifying enzyme. *Structure*. *10*, 1569-1580.
141. Schoenhofen, I. C., Lunin, V. V., Julien, J. P., Li, Y., Ajamian, E., Matte, A., Cygler, M., Brisson, J. R., Aubry, A., Logan, S. M., Bhatia, S., Wakarchuk, W. W., and Young, N. M. (2006) Structural and functional characterization of PseC, an aminotransferase involved in the biosynthesis of pseudaminic acid, an essential flagellar modification in *Helicobacter pylori*. *J. Biol. Chem.* *281*, 8907-8916.
142. Burgie, E. S. and Holden, H. M. (2007) Molecular architecture of DesI: a key enzyme in the biosynthesis of desosamine. *Biochemistry*. *46*, 8999-9006.
143. Dow, G. T., Gilbert, M., Thoden, J. B., and Holden, H. M. (2017) Structural investigation on WlaRG from *Campylobacter jejuni*: a sugar aminotransferase. *Protein Sci.* *26*, 586-599.
144. Wang, F., Singh, S., Zu, W., Helmich, K. E., Miller, M. D., Cao, H., Bingman, C. A., Thorson, J. S., and Phillips, G. N. (2015) Structural basis for the stereochemical control of amine installation in nucleotide sugar aminotransferases. *ACS Chem. Biol.* *10*, 2048-2056.
145. Thoden, J. B., Schäffer, C., Messner, P., and Holden, H. M. (2009) Structural analysis of QdtB, an aminotransferase required for the biosynthesis of dTDP-3-acetamido-3,6-dideoxy- α -D-glucose. *Biochemistry*. *48*, 1553-1561.
146. Cook, P. D., Carney, A. E., and Holden, H. M. (2008) Accomodation of GDP-linked sugars in the active site of GDP-perosamine synthase. *Biochemistry*. *47*, 10685-10693.
147. Riegert, A. S., Young, N. M., Watson, D. C., Thoden, J. B., and Holden, H. M. (2015) Structure of the external aldimine form of PglE, an aminotransferase required for *N,N'*-diacetylbacillosamine biosynthesis. *Protein Sci.* *24*, 1609-1616.
148. Tomshich, S. V., Gorshkova, R. P., and Ovodov, Y. S. (1987) Structural studies on lipopolysaccharide from *Y. enterocolitica* serovar O:8. *Khimia Prirodnikh Soedinenii.*, 657-664.
149. Holden, H. M., Rayment, I., and Thoden, J. B. (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J. Biol. Chem.* *278*, 43885-43888.
150. Graninger, M., Nidetzky, B., Heinrichs, D. E., Whitfield, C., and Messner, P. (1999) Characterization of dTDP-4-dehydrorhamnose 3,5-epimerase and dTDP-4-dehydrorhamnose reductase, required for dTDP-L-rhamnose biosynthesis in *Salmonella enterica* serovar typhimurium LT2. *J. Biol. Chem.* *274*, 25069-25077.
151. Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kalberg, Y., Persson, B., and Jörnvall, H. (2003) Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem. Biol. Interact.* *143-144*, 247-253.
152. Giraud, M. F., Leonard, G. A., Field, R. A., Berlind, C., and Naismith, J. H. (2000) RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat. Struct. Biol.* *7*, 398-402.
153. Dong, C., Major, L. L., Srikannathan, V., Errey, J. C., Giraud, M. F., Lam, J. S., Graninger, M., Messner, P., McNeil, M. R., Field, R. A., Whitfield, C., and Naismith, J. H. (2007) RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J. Mol. Biol.* *365*, 146-159.

154. Liu, Y., Thoden, J. B., Kim, J., Berger, E., Gulick, A. M., Ruzicka, F. J., Holden, H. M., and Frey, P. A. (1997) Mechanistic roles of tyrosine 149 and serine 124 in UDP-galactose 4-epimerase from *Escherichia coli*. *Biochemistry*. *36*, 10675-10684.
155. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose-4-epimerase from *Escherichia coli*: implications for the catalytic mechanism. *Biochemistry*. *35*, 5137-5144.
156. Nelsestuen, G. L. and Kirkwood, S. (1971) The mechanism of action of the enzyme uridine diphosphoglucose 4-epimerase: proof of an oxidation-reduction mechanism with direct transfer of hydrogen between substrate and the B-position of the enzyme-bound pyridine nucleotide. *J. Biol. Chem.* *246*, 7533-7543.
157. Lau, S. T. B. and Tanner, M. E. (2008) Mechanism and active site residues of GDP-fucose synthase. *J. Am. Chem. Soc.* *130*, 17593-17602.
158. Major, L. L., Wolucka, B. A., and Naismith, J. H. (2005) Structure and function of GDP-mannose-3',5'-epimerase: an enzyme which performs three chemical reactions at the same active site. *J. Am. Chem. Soc.* *127*, 18309-18320.
159. Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., and Goodenough, P. W. (2001) Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* *26*, 740-746.
160. Thorson, J. S., Lo, S. F., Ploux, O., He, Z., and Liu, H. W. (1994) Studies of the biosynthesis of 3,6-dideoxyhexoses: molecular cloning and characterization of the *asc* (ascarylose) region from *Yersinia pseudotuberculosis* serogroup VA. *J. Bacter.* *176*, 5483-5493.
161. Merkel, A. B., Major, L. L., Errey, J. C., Burkart, M. D., Field, R. A., Walsh, C. T., and Naismith, J. H. (2004) The position of a key tyrosine in dTDP-4-keto-6-deoxy-D-glucose-5-epimerase (EvaD) alters the substrate profile for this RmlC-like enzyme. *J. Biol. Chem.* *279*, 32684-32691.
162. Jakimowicz, P., Tello, M., Freel Meyers, C. L., Walsh, C. T., Buttner, M. J., Field, R. A., and Lawson, D. M. (2006) The 1.6-Å resolution crystal structure of NovW: A 4-keto-6-deoxy sugar epimerase from the novobiocin biosynthetic gene cluster of *Streptomyces spheroides*. *Proteins*. *63*, 261-265.
163. Tello, M., Jakimowicz, P., Errey, J. C., Freel Meyers, C. L., Walsh, C. T., Buttner, M. J., Lawson, D. M., and Field, R. A. (2006) Characterisation of *Streptomyces spheroides* NovW and revision of its functional assignment to a dTDP-6-deoxy-D-xylo-4-hexulose 3-epimerase. *Chem. Commun. (Camb)*. *14*, 1079-1081.
164. Kubiak, R. L., Phillips, R. K., Zmudka, M. W., Ahn, M. R., Maka, E. M., Pyeat, G. L., Roggensack, S. J., and Holden, H. M. (2012) Structural and functional studies on a 3'-epimerase involved in the biosynthesis of dTDP-6-deoxy-D-allose. *Biochemistry*. *51*, 9375-9383.
165. Dong, C., Major, L. L., Allen, A., Blankenfeldt, W., Maskell, D., and Naismith, J. H. (2003) High-resolution structures of RmlC from *Streptococcus suis* in complex with substrate analogs locate the active site of this class of enzyme. *Structure*. *11*, 715-723.
166. Li, W., Xin, Y., McNeil, M. R., and Ma, Y. (2006) *rmlB* and *rmlC* genes are essential for growth in mycobacteria. *Biochem. Biophys. Res. Commun.* *342*, 170-178.
167. Rahim, R., Burrows, L. L., Monteiro, M. A., Perry, M. B., and Lam, J. S. (2000) Involvement of the *rml* locus in core oligosaccharide and O polysaccharide assembly in *Pseudomonas aeruginosa*. *Microbiology*. *146*, 2804-2814.

168. Yamashita, Y., Tomihisa, K., Nakano, Y., Shimazaki, Y., Oho, T., and Koga, T. (1999) Recombination between *gtfB* and *gtfC* is required for survival of a dTDP-rhamnose synthesis-deficient mutant of *Streptococcus mutans* in the presence of sucrose. *Infect. Immun.* *67*, 3693-3697.
169. Szymanski, C. M., St. Michael, F., Jarrell, H. C., Li, J., Gilbert, M., Larocque, S., Vinogradov, E., and Brisson, J. R. (2003) Detection of conserved *N*-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. *J. Biol. Chem.* *278*, 24509-24520.
170. van Alphen, L. B., Wenzel, C. Q., Richards, M. R., Fodor, C., Ashmus, R. A., Stahl, M., Karlyshev, A. V., Wren, B. W., Stintzi, A., Miller, W. G., Lowary, T. L., and Szymanski, C. M. (2014) Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One.* *9*, e87051.
171. Young, K. T., Davis, L. M., and DiRita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* *5*, 665-679.
172. Ang, C. W., Laman, J. D., Willison, H. J., Wagner, E. R., Endtz, H. P., de Klerk, M. A., Tio Gillen, A. P., van den Braak, N., Jacobs, B. C., and van Doorn, P. A. (2002) Structure of *Campylobacter jejuni* lipopolysaccharides determines antiganglioside specificity and clinical features of Guillain-Barré and Miller Fisher patients. *Infect. Immun.* *70*, 1202-1208.
173. Jacobs, B. C., Rothbarth, P. H., van der Meché, F. G. A., Herbrink, P., Schmitz, P. I. M., de Klerk, M. A., and van Doorn, P. A. (1998) The spectrum of antecedent infections in Guillain-Barré syndrome: a case-control study. *Neurology.* *51*, 1110-1115.
174. Karlyshev, A. V., Ketley, J. M., and Wren, B. W. (2005) The *Campylobacter jejuni* glycome. *FEMS Microbiol. Rev.* *29*, 377-390.
175. Logan, S. M., Trust, T. J., and Guerry, P. (1989) Evidence for posttranslational modification and gene duplication of *Campylobacter* flagellin. *J. Bacter.* *171*, 3031-3038.
176. Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* *32*, 1022-1030.
177. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., and Monteiro, M. A. (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines *Front. Cell Infect. Microbiol.* *2*, doi: 10.3389.
178. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J., and Wren, B. W. (2000) Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol. Microbiol.* *35*, 529-541.
179. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holfoyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M. J., Penn, C. W., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H. M., Whitehead, S., and Barrell, B. G. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature.* *403*, 665-668.
180. Bacon, D. J., Szymanski, C. M., Burr, D. H., Silver, R. P., Alm, R. A., and Guerry, P. (2001) A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol. Microbiol.* *40*, 769-777.

181. Bachtiar, B. M., Coloe, P. J., and Fry, B. N. (2007) Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 8116 and its involvement in bacterium-host interactions. *FEMS Immunol. Med. Microbiol.* *49*, 149-154.
182. Wong, A., Lange, D., Houle, S., Arbatsky, N. P., Valvano, M. A., Knirel, Y. A., Dozois, C. M., and Creuzenet, C. (2015) Role of capsular modified heptose in the virulence of *Campylobacter jejuni*. *Mol. Microbiol.* *96*, 1136-1158.
183. Holst Sørensen, M. C., van Alphen, L. B., Fodor, C., Crowley, S. M., Christensen, B. B., Szymanski, C. M., and Brønsted, L. (2012) Phase variable expression of capsular polysaccharide modifications allows *Campylobacter jejuni* to avoid bacteriophage infection in chickens. *Front. Cell Infect. Microbiol.* *2*, Article 11.
184. Pequegnat, B., Laird, R. M., Ewing, C. P., Hill, C. L., Omari, E., Poly, F., Monteiro, M. A., and Guerry, P. (2017) Phase-variable changes in the position of the *O*-methyl phosphoramidate modifications on the polysaccharide capsule of *Campylobacter jejuni* modulate serum resistance. *J. Bacteriol.* *199*, e00027-17.
185. Karlyshev, A. V., Champion, O. L., Churcher, C., Brisson, J. R., Jarrell, H. C., Gilbert, M., Brochu, D., St. Michael, F., Li, J., Wakarchuk, W. W., Goodhead, I., Sanders, M., Stevens, K., White, B., Parkhill, J., Wren, B. W., and Szymanski, C. M. (2005) Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* *55*, 90-103.
186. Holst Sørensen, M. C., van Alphen, L. B., Harboe, A., Li, J., Christensen, B. B., Szymanski, C. M., and Brønsted, L. (2011) Bacteriophage F336 recognizes the capsular phosphoramidate modification of *Campylobacter jejuni* NCTC11168. *J. Bacteriol.* *193*, 6742-6749.
187. Kim, S., Vela, A., Clohisey, S. M., Athanasiadou, S., Kaiser, P., Stevens, M. P., and Vervelde, L. (2018) Host-specific differences in the response of cultured macrophages to *Campylobacter jejuni* capsule and *O*-methyl phosphoramidate mutants. *Vet. Res.* *49*, s13567-017-0501-y.
188. Maue, A. C., Mohawk, K. L., Giles, D. K., Poly, F., Ewing, C. P., Jiao, Y., Lee, G., Ma, Z., Monteiro, M. A., Hill, C. L., Ferdberber, J. S., Porter, C. K., Trent, M. S., and Guerry, P. (2013) The polysaccharide capsule of *Campylobacter jejuni* modulates the host immune response. *Infect. Immun.* *81*, 665-672.
189. Ellington, W. R. (2001) Evolution and physiological roles of phosphagen systems. *Annu. Rev. Physiol.* *63*, 289-325.
190. Attwood, P. V., Piggott, M. J., Zu, X. L., and Besant, P. G. (2007) Focus on phosphohistidine. *Amino Acids.* *32*, 145-156.
191. Lehman, I. R. (1974) DNA ligase: structure, mechanism, and function. *Science.* *186*, 790-797.
192. Mitchell, R. E. (1976) Isolation and structure of a chlorosis-inducing toxin of *Pseudomonas phaseolicola*. *Phytochemistry.* *15*, 1941-1947.
193. Guijarro, J. I., González-Pastor, J. E., Baleux, F., Millán, J. L. S., Castilla, M. A., Rico, M., Moreno, F., and Delepierre, M. (1995) Chemical structure and translation inhibition studies of the antibiotic microcin C7. *J. Biol. Chem.* *270*, 23520-23532.
194. Pietrowska-Borek, M., Nuc, K., and Guranowski, A. (2015) Exogenous adenosine 5'-phosphoramidate behaves as a signal molecule in plants; it augments metabolism of phenylpropanoids and salicylic acid in *Arabidopsis thaliana* seedlings. *Plant Physiol. Biochem.* *94*, 144-152.

195. Umezawa, S., Tatsuta, K., Izawa, O., Tsuchiya, T., and Umezawa, H. (1972) A new microbial metabolite phosphoramidon (isolation and structure). *Tetrahedron Lett.* *13*, 97-100.
196. Fukuhara, K., Murao, S., Nozawa, T., and Hatano, M. (1982) Structural elucidation of talopeptin (MK-1), a novel metallo proteinase inhibitor produced by *Streptomyces mozunensis* MK-23. *Tetrahedron Lett.* *23*, 2319-2322.
197. Ogita, T., Gunji, S., Fukazawa, Y., Terahara, A., Kinoshita, T., Nagaki, H., and Beppu, T. (1983) The structures of fosfazinomycins A and B. *Tetrahedron Lett.* *24*, 2283-2286.
198. Cáceres, T., Megharaj, M., Venkateswarlu, K., Sethunathan, N., and Naidu, R., *Fenamiphos and related organophosphorus pesticides: environmental fate and toxicology*, in *Reviews of Environmental Contamination and Toxicology*, D. Whitacre, Editor. 2010, Springer: New York, NY.
199. Pai, E. F., Kregel, U., Petsko, G. A., Goody, R. S., Kabsch, W., and Wittinghofer, A. (1990) Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* *9*, 2351-2359.
200. Anthony, K., Feng, L., Arechavala-Gomez, V., Guglieri, M., Straub, V., Bushby, K., Cirak, S., Morgan, J., and Muntoni, F. (2012) Exon skipping quantification by quantitative reverse-transcription polymerase chain reaction in Duchenne Muscular Dystrophy patients treated with the antisense oligomer Eteplirsen. *Hum. Gene. Ther. Methods.* *23*, 336-345.
201. Cahn, P., Rolon, M. J., Gun, A. M., Ferrari, I., Dibirdik, I., Qazi, S., D'Cruz, O., Sahin, K., and Uckun, F. (2012) Preclinical and first-in-human phase I clinical evaluation of Stampidine, a potent anti-HIV pharmaceutical drug candidate. *J. AIDS Clin. Res.* *3*, doi: 10.4172.
202. McNally, D. J., Lamoureaux, M., Li, J., Kelly, J., Brisson, J. R., Szymanski, C. M., and Jarrell, H. C. (2006) HR-MAS NMR Studies of ¹⁵N-labeled cells confirm the structure of the *O*-methyl phosphoramidate CPS modification in *Campylobacter jejuni* and provide insight into its biosynthesis. *Can. J. Chem.* *84*, 676-684.
203. Herzberg, O., Chen, C. C. H., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano, D. (1996) Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci. USA.* *93*, 2652-2657.
204. Berman, K. M. and Cohn, M. (1970) Phosphoenolpyruvate synthetase of *Escherichia coli*. *J. Biol. Chem.* *245*, 5309-5318.
205. Massière, F. and Badet-Denisot, M. A. (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol. Life Sci.* *54*, 205-222.
206. Kwak, B. Y., Zhang, Y. M., Yun, M., Heath, R. J., Rock, C. O., Jackowski, S., and Park, H. W. (2002) Structure and mechanism of CTP:phosphocholine cytidylyltransferase (LicC) from *Streptococcus pneumoniae*. *J. Biol. Chem.* *277*, 4343-4350.
207. Leyh, T. S., Taylor, J. C., and Markham, G. D. (1988) The sulfate activation locus of *Escherichia coli* K12: cloning, genetic, and enzymatic characterization. *J. Biol. Chem.* *263*, 2409-2416.
208. Poulin, M. B., Nothaft, H., Hug, I., Feldman, M. F., Szymanski, C. M., and Lowary, T. L. (2010) Characterization of a bifunctional pyranose-furanose mutase from *Campylobacter jejuni* 11168. *J. Biol. Chem.* *285*, 493-501.
209. Feinstein, N. W., Allen, S., and Jenkins, E. (2013) Outside the pipeline: reimagining science education for nonscientists. *Science.* *340*, 314-317.

Chapter 2:

The Structure of Glucose-1-Phosphate Thymidyltransferase from *Mycobacterium tuberculosis* Reveals the Location of an Essential Magnesium Ion in the RmlA-type Enzymes

This work was previously published as: Brown, H. A., Thoden, J. B., Tipton, P. A., and Holden, H. M. (2018) The Structure of Glucose-1-Phosphate Thymidyltransferase from *Mycobacterium tuberculosis* Reveals the Location of an Essential Magnesium Ion in the RmlA-type Enzymes. *Protein Sci.* 27, 441-450.

2.1 Abstract

Tuberculosis, caused by the bacterium *Mycobacterium tuberculosis*, continues to be a major threat to populations worldwide. Whereas the disease is treatable, the drug regimen is arduous at best with the use of four antimicrobials over a six-month period. There is clearly a pressing need for the development of new therapeutics. One potential target for structure-based drug design is the enzyme RmlA, a glucose-1-phosphate thymidyltransferase. This enzyme catalyzes the first step in the biosynthesis of L-rhamnose, which is a deoxysugar critical for the integrity of the bacterium's cell wall. Here we report the X-ray structures of *M. tuberculosis* RmlA in complex with either dTTP or dTDP-glucose to 1.6 Å and 1.85 Å resolution, respectively. In the RmlA/dTTP complex, two magnesium ions were observed binding to the nucleotide, both ligated in octahedral coordination spheres. In the RmlA/dTDP-glucose complex, only a single magnesium ion was observed. Importantly, for RmlA-type enzymes with known three-dimensional structures, not one model shows the position of the magnesium ion bound to the nucleotide-linked sugar. As such, this investigation represents the first direct observation of the manner in which a magnesium ion is coordinated to the RmlA product and thus has important ramifications for structure-based drug design. In the past, molecular modeling procedures have been employed to derive a three-dimensional model of the *M. tuberculosis* RmlA for drug design. The X-ray structures presented herein provide a superior molecular scaffold for such endeavors in the treatment of one of the world's deadliest diseases.

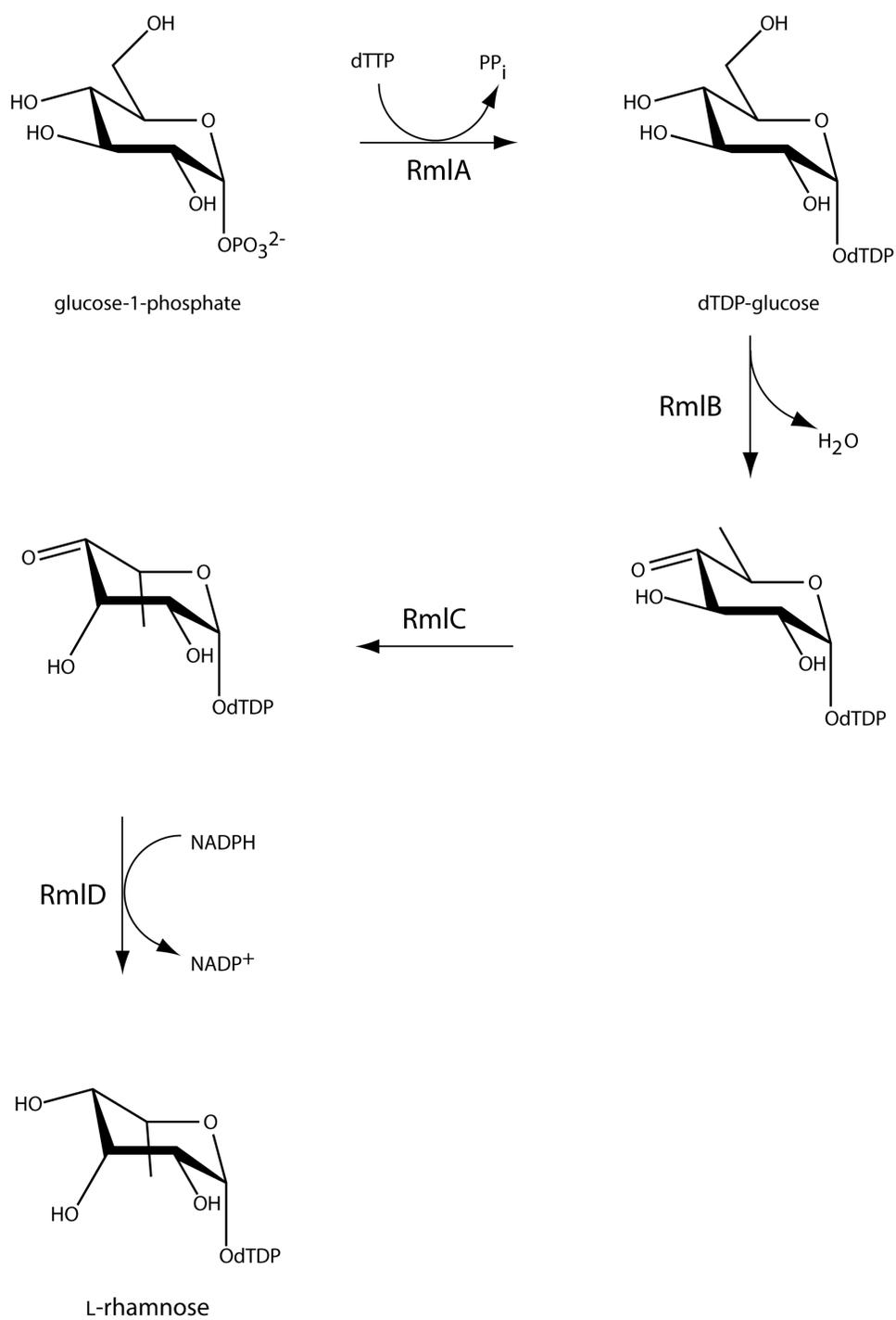
2.2 Introduction

Mycobacterium tuberculosis, a facultative intracellular pathogen, is the causative agent of tuberculosis. According to the World Health Organization, *M. tuberculosis* resulted in the deaths of 1.8 million people worldwide in 2015 alone (1). Although the number of tuberculosis deaths fell by 22% between 2000 and 2015, there is still cause for concern given the appearance of multidrug-resistant and extensively drug-resistant bacterial strains (2).

The cell wall of *M. tuberculosis*, known to be essential for viability, consists of three components: the peptidoglycan, the arabinogalactan, and the outermost layer composed of mycolic acids (3). The arabinogalactan is connected to the peptidoglycan by a disaccharide composed of α -L-rhamnosyl-(1 \rightarrow 3)- α -D-N-acetylglucosaminosyl-1-phosphate (4). The precursor to the rhamnosyl moiety is dTDP-L-rhamnose, which is synthesized as outlined in Scheme 2.1 (5). L-rhamnose is found not only in *M. tuberculosis*, but also in a variety of pathogenic bacteria including but not limited to *Salmonella enterica*, *Shigella flexneri*, and various *Escherichia coli* strains (6). As can be seen in Scheme 2.1, four enzymes, RmlA, RmlB, RmlC, and RmlD, are required for the production of dTDP-L-rhamnose. Given that L-rhamnose is not found in humans, these enzymes have served as potential targets for drug design (6, 7).

The focus of this investigation is on RmlA, which catalyzes the first step in the pathway, namely the conversion of glucose-1-phosphate and dTTP to dTDP-glucose and pyrophosphate. Thus far the enzymes from *Pseudomonas aeruginosa*, *S. enterica*, and *E. coli* have been biochemically and structurally characterized (8-10). These enzymes are known to be allosterically regulated by dTDP-rhamnose (11), and the allosteric binding pockets have been well characterized (8-10).

Given that the *rmlA* gene has been shown to be essential for mycobacterial growth, it is somewhat surprising that the crystal structure of the enzyme from *M. tuberculosis* has never been described (12). Indeed, two recent reports focusing on possible inhibitors of the enzyme from *M. tuberculosis* relied solely upon molecular modeling procedures (13, 14). Here we report a functional and structural investigation of *M. tuberculosis* RmlA. Importantly, for the first time we demonstrate the manner in which enzymes in this superfamily coordinate a magnesium ion through their nucleotide-linked sugar products.



Scheme 2.1 Pathway for the biosynthesis of dTDP-rhamnose.

2.3 Materials and Methods

2.3.1 Cloning, expression, and purification

The gene encoding RmlA (*rv0334*) from *M. tuberculosis* H37Rv was synthesized by Integrated DNA technologies using codons optimized for protein expression in *E. coli*. The gene was cloned into a pET31b vector (Novagen) using NdeI and XhoI restriction sites to yield a construct with a C-terminal His₆ tag. The pET31b-*rmlA* plasmid was utilized to transform Rosetta2(DE3) *E. coli* cells (Novagen). The cultures were grown in lysogeny broth supplemented with ampicillin and chloramphenicol (100 mg/L and 25 mg/L concentrations, respectively) at 37 °C with shaking until an optical density of 0.9 was reached at 600 nm. The flasks were cooled in cold water, and the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and allowed to express protein at 21 °C for 20 h.

The cells were harvested by centrifugation and frozen as pellets in liquid nitrogen. The pellets were subsequently disrupted by sonication on ice in a lysis buffer composed of 50 mM sodium phosphate, 20 mM imidazole, 10% glycerol, and 300 mM sodium chloride, pH 8.0. The lysate was cleared by centrifugation, and RmlA with a C-terminal His-tag was purified at 4 °C utilizing nickel nitrilotriacetic acid resin (Qiagen) according to the manufacturer's instructions. All buffers were adjusted to pH 8.0 and contained 50 mM sodium phosphate, 300 mM sodium chloride, and imidazole concentrations of 25 mM for the wash buffer and 250 mM for the elution buffer. Following purification, the protein was dialyzed against 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl at 4 °C, and concentrated to 9 mg/mL based on the calculated extinction coefficient of 0.71 (mg/mL)⁻¹ cm⁻¹.

2.3.2 Protein crystallization and X-ray structural analyses

Crystallization conditions were surveyed by the hanging drop method of vapor diffusion using a sparse matrix screen developed in the Holden laboratory. Initial experiments were conducted with the enzyme in complex with either 10 mM MgCl₂ and 5 mM dTTP (Sigma-Aldrich) or 10 mM MgCl₂ and 10 mM dTDP-glucose at room temperature. The dTDP-glucose ligand was prepared as previously described (15). X-ray diffraction quality crystals grown in the presence of dTTP reached a maximal size of approximately 0.6 mm x 0.2 mm x 0.2 mm after one week. Specifically, they were grown by mixing 1:1 the protein sample at 9 mg/mL containing 10 mM MgCl₂ and 5 mM dTTP with a precipitant solution composed of 18 - 22% poly(ethylene glycol) 8000 and 100 mM MOPS (pH 7.0). The crystals belonged to the space group *I4* with unit cell dimensions of $a = b = 96.4 \text{ \AA}$, and $c = 151.9 \text{ \AA}$. The asymmetric unit contained two subunits. Prior to X-ray data collection, the crystals were transferred step-wise into a solution composed of 25% poly(ethylene glycol) 8000, 250 mM NaCl, 100 mM MOPS (pH 7.0), 10 mM MgCl₂, 5 mM dTTP, and 15% ethylene glycol.

X-ray diffraction quality crystals of the RmlA/dTDP-glucose complex grew to a maximal size of approximately 0.8 mm x 0.2 mm x 0.4 mm after one to two weeks. They were grown by mixing 1:1 the protein sample at 9 mg/mL containing 10 mM MgCl₂ and 10 mM dTDP-glucose with a precipitant solution composed of 20% poly(ethylene glycol) 3350, 200 mM KCl, and 100 mM CHES (pH 9.0). These crystals belonged to the space group *P2₁2₁2₁* with unit cell dimensions of $a = 72.3 \text{ \AA}$, $b = 111.3 \text{ \AA}$, and $c = 290.4 \text{ \AA}$. The asymmetric unit contained eight subunits. Prior to X-ray data collection, the crystals were transferred step-wise into a solution composed of 25% poly(ethylene glycol) 3350, 250 mM KCl, 250 mM NaCl, 100 mM CHES (pH 9.0), 10 mM MgCl₂, 10 mM dTDP-glucose, and 15% ethylene glycol.

X-ray data sets were collected at the Structural Biology Center beamline 19-BM at a wavelength of 0.9794 Å (Advanced Photon Source). The X-ray data were processed and scaled with HKL3000 (16). Relevant X-ray data collection statistics are provided in Table 2.1. The structure of the RmlA/dTTP complex was solved via molecular replacement with the software package PHASER (17) using PDB entry code 1H5S (*E. coli* RmlA) (10) as a search probe. Iterative cycles of model building with COOT (18) and refinement with REFMAC (19) reduced the R_{work} and R_{free} to 19.5% and 22.7%, respectively, from 50 to 1.60 Å resolution. The structure of RmlA in the presence of dTDP-glucose was solved via molecular replacement using the RmlA/dTTP complex coordinates as a search probe. Iterative cycles of model building with COOT and refinement with REFMAC reduced the R_{work} and R_{free} to 17.1% and 21.8%, respectively, from 50 to 1.85 Å resolution. Relevant refinement statistics are listed in Table 2.2.

Table 2.1 X-ray Data Collection Statistics.

	RmlA/dTTP Complex	RmlA/dTDP-glucose Complex
resolution limits	50 – 1.60 (1.66 – 1.60) ^b	50.0 – 1.85 (1.92 – 1.85)
Space Group	<i>I4</i>	<i>P2₁2₁2₁</i>
Unit Cell		
<i>a</i> (Å)	96.4	72.3
<i>b</i> (Å)	96.4	111.3
<i>c</i> (Å)	151.9	290.4
number of independent reflections	89386 (8778)	195396 (18536)
completeness (%)	98.5 (97.0)	97.7 (93.8)
redundancy	5.9 (3.7)	6.5 (4.5)
avg <i>I</i> /avg σ (<i>I</i>)	55.7 (5.9)	45.8 (14.9)
R_{sym} (%) ^a	4.9 (19.8)	10.3 (17.9)
Wilson B factor (Å ²)	20.0	19.9

$${}^a R_{\text{sym}} = (\sum |I - \bar{I}| / \sum I) \times 100.$$

^bStatistics for the highest resolution bin.

Table 2.2 Refinement Statistics.

	RmlA/dTTP Complex	RmlA/dTDP-glucose Complex
resolution limits (Å)	50 – 1.60	50 – 1.85
^a R-factor (overall)%/no. reflections	19.6/89386	17.4/195366
R-factor (working)%/no. reflections	19.5/84972	17.1/185587
R-factor (free)%/no. reflections	22.7/4414	21.8/9779
number of protein atoms	4217	17583
number of heteroatoms	466	2003
average B values		
protein atoms (Å ²)	32.4	26.8
ligands (Å ²)	28.9	23.6
solvent (Å ²)	33.3	29.4
weighted RMS deviations from ideality		
bond lengths (Å)	0.013	0.013
bond angles (°)	1.7	1.7
general planes (°)	0.007	0.007
Ramachandran regions (%)^b		
most favored	98.3	98.8
additionally allowed	1.5	0.8
generously allowed	0.2	0.4

^aR-factor = $(\sum|F_o - F_c| / \sum|F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

^bDistribution of the Ramachandran angles according to PROCHECK (20).

2.3.3 Kinetic analyses

Inosine, microbial xanthine oxidase, and human purine nucleoside phosphorylase were purchased from Sigma Aldrich.

RmlA activity was monitored spectrophotometrically by following an increase in absorbance at 290 nm concomitant with xanthine formation in a modified coupled assay first reported by Forget, S.M. *et al.*, 2015 (21). Reactions were monitored with a Beckman DU 640B spectrophotometer. Briefly, pyrophosphate released by RmlA is hydrolyzed to free phosphate by inorganic pyrophosphatase. Purine nucleoside phosphorylase then reacts with inosine and phosphate to yield hypoxanthine. Finally, hypoxanthine is oxidized to xanthine via xanthine oxidase, and its formation is monitored at 290 nm with an extinction coefficient of $12200 \text{ M}^{-1} \text{ cm}^{-1}$.

100 μL reactions were set up at room temperature with 50 mM Tris-HCl (pH 7.5), 1 mM inosine, 10 U/mL of *E. coli* inorganic pyrophosphatase, 1 μM purine nucleoside phosphorylase, and 1.5 U/mL of xanthine oxidase. The required inorganic pyrophosphatase with a C-terminal tag was purified in the laboratory according to standard procedures. The K_M of the enzyme for dTTP was determined by holding the glucose-1-phosphate concentration constant at 400 μM and varying the dTTP concentrations from 1 to 200 μM . The K_M of the enzyme for glucose-1-phosphate was determined by holding the dTTP concentration constant at 100 μM and varying the glucose-1-phosphate concentrations from 1 to 250 μM .

As a control, all reactions were initially set up without RmlA to monitor the background rate due to phosphate contamination in the starting materials. Reactions were initiated with the addition of RmlA to a final concentration of 0.5 $\mu\text{g/mL}$. All data were fitted by initial velocity Michaelis – Menten kinetics to the following equation: $v = (V_{\text{max}}[S])/(K_M + [S])$. The k_{cat} values were calculated according to the equation $k_{\text{cat}} = V_{\text{max}}/[E]_t$. Relevant kinetic parameters are listed in Table 2.3.

Table 2.3 Kinetic Parameters.

Substrate	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($\text{M}^{-1}\text{s}^{-1}$)
dTTP	10.6 ± 0.7	7.9 ± 0.2	$(7.5 \pm 0.5) \times 10^5$
glucose-1-phosphate	32.7 ± 2.9	7.3 ± 0.2	$(2.2 \pm 0.2) \times 10^5$

2.4 Results and Discussion

2.4.1 Overall architecture of *M. tuberculosis* RmlA

The first structure determined in this investigation was that of RmlA in complex with dTTP. The crystals utilized for the X-ray analysis belonged to the space group *I4* with two subunits in the asymmetric unit. RmlA functions as homotetramer with 222-symmetry. Thus, one of the twofold axes of the tetramer was coincident to a crystallographic dyad. The structure was determined to 1.6 Å resolution and refined to an overall *R*-factor of 19.6%. Tyr 29, in each subunit, adopts ϕ , ψ angles of $\sim 68^\circ$ and -81° , respectively, which are outside of the allowed regions in the Ramachandran plot. The electron densities for Tyr 29 in each subunit were unambiguous, however. It is located in a tight turn connecting β -strands two and three of the subunit.

Shown in Figure 2.1 (a) is a ribbon representation of the tetramer with the crystallographic axis indicated by the red line. The tetramer has overall dimensions of ~ 77 Å x 78 Å x 84 Å. Subunit 1 in the asymmetric unit extends from Met 1 to Glu 286 with breaks in the polypeptide chain backbone between Phe 147 to Leu 154 and Ala 159 to Pro 163. Subunit 2 extends from Met 1 to Glu 286 with breaks in the polypeptide chain backbone between Glu 146 to Ser 155 and Glu 157 to Ser 165. The α -carbons for the two subunits in the asymmetric unit superimpose with a root-mean-square deviation of 0.2 Å.

There was clear electron density in each subunit for a dTTP molecule in the active site and a dTDP molecule in the allosteric binding pocket. The active site and the allosteric binding pocket, within a subunit, are separated by ~ 20 Å. Representative electron density for the dTTP ligand and its associated magnesium ions, as observed in subunit 1, is presented in Figure 2.1 (b), whereas that for the dTDP molecule is displayed in Figure 2.1 (c).

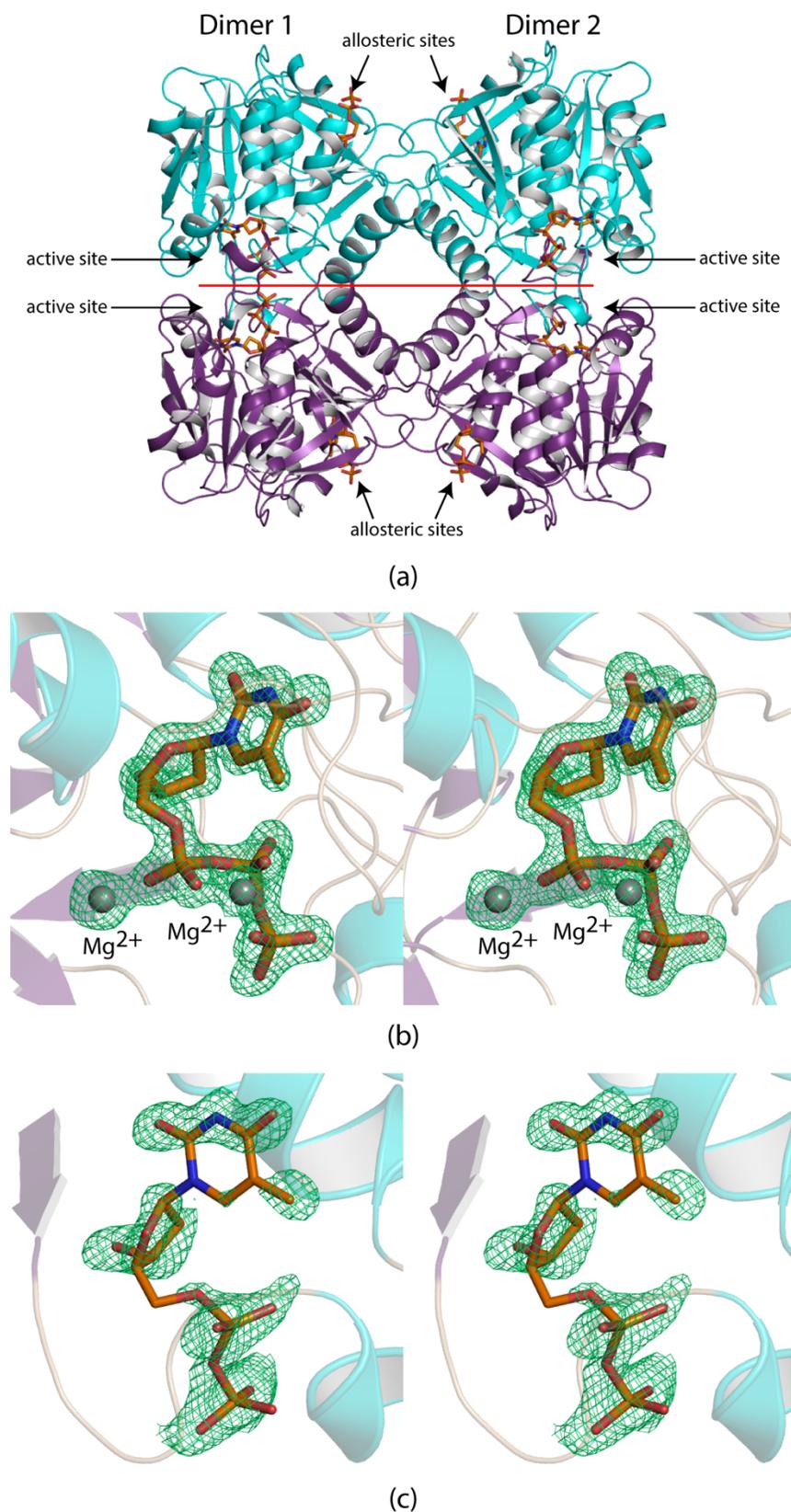


Figure 2.1 Overall structure of *M. tuberculosis* RmlA. Shown in (a) is a ribbon representation of the RmlA tetramer, which displays 222 -symmetry. One of the twofold axes of the tetramer, indicated by the red line, packed along a crystallographic dyad thereby reducing the contents of the asymmetric unit to half a tetramer. Electron density corresponding to the bound dTTP in the active site of subunit 1 is shown in stereo in (b). The electron density map was calculated with $(F_o - F_c)$ coefficients and contoured at 3σ . The ligands were not included in the X-ray coordinate file used to calculate the omit map, and thus there is no model bias. Shown in (c) is the electron density corresponding to the bound dTDP ligand in the allosteric binding pocket (subunit 1). The electron density map was calculated as described in (b) and contoured at 3σ . As can be seen, the density is weaker most likely due to the lower occupancy of the ligand. This figure and Figures 2.2, 2.3, 2.4, and 2.6 were prepared with PyMOL (22).

A ribbon representation of subunit 1 is depicted in Figure 2.2 (a). The overall molecular architecture of the subunit is built around 14 β -strands and 11 α -helices. Pro 16, which lies between the first β -strand and the first α -helix adopts the *cis* conformation. It is not close to the active site but rather is situated at a subunit:subunit interface. The predominant feature of the subunit is a seven-stranded mixed β -sheet with the active site positioned at the end of the β -sheet. There are three additional anti-parallel β -sheets (one containing three β -strands and the other two composed of two β -strands). The allosteric binding pocket is surrounded primarily by α -helices [Fig. 2.2 (a)].

A close-up view of the active site is displayed in Figure 2.2 (b). The side chain of Gln 80 plays a key role in anchoring the thymine ring into the active site cleft. Arg 13 and Lys 23 interact through their side chains with the γ - and α -phosphoryl groups of the dTTP ligand, respectively. One of the magnesium ions, referred to as MgA, is octahedrally coordinated by an α -phosphoryl oxygen, the side chains of Asp 108 and Asp 222 and three water molecules. The average ligand:metal bond length is 2.1 Å. The second magnesium ion, MgB, is octahedrally coordinated by an α -, a β -, and a γ -phosphoryl oxygen and three waters. Again, the average ligand:metal bond length is 2.1 Å. Interestingly, in the structure of the *E. coli* RffH, only one magnesium ion was observed binding in the enzyme/dTTP complex (corresponding to MgA in the *M. tuberculosis* RmlA) (23). The *E. coli* RffH is also a glucose-1-phosphate thymidyltransferase and demonstrates 59% and 74% amino acid sequence identity and similarity, respectively, to the *M. tuberculosis* RmlA.

Shown in Figure 2.2 (c) is a close-up view of the allosteric binding pocket. Most of the ligand:protein hydrogen bonding interactions are provided by backbone amide groups (Gly 113, Gly 115, Gly 117, Thr 118, Ser 119, and Ser 248). There are two side chains, Thr 118 and Ser 119, that participate in hydrogen bonding interactions with the pyrophosphoryl group.

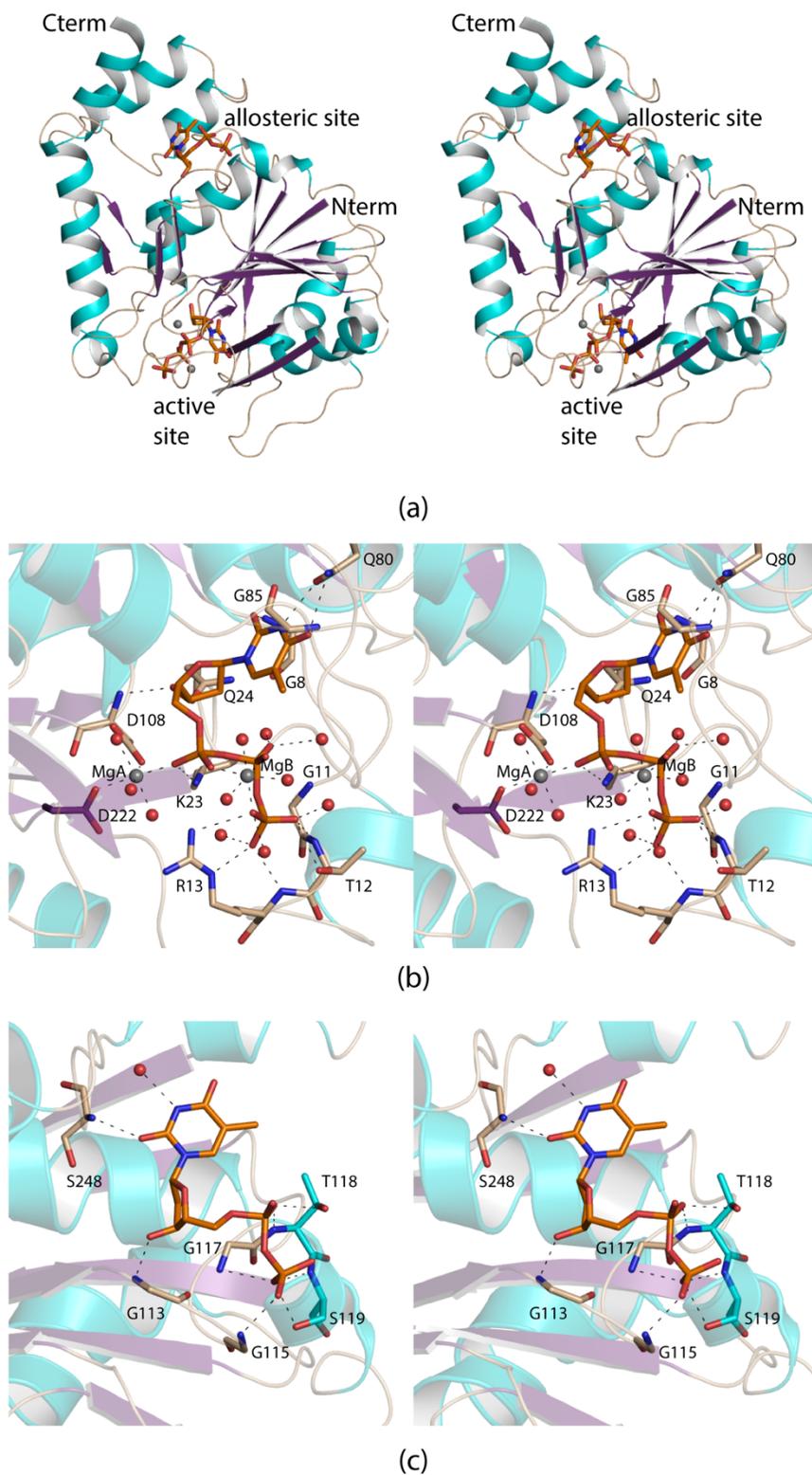


Figure 2.2 Structure of the *M. tuberculosis* RmlA subunit. A ribbon representation of the RmlA subunit 1 is presented in (a). The bound ligands are drawn in stick representations, and the magnesium ions are depicted as black spheres. Shown in (b) is the region of the active site surrounding the dTTP/Mg²⁺ complex. Water molecules are depicted as red spheres. Potential hydrogen bonding interactions, within 3.2 Å, are indicated by the dashed lines. Presented in (c) are the observed interactions between the dTDP ligand and the protein in the allosteric binding pocket.

2.4.2 Structure of the *M. tuberculosis* RmlA/dTDP-glucose complex

The second structure determined in this investigation was that of the enzyme in complex with its product, dTDP-glucose. The crystals used in this analysis belonged to the space group $P2_12_12_1$ with two tetramers in the asymmetric unit. The α -carbons for the individual subunits superimpose with root-mean-square deviations of between 0.2 Å (subunits 3 and 4) to 0.4 Å (subunits 5 and 8). There are little structural perturbations upon binding the dTDP-glucose product. Indeed, the α -carbons for the RmlA/dTTP and the RmlA/dTDP-glucose subunits superimpose with a root-mean-square deviation of 0.3 Å.

The observed electron density corresponding to the bound ligand and a magnesium ion in subunit 1 is shown in Figure 2.3 (a). A close-up view of the active site is presented in Figure 2.3 (b). In this complex only one magnesium ion is observed binding, namely that coordinated by Asp 108 and Asp 222. In addition to these amino acid residues, the other ligands to the metal include an α - and a β -phosphoryl oxygen and two water molecules. This type of metal binding is similar to that previously observed for the *Salmonella typhi* glucose-1-phosphate cytidylyltransferase with bound CDP-glucose (24). The pyranosyl moiety of the dTDP-glucose is anchored into the active site by the side chain of Glu 158, the backbone amide group of Gly 143, and the carbonyl oxygen of Val 169. As in the first RmlA/dTTP structure, dTDP is observed in the allosteric binding pocket.

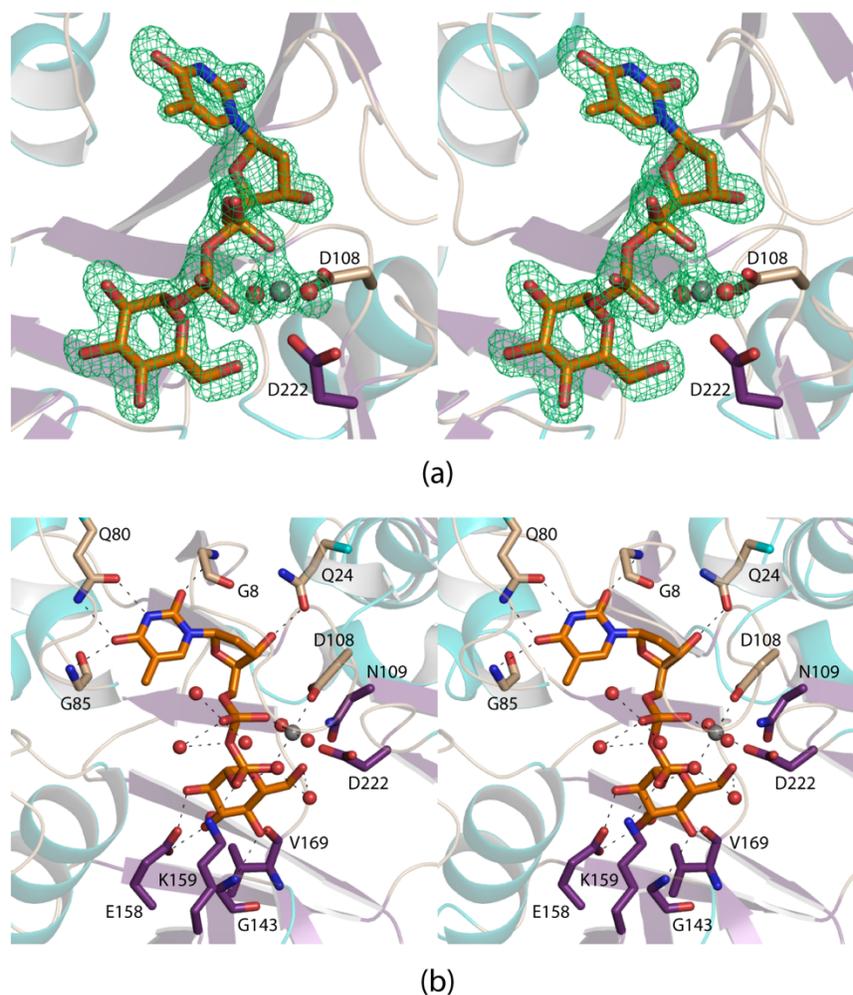


Figure 2.3 Close-up stereo views of the active site with bound dTDP-glucose. The observed electron density in subunit 1 is presented in (a). The map was calculated and contoured as described in Figure legend 2.1. Potential hydrogen bonding interactions between the protein and the dTDP-glucose ligand are depicted as dashed lines in (b). The bound magnesium ion is represented by the gray sphere whereas ordered water molecules are depicted as red spheres.

RmlA-type enzymes are known to be magnesium-dependent (25). Surprisingly, of all the entries for RmlA-type enzymes in the Protein Data Bank, not one has a bound nucleotide-linked sugar with its associated magnesium ion. As such, this investigation represents the first direct observation of the manner in which a magnesium ion is coordinated to the RmlA product. In one molecular dynamics study of the *M. tuberculosis* RmlA, a model was produced in order to aid in the development of novel inhibitors (14). Coordinates for the *M. tuberculosis* RmlA model were deposited with the Protein Model Database (PM0076036). The α -carbons for the “RmlA model” and the complex determined in this investigation superimpose with a root-mean-square deviation of 0.9 Å. Shown in Figure 2.4 (a) is a superposition of the ribbon drawings for the two models. Overall the molecular architectures are similar. The issue, however, arises in the active sites. The modeling never took into account the position of the magnesium ion, and specifically that Asp 222 serves as a ligand to the metal. As can be seen in Figure 2.4 (b), the net result is that in the theoretical model, Asp 222 is splayed away from the active site. This type of detail has profound consequences for structure-based drug design.

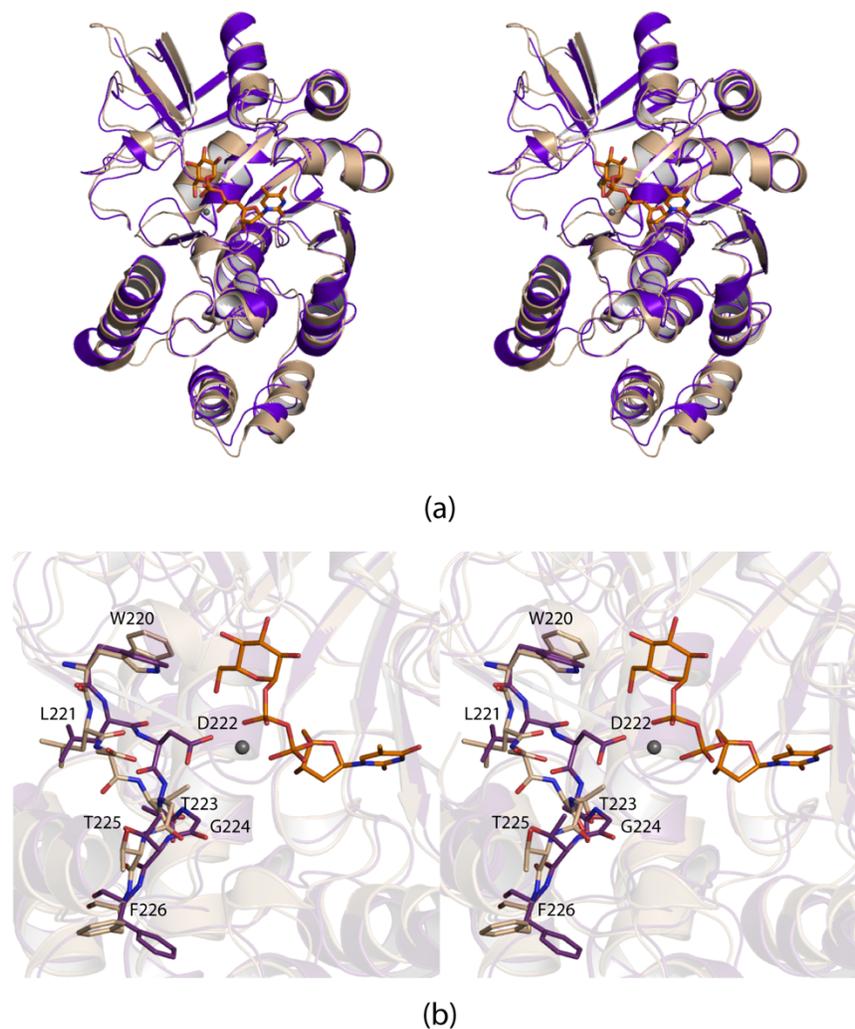
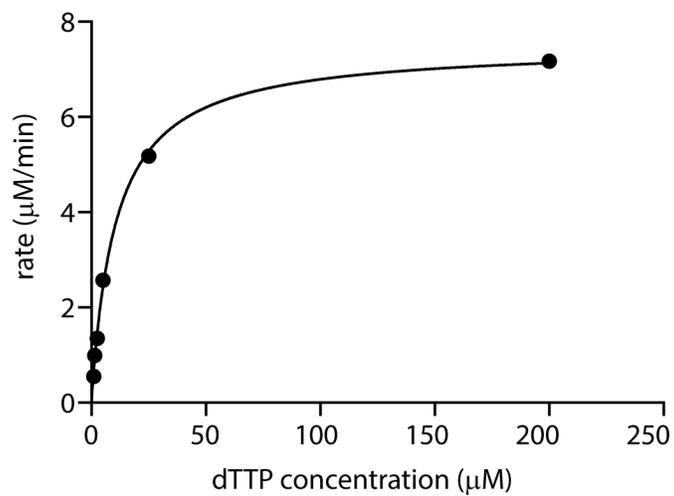


Figure 2.4 Comparison of the experimentally determined RmlA model presented here with that based on the homology modeling previously described. A superposition of the ribbon representations for the two models in stereo is presented in (a) with the experimentally determined model highlighted in violet. Overall the folds are similar as would be expected. One of the major differences, however, is the positioning of the loop delineated by Trp 220 to Phe 226. As can be seen in (b), in the theoretical model the Asp 222 side-chain is positioned away from the active site.

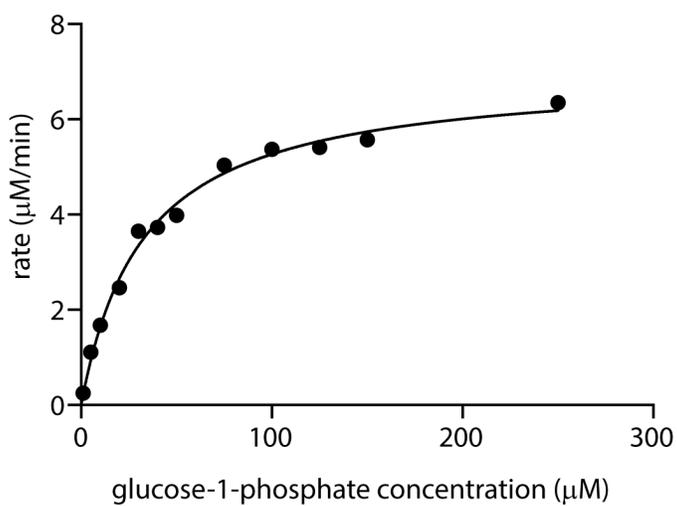
2.4.3 Kinetic analysis of *M. tuberculosis* RmlA

The kinetic parameters for *M. tuberculosis* RmlA were determined by a coupled spectrophotometric assay as described in the Materials and Methods. Shown in Figures 2.5 (a) and 2.5 (b) are plots of initial velocities versus either dTTP or glucose-1-phosphate concentrations, respectively. As can be seen, RmlA demonstrates classical Michaelis-Menten kinetics. The observed K_M values of $10.6 \pm 0.7 \mu\text{M}$ and $32.7 \pm 2.9 \mu\text{M}$ for dTTP and glucose-1-phosphate, respectively, are similar to those values reported for the enzyme from *E. coli* (10). *M. tuberculosis* RmlA displays catalytic efficiencies of $7.5 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $2.2 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for dTTP and glucose-1-phosphate, respectively.

The kinetic mechanisms for the *E. coli* thymidyltransferase and the *S. typhi* cytidyltransferase have been previously reported to be sequential-ordered Bi Bi (10, 26). A superposition of the coordinates for the *M. tuberculosis* RmlA/dTTP complex onto those of the *M. tuberculosis* RmlA/dTDP-glucose is presented in Figure 2.6. Most likely, the binding of the glucose moiety in dTDP-glucose mimics what would occur for the substrate, glucose-1-phosphate. If that is, indeed, the case, then the phosphoryl group of glucose-1-phosphate would be $\sim 3.1 \text{ \AA}$ from the α -phosphorus of dTTP. It can thus be envisioned how the substrate, glucose-1-phosphate, is in line for a direct attack at the α -phosphorus of the dTTP molecule. The pyrophosphate group, with its associated magnesium ion, would function as an ideal leaving group.



(a)



(b)

Figure 2.5 Plots of initial velocities versus substrate concentrations. Shown in (a) is a graph of the initial rate versus dTTP concentration. The initial rate versus glucose-1-phosphate concentration is shown as a graph in (b).

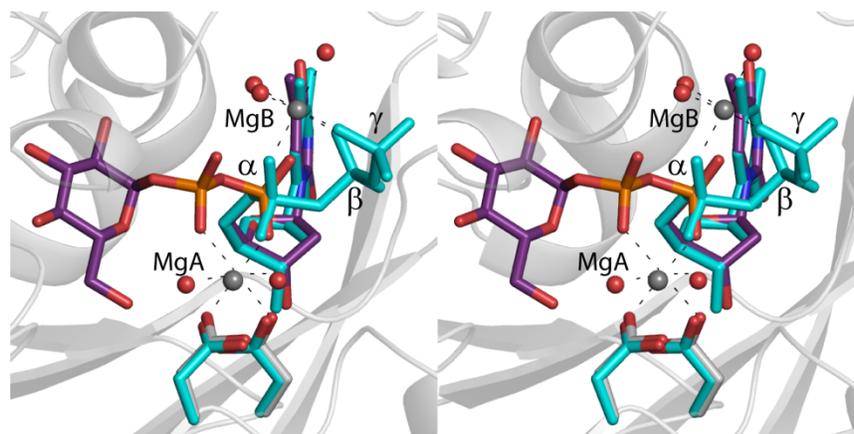


Figure 2.6 Superposition of the ligands in the RmlA/dTTP and the RmlA/dTDP-glucose complexes. The dTDP-glucose is highlighted in violet bonds whereas the dTTP is colored in cyan.

In summary, the structure of the *M. tuberculosis* thymidyltransferase has now been determined to high resolution and the coordination geometry around the catalytically required magnesium ions defined. The high-resolution models presented herein thus provide new and important data for the subsequent design of novel therapeutic inhibitors. Treatment for tuberculosis typically requires extensive use of isoniazid, rifampicin, ethambutol, and pyrazinamide for two months followed by an additional four months of isoniazid and rifampicin (27). Without the development of shorter and more effective treatments, tuberculosis will remain a worldwide disease responsible for both significant illness and economic devastation.

2.5 References

1. 2015; Available from: <http://www.who.int/mediacentre/factsheets/fs104/en/>.
2. Brouqui, P., Quenard, F., and Drancourt, M. (2017) Old antibiotics for emerging multidrug-resistant/extensively drug-resistant tuberculosis (MDR/XDR-TB). *Int. J. Antimicrob. Agents.* *49*, 554-557.
3. Abrahams, K. A. and Besra, G. S. (2016) Mycobacterial cell wall biosynthesis: a multifaceted antibiotic target. *Parasitology.* *15*, 1-18.
4. McNeil, M., Daffé, M., and Brennan, P. J. (1990) Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. *J. Biol. Chem.* *265*, 18200-18206.
5. Shibaev, V. N. (1986) Biosynthesis of bacterial polysaccharide chains composed of repeating units. *Adv. Carbohydr. Chem. Biochem.* *44*, 277-339.
6. Giraud, M. F. and Naismith, J. H. (2000) The rhamnose pathway. *Curr. Opin. Struct. Biol.* *10*, 687-696.
7. Ma, Y., Stern, R. J., Scherman, M. S., Vissa, V. D., Yan, W., Jones, V. C., Zhang, F., Franzblau, S. G., Lewis, W. H., and McNeil, M. R. (2001) Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob. Agents. Chemother.* *45*, 1407-1416.
8. Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.* *19*, 6652-6663.
9. Barton, W. A., Lesniak, J., Biggins, J. B., Jeffrey, P. D., Jiang, J., Rajashankar, K. R., Thorson, J. S., and Nikolov, D. B. (2001) Structure, mechanism and engineering of a nucleotidyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* *8*, 545-551.
10. Zuccotti, S., Zanardi, D., Rosano, C., Sturla, L., Tonetti, M., and Bolognesi, M. (2001) Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidyltransferase. *J. Mol. Biol.* *313*, 831-843.
11. Melo, A. and Glaser, L. (1965) The nucleotide specificity and feedback control of thymidine diphosphate D-glucose pyrophosphorylase. *J. Biol. Chem.* *240*, 398-405.
12. Qu, H., Xin, Y., Dong, X., and Ma, Y. (2007) An *rmlA* gene encoding D-glucose-1-phosphate thymidyltransferase is essential for mycobacterial growth. *FEMS Microbiol. Lett.* *275*, 237-243.

13. Mansuri, R., Ansari, M. Y., Singh, J., Rana, S., Sinha, S., Sahoo, G. C., Dikhit, M. R., and Das, P. (2016) Computational elucidation of structural basis for ligand binding with *Mycobacterium tuberculosis* glucose-1-phosphate thymidyltransferase (RmlA). *Curr. Pharm. Biotechnol.* *17*, 1089-1099.
14. Harathi, N., Pulaganti, M., Anuradha, C. M., and Chitta, S. K. (2016) Inhibition of *Mycobacterium*-RmlA by molecular modeling, dynamics simulation, and docking. *Adv. Bioinformatics.* *2016*, 9841250.
15. Takahashi, H., Liu, Y. N., and Liu, H. W. (2006) A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J. Am. Chem. Soc.* *128*, 1432-1433.
16. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* *276*, 307-326.
17. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658-674.
18. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* *60*, 2126-2132.
19. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D. Biol. Crystallogr.* *53*, 240-255.
20. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* *231*, 1049-1067.
21. Forget, S. M., Jee, A., Smithen, D. A., Jagdhane, R., Anjum, S., Beaton, S. A., Palmer, D. R. J., Syvitski, R. T., and Jakeman, D. L. (2015) Kinetic evaluation of glucose-1-phosphate analogues with a thymidyltransferase using a continuous coupled enzyme assay. *Org. Biomol. Chem.* *13*, 866-875.
22. DeLano, W. L., The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. 2002.
23. Sivaraman, J., Sauv e, V., Matte, A., and Cygler, M. (2002) Crystal structure of *Escherichia coli* glucose-1-phosphate thymidyltransferase (RffH) complexed with dTTP and Mg²⁺. *J. Biol. Chem.* *277*, 44214-44219.
24. Koropatkin, N. M. and Holden, H. M. (2004) Molecular structure of α -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *279*, 44023-44029.
25. Bernstein, R. L. and Robbins, P. W. (1965) Control aspects of uridine 5'-diphosphate glucose and thymidine 5'-diphosphate glucose synthesis by microbial enzymes. *J. Biol. Chem.* *240*, 391-397.

26. Koropatkin, N. M., Cleland, W. W., and Holden, H. M. (2005) Kinetic and structural analysis of α -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *280*, 10774-10780.
27. Cole, S. T. (2016) Inhibiting *Mycobacterium tuberculosis* within and without. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *371*, pii: 20150506.

2.6 Acknowledgements

A portion of the research described here was performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source (U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357). We gratefully acknowledge Drs. Randy Alkire and Krzysztof Lazarski for assistance during X-ray data collection.

Chapter 3:

The *Mycobacterium tuberculosis* Complex has a Pathway for the Biosynthesis of 4-Formamido-4,6-Dideoxy-D-Glucose

This work was previously published as: Brown, H. A., Vinogradov, E., Gilbert, M., and Holden, H. M. (2018) The *Mycobacterium tuberculosis* Complex has a Pathway for the Biosynthesis of 4-Formamido-4,6-Dideoxy-D-Glucose. *Protein Sci.* [In press].

3.1 Abstract

Recent studies have demonstrated that the O-antigens of some pathogenic bacteria such as *Brucella abortus*, *Francisella tularensis*, and *Campylobacter jejuni* contain quite unusual *N*-formylated sugars (3-formamido-3,6-dideoxy-D-glucose or 4-formamido-4,6-dideoxy-D-glucose). Typically, four enzymes are required for the formation of such sugars: a thymidyltransferase, a 4,6-dehydratase, a pyridoxal 5'-phosphate or PLP-dependent aminotransferase, and an *N*-formyltransferase. To date, there have been no published reports of *N*-formylated sugars associated with *Mycobacterium tuberculosis*. A recent investigation from our laboratories, however, has demonstrated that one gene product from *M. tuberculosis*, Rv3404c, functions as a sugar *N*-formyltransferase. Given that *M. tuberculosis* produces L-rhamnose, both a thymidyltransferase (Rv0334) and a 4,6-dehydratase (Rv3464) required for its formation have been identified. Thus, there is one remaining enzyme needed for the production of an *N*-formylated sugar in *M. tuberculosis*, namely a PLP-dependent aminotransferase. Here we demonstrate that the *M. tuberculosis* *rv3402c* gene encodes such an enzyme. Our data prove that *M. tuberculosis* contains all of the enzymatic activities required for the formation of dTDP-4-formamido-4,6-dideoxy-D-glucose. Indeed, the *rv3402c* gene product likely contributes to virulence or persistence during infection, though its temporal expression and location remain to be determined.

3.2 Introduction

The causative agent of tuberculosis, *Mycobacterium tuberculosis*, is a facultative intracellular pathogen. It exhibits a complex life cycle with its human host that is still not fully understood. Indeed, the human immune response is often only partially effective, leading to bacteria that exist in nonproliferative persistent states (1). These bacteria have apparently evolved a number of strategies to survive within human alveolar macrophages (2). Strikingly, only 5 to 15% of those individuals with “latent” infections will develop active infections in the course of their lifetimes, and as a consequence they function as “repositories” for the bacterium (3). This is especially concerning given that tuberculosis remains one of the top 10 causes of death worldwide according to the World Health Organization.

Within recent years, a large number of *M. tuberculosis* virulence genes have been identified, in part through the use of transposon mutant libraries and *in vivo* screening techniques (4). Some of these genes encode putative enzymes involved in lipid biosynthetic pathways whereas others likely encode cell surface proteins or regulators of signal transduction systems (4). The focus of this investigation is on the open reading frame *rv3402c*. This particular gene has been shown to be upregulated in *M. tuberculosis* cultures deficient in iron (5). In addition, a recent report has demonstrated that the Rv3402c protein, when expressed in *Mycobacterium smegmatis*, enhances the organism’s intracellular persistence within macrophages (6).

We recently demonstrated that another protein, Rv3404c from *M. tuberculosis* H37Rv, functions as a sugar *N*-formyltransferase by transferring a formyl group from *N*¹⁰-formyltetrahydrofolate to dTDP-4-amino-4,6-dideoxy-D-glucose (7). Our results were especially intriguing given that there have been no reports in the literature regarding the existence of *N*-formylated sugars in *M. tuberculosis*. Typically, these unusual sugars are found on the

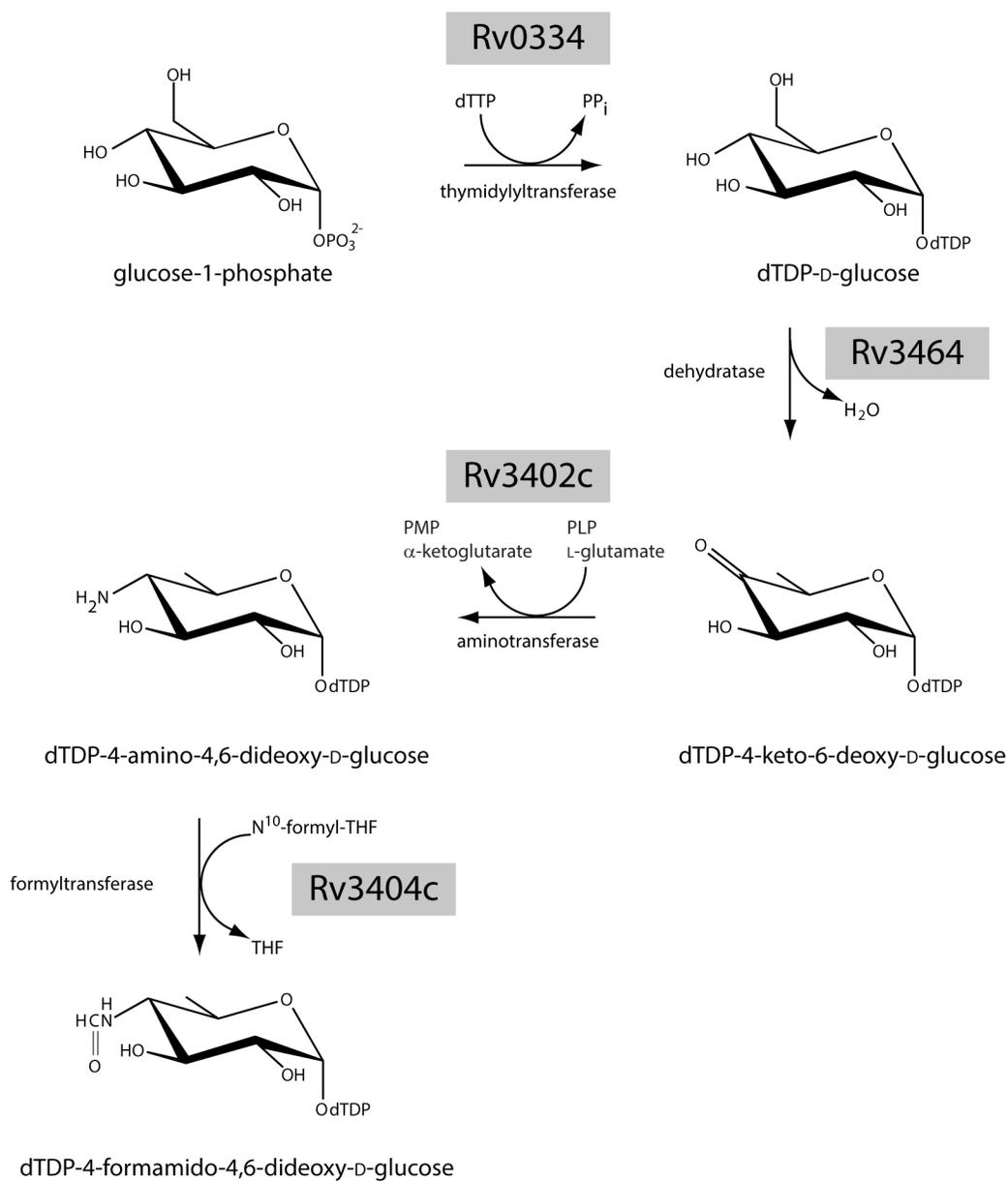
lipopolysaccharides of pathogenic Gram-negative bacteria such as *Campylobacter jejuni*, *Francisella tularensis*, *Providencia alcalifaciens*, *Salmonella enterica*, and *Brucella melitensis* (8-15).

Shown in Scheme 3.1 is the accepted biosynthetic pathway by which *N*-formylated sugars are produced in bacteria. Rv3404c catalyzes the last step in the pathway. The question then arises as to whether *M. tuberculosis* contains the additional enzymatic machinery required for the production of *N*-formylated sugars. As can be seen in Scheme 3.1, the first step involves the attachment of glucose-1-phosphate to a dTMP moiety via the action of a glucose-1-phosphate thymidyltransferase. There is an enzyme in *M. tuberculosis*, encoded by the *rv0334* gene, that catalyzes such a reaction. It is specifically involved in the biosynthesis of L-rhamnose (16), and its three-dimensional structure has now been determined to high resolution (17). In the second step of the pathway, dTDP-D-glucose is dehydrated by the action of a 4,6-dehydratase to yield dTDP-4-keto-6-deoxy-D-glucose. Again, this reaction is also required for the production of L-rhamnose in *M. tuberculosis*, and not surprisingly the gene encoding for an enzyme with such activity has been identified, namely *rv3464* (16).

Thus, the only enzyme missing from the pathway outlined in Scheme 3.1 is a pyridoxal 5'-phosphate or PLP-dependent aminotransferase. Our initial analysis of the *M. tuberculosis* H37Rv genome identified the Rv3402c protein as a potential candidate for such activity. A BLAST search of the Rv3402c amino acid sequence against the Protein Data Bank indicates that it belongs to the Type I aminotransferase superfamily (18). The closest structural relative of Rv3402c that has been biochemically and structurally characterized is DesI from *Streptomyces venezuelae* (19). Rv3402c and DesI demonstrate amino acid sequence identities and similarities of 29% and 39%, respectively (Fig. 3.1). Importantly, DesI catalyzes the formation of dTDP-4-amino-4,6-dideoxy-

D-glucose from dTDP-4-keto-6-deoxy-D-glucose. It can thus be postulated that Rv3402c catalyzes the third reaction in Scheme 3.1. In keeping with this hypothesis, Rv3402c has been previously predicted to be involved in the biosynthesis of lipopolysaccharide-like molecules (5).

Here we demonstrate that Rv3402c is, indeed, a PLP-dependent aminotransferase that functions on dTDP-4-keto-6-deoxy-D-glucose to produce dTDP-4-amino-4,6-dideoxy-D-glucose. Our results provide strong evidence that at some point in the life cycle of *M. tuberculosis*, *N*-formylated sugars are produced. Whether these sugars are important for virulence remains an open question. However, the loss of activity of a sugar *N*-formyltransferase in *Brucella abortus* results in a bacterial strain with attenuated pathogenicity (13).



Scheme 3.1 Possible pathway for the *in vivo* production of dTDP-4-formamido-4,6-dideoxy-D-glucose.

Rv3402c	MKIRTLSGSVLEPPSAVRATPGTSMKLEPGGSTIPK--IPFIRPSFPGPAELAEDFVQI	58
DesI	MK-----SALSDLAFF-GGPAAFDQPLLVGRPNRIDRARLYERLDRA	41
Rv3402c	AQANWYTNFGPNERRFARALRDYLGPHLHVATLANGTLALLAALHVSFGAGTRDRYLLMP	118
DesI	LDSQWLSNNGGPLVREFEERVAGLAGVRHAVATCNA-T----AGLQLLAHAAGLTGEVIMP	96
Rv3402c	SFTFVGVQAALWTGYRPWFIDIDANTWQPCVHSARAVIERFRDRIAGILLANVFGVGNP	178
DesI	SMTFAATPHALRWIGLTPVFADIDPDTGNLDPDQVAAAVTPRTSAVVG---HLWGRPCA	153
Rv3402c	QISVWEELAAEWELPIVLD SAAGFGSTYADGERLGGRGACEIFS FHAT K PFVAVGEGGALV	238
DesI	-ADQLRKVADEHGLRLYFDAAHALGC-AVDGRPAGSLGDAEVFSFHATKAVNAFEGGAVV	211
Rv3402c	SRDPRLVEHAYKFQNFGLVQTRESIQLGMN NGKLSEISAAIGLRQLVGLDRRLASRRK VLE	298
DesI	TDDADLAARIRALHNF GF DLPGGSPAGGTNAKMSEAAAAMGLTSLDAFPEVIDRNRNHA	271
Rv3402c	CYRTGMADAGVRFQDNANVA-----SLCFASACCTSADHKA AVL GSLRRHAIEARD Y YN	352
DesI	AYREHLADLPGV L VADHDRHGLNNHQYVIVEIDEATTGIHRDLVMEVLKAEGVHTRAYFS	331
Rv3402c	PPQHRHPYFVTNAELVESTDLAVTADICSRIVSLPVHDHMAPDDVARVVA AV QEA EV RGE	412
DesI	PGCHELEPYRGQP----HAPLPHTERLAARVLSLPTGTAIGDDDIRRVADLLRLCATRGR	387
Rv3402c	-----	412
DesI	ELTARHRDTAPAPLAAPQTSTPTIGRSRGLE	418

Figure 3.1 Alignment of Rv3402c with DesI. The alignment was prepared with Clustal Omega. The conserved lysine in DesI that forms the Schiff base with the pyridoxal 5'-phosphate cofactor is highlighted in red. It corresponds to Lys 227 in Rv3402c. Those side chains in Des I that participate in hydrogen bonding interactions with the external aldimine are highlighted in yellow whereas those residues that provide backbone atoms for hydrogen bonding to the ligand are colored in green.

3.3 Materials and Methods

3.3.1 Cloning, expression, and purification of Rv3402c

The gene encoding Rv3402c from *M. tuberculosis* H37Rv was synthesized by Integrated DNA Technologies using codons optimized for protein expression in *E. coli*. The gene was placed into plasmid pKLD116, a pET31b derivative containing a His₆-tagged maltose-binding protein (MBP) followed by a TEV protease cleavage site (20). The pKLD116-*rv3402c* plasmid was utilized to transform Rosetta2(DE3) *E. coli* cells (Novagen). The cultures were grown in lysogeny broth supplemented with ampicillin and chloramphenicol (100 mg/L and 25 mg/L concentration, respectively) at 37 °C with shaking until an optical density of 0.9 was reached at 600 nm. The flasks were cooled in ice water, and the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and allowed to express protein at 21 °C for 20 h.

The cells were harvested by centrifugation and frozen as pellets in liquid nitrogen. The pellets were subsequently disrupted by sonication on ice in a lysis buffer composed of 50 mM sodium phosphate, 20 mM imidazole, 20% ethylene glycol, and 300 mM sodium chloride, pH 8.0. The lysate was cleared by centrifugation, and MBP-Rv3402c was purified at 4 °C utilizing nickel nitrilotriacetic acid resin (Qiagen) according to the manufacturer's instructions. All buffers were adjusted to pH 8.0 and contained 50 mM sodium phosphate, 300 mM sodium chloride, and imidazole concentrations of 25 mM for the wash buffer and 250 mM for the elution buffer. The column was first washed in buffer containing 20% ethylene glycol and subsequently washed with buffer lacking it. The protein was pooled and mixed in a 30:1 molar ratio (MBP-Rv3402c protein:TEV protease) based on a calculated extinction coefficient for the MBP-Rv3402c protein of 0.76 (mg/mL)⁻¹cm⁻¹. The recombinant protein was allowed to digest at 4 °C for 48 h. Uncleaved protein, His₆-MBP and the TEV protease were removed by two passages over Ni-nitrilotriacetic

acid resin. Cleaved protein was collected and dialyzed against 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl at 4 °C and concentrated to ~6 mg/mL based on a calculated extinction coefficient of 0.93 (mg/mL)⁻¹cm⁻¹.

3.3.2 Site-directed mutagenesis

The K227A variant of MBP-Rv3402c was generated via the QuikChange method of Stratagene. It was expressed and purified as described for the wild-type MBP-Rv3402c construct.

3.3.3 Enzymatic assay

Reactions containing 50 mM HEPPS (pH 8.0), 4 mM dTDP-D-glucose, 0.1 mM PLP, 40 mM sodium glutamate, and 0.5 mg/mL *E. coli* RmlB (a 4,6-dehydratase) were set up with or without 1 mg/mL Rv3402c or the K227A variant and allowed to incubate for 24 hours at room temperature. The reactions were stopped by removing the enzyme(s) via filtration through a 30 kDa cutoff filter. The mixtures were diluted 30x with water and examined by HPLC. Initially, small samples of the mixtures were examined with a 12-column volume, 0.0 – 1.0 M gradient of ammonium acetate (pH 4.0) on a 1 mL Resource-Q column. The reaction mixture containing Rv3402c demonstrated a decrease in the ketosugar (retention volume of 14.8 mL) concomitant with the appearance of a peak at 9.2 mL (approximately 300 mM ammonium acetate), corresponding to the expected dTDP-aminosugar. The reaction lacking Rv3402c resulted in no decrease in the peak at 14.8 mL. dTMP contamination was present in both samples (retention volume of 10.5 mL). Subsequently, the remainder of the Rv3402c reaction mixture was separated on a 6 mL Resource-Q column using the same gradient described above. Fractions containing the expected dTDP-aminosugar were pooled and lyophilized for ¹H NMR analysis. The required RmlB enzyme for the assay was purified in the laboratory.

3.3.4 Sugar product analysis

Prior to lyophilization of the dTDP-aminosugar, a 300 μ L sample was removed for mass spectrometry analysis. The identity of the Rv3402c product was confirmed as a dTDP-aminosugar by electrospray ionization mass spectrometry in the negative ion mode (Mass Spectrometry/Proteomic Facility at the University of Wisconsin).

3.3.5 NMR spectroscopy

NMR spectra (^1H , gCOSY, TOCSY, ^1H - ^{13}C HSQC, ^1H - ^{31}P HSQC) were recorded using a Bruker AVANCE III 600 MHz spectrometer at 25 $^{\circ}\text{C}$ in D_2O . Spectra were referenced to an internal acetone standard (δH 2.23 ppm and δc 31.45 ppm).

3.4 Results and Discussion

Rv3402c was cloned and over-expressed in *Escherichia coli* utilizing a MBP tag to aid in the *in vivo* protein folding and solubility. The tag was ultimately removed using TEV protease, and Rv3402c was first tested for activity using an HPLC assay that included it and dTDP-D-glucose, *E. coli* RmlB (which is functionally equivalent to *M. tuberculosis rv3464*), PLP, and glutamate (Scheme 3.1). As a negative control, the assay was also conducted in the absence of Rv3402c. Shown in Figures 3.2 (a) and 3.2 (b) are the corresponding HPLC traces without and with the addition of Rv3402c to the reaction mixture. As can be seen, in the absence of Rv3402c, one major peak is observed at 14.8 mL. On the basis of our laboratory standards, this peak corresponds to either dTDP-D-glucose or dTDP-4-keto-6-deoxy-D-glucose. The HPLC trace of the reaction mixture in the presence of Rv3402c, however, shows a new peak at 9.2 mL [Fig. 3.2 (b)]. A sample from this peak was subjected to electrospray ionization mass spectrometry in the negative ion mode. The spectrum is shown in Figure 1c. There is a major peak at a $m/z = 546.0883$, which would correspond to a dTDP-4-aminosugar in its protonated state.

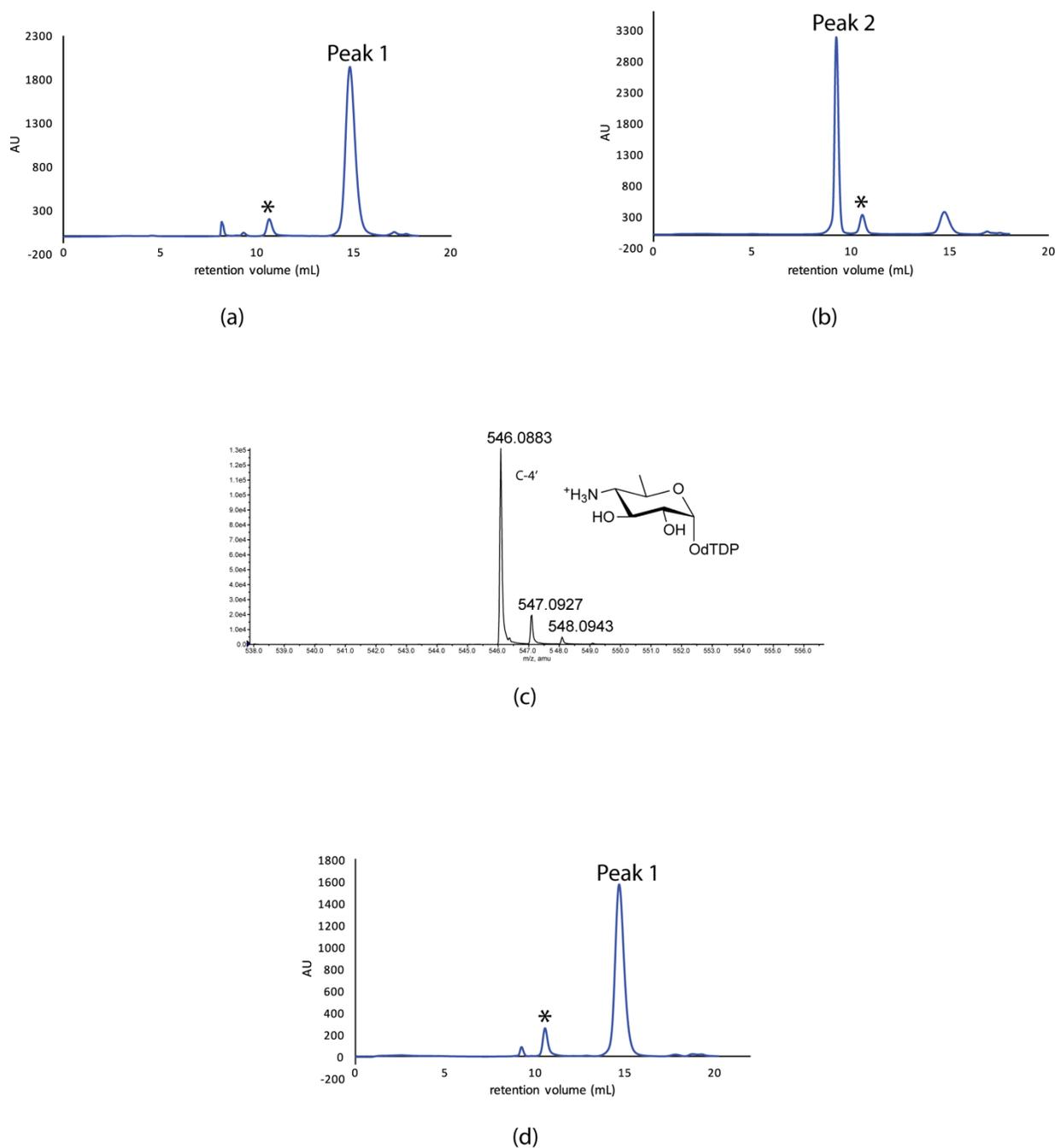


Figure 3.2 HPLC assay of Rv3402c activity and mass spectrum of the Rv3402c product. Aminotransferase activity was assessed in the absence (a) and presence (b) of Rv3402c. (a) Peak 1 (14.8 mL) corresponds to either dTDP-D-glucose or dTDP-4-keto-6-deoxy-D-glucose. (b) Peak 2 (9.2 mL) corresponds to a dTDP-aminosugar, although the actual stereochemistry about the C-4' carbon cannot be determined from these data. The peaks in (a), (b), and (d), marked by the asterisks, indicate the presence of dTMP, which is often a contaminant in dTDP-D-glucose preparations. Shown in (c) is the ESI mass spectrum corresponding to Peak 2 in the elution profile displayed in (b). Finally, the elution profile for the activity assay of the K227A protein variant is presented in (d).

Rv3402c was difficult to purify to homogeneity so in order to ensure that the production of the aminosugar was not due to the presence of contaminating *E. coli* aminotransferases, a variant of Rv3402c was produced via site-directed mutagenesis. Specifically, on the basis of amino acid sequence homology and the known catalytic properties of the Type I aminotransferases, Lys 227 in Rv3402c was predicted to be the conserved residue required for formation of the internal aldimine (18). Indeed, the HPLC trace of the reaction mixture containing the K227A variant [Fig. 3.2 (d)] is comparable to that of the reaction with no added aminotransferase. Whereas a small amount of aminosugar was produced by the K227A variant, there are reports of other aminotransferases that still show activity when their catalytic lysines are mutated. Indeed, PseC from *Helicobacter pylori* exhibited 12% activity when its catalytic lysine was mutated to an arginine (21). By comparison, the K227A Rv3402c variant exhibited approximately 2% activity of the wild-type protein, strongly suggesting that Lys 227 is necessary for catalysis, and that aminosugar formation is due to Rv3402c and not to a contaminating protein. Furthermore, contaminating proteins that persisted during purification were subjected to mass spectrometry analysis and determined to be the *E. coli* chaperones DnaJ and the large subunit of GroEL. Finally, the only known enzyme in the laboratory-strain of *E. coli* that uses dTDP-4-keto-6-deoxy-D-glucose as a substrate is WecE, which synthesizes an aminosugar having a different stereochemistry about the C-4' carbon position as compared to the Rv3402c product (22).

The HPLC assay and the mass spectrophotometric data indicate that Rv3402c catalyzes a PLP-dependent amination of dTDP-4-keto-6-deoxy-D-glucose. The final dTDP-sugar product was analyzed by 2D NMR (Table 3.1). The resulting monosaccharide had all vicinal coupling constants characteristic for the α -glucopyranose configuration and the ^{13}C chemical shifts expected for 4-amino-4,6-dideoxy- α -glucopyranoside, thus confirming that the Rv3402c product is dTDP-4-

amino-4,6-dideoxy- α -D-glucofuranose.

As noted in the introduction, previous biochemical studies have utilized a transposon library made in *M. tuberculosis* H37Rv to determine the genetic requirements for mycobacterial growth *in vitro* and for bacterial survival during infection (23). Both Rv0334 and Rv3464 (Scheme 3.1) were found to be essential for growth, which is consistent with their conservation across *Mycobacterium* spp. and with their role in the synthesis of L-rhamnose, an essential component of the mycobacterial cell envelope. Rv3402c and Rv3404c, however, were found to be nonessential for both *in vitro* and *in vivo* growth. It should be noted that Rv3402c and Rv3404c have a more restricted distribution than Rv0334 and Rv3464, being absent in rapidly growing *Mycobacterium* spp. as well as in *M. leprae*, *M. simiae* and *M. kansasii*. Rv3402c and Rv3404c are conserved within the *M. tuberculosis* complex (Table 3.2), except for strains with RD16 (region of difference 16) such as *M. bovis* BCG-Moreau where there is a deletion going from Rv3400 to Rv3405c (24). A non-sense mutation is observed in Rv3402c from lineages 5- and 6- as well as in animal lineages of the *M. tuberculosis* complex, causing it to be present as a pseudogene in these strains (25). Notably a significant homologue (71-75% identity) of Rv3402c is found in the *M. avium* complex whereas Rv3404c is absent in that group of strains.

Table 3.1 NMR Spectroscopic data of dTDP-4-Amino-4,6-Dideoxyglucose (dTDP-Qui4N) Produced Using Rv3402c.

	H/C-1 (ppm)	$J_{n,n+1}$ (Hz)	H/C-2 (ppm)	$J_{n,n+1}$ (Hz)	H/C-3 (ppm)	$J_{n,n+1}$ (Hz)	H/C-4 (ppm)	$J_{n,n+1}$ (Hz)	H/C-5 (ppm)	$J_{n,n+1}$ (Hz)	H/C-6 (ppm)
Qui4N	5.59	3.48	3.62	9.6	3.93	10.4	3.02	10.4	4.29	6.3	1.36
	96.2		73.0		69.8		58.1		66.8		18.0
2dRib	6.34		2.37 ^a		4.62		4.19		4.17 ^a		
	86.2		39.8		72.2		86.5		66.6		
Thymine											7.73
											138.5

^a Identical shifts were determined for the two protons at these positions.

Table 3.2 Summary of the Occurrence of dTDP-4-Formamido-4,6-Dideoxy-D-Glucose Biosynthetic Genes in *Mycobacterium* spp.

^a ORF#	Activity	Rapidly growing <i>Mycobacterium</i> spp.	<i>M. tuberculosis</i> complex	Comments
Rv0334	Glucose-1-phosphate thymidyltransferase	Conserved	Conserved	^b Essential for growth
Rv3464	dTDP-D-glucose 4,6-dehydratase	Conserved	Conserved	^b Essential for growth
Rv3402c	4-aminotransferase	Absent	Conserved	^c Impact on pathogenicity
Rv3404c	4- <i>N</i> -formyltransferase	Absent	Conserved	^d Impact on pathogenicity

^a According to *M. tuberculosis* H37Rv

^b Reference (23)

^c Reference (6)

^d Reference (26)

While it is not possible to conclude if the presence of Rv3402c and/or Rv3404c contributes to specific clinical presentations, there are a few studies suggesting that they have a role in determining the extent of virulence. Indeed, the *rv3402c* gene has been shown experimentally to be regulated by IdeR, an iron-responsive DNA-binding protein known to control iron acquisition (5). Notably, the mRNA level of the *rv3402c* gene was induced 17-fold in cultures of *M. tuberculosis* starved for iron, and it was induced 25-fold during infection of human THP-1 macrophages. Although Rv3402c clearly does not play a role in iron metabolism, it was postulated that *M. tuberculosis* might alter its membrane structure under iron-limiting conditions such as those found during intracellular growth or infection. Rv3402c was also found to enhance mycobacterial survival within macrophages and to modulate host pro-inflammatory cytokine production (6). Indeed, heterologous expression of the *rv3402c* gene in the non-pathogenic *M. smegmatis* strain showed that Rv3402c enhanced its intracellular survival in human and mice macrophages.

Importantly, the gene encoding Rv3404c was identified as one of 56 *M. tuberculosis* genes preferentially expressed in mouse lung in an investigation using a promoter trap strategy (27). This regulation may indicate a role for growth or persistence in that particular environment. In another study, McAdam *et al.* (26), generated insertional mutants in 351 different ORFs of *M. tuberculosis* H37Rv in order to identify virulence factors. Strikingly, SCID mice infected with a mutant form of Rv3404c presented significant increases in survival times when compared to mice infected with the wild-type strain, thus identifying Rv3404c as a virulence candidate for further investigation. The *rv3404c* gene was also distinguished as one of 21 genes whose expression was able to discriminate between isoniazid-, thiolactomycin-, or triclosan-treated *M. tuberculosis* (28). The exact role of Rv3404c in the response to drug treatment was not characterized, however.

The locations and structural characterizations of putative glycolipids or other glycoconjugates containing *N*-formylated sugars in *M. tuberculosis* remain major experimental challenges due to the complexity of the cell wall envelope. Nevertheless, the conservation of the genes involved in the bioproduction of dTDP-4-formamido-4,6-dideoxy-D-glucose in the *M. tuberculosis* complex and their impact on pathogenicity certainly support the targeting of this biosynthetic pathway as a strategy to develop novel anti-tuberculosis therapeutics.

3.5 References

1. Ernst, J. D. (2012) The immunological life cycle of tuberculosis. *Nat. Rev. Immunol.* *12*, 581-591.
2. Meena, L. S. and Rajni (2010) Survival mechanisms of pathogenic *Mycobacterium tuberculosis* H37Rv. *FEBS J.* *277*, 2416-2427.
3. Getahun, H., Matteelli, A., Chaisson, R. E., and Raviglione, M. (2015) Latent *Mycobacterium tuberculosis* Infection. *N. Engl. J. Med.* *372*, 2127-2135.
4. Forrellad, M. A., Klepp, L. I., Gioffré, A., Sabio y Garcia, J., Morbidoni, H. R., de la Paz Santangelo, M., Cataldi, A. A., and Bigi, F. (2013) Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence.* *4*, 1-64.
5. Gold, B., Rodriguez, G. M., Marras, S. A. E., Pentecost, M., and Smith, I. (2001) The *Mycobacterium tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Mol. Microbiol.* *42*, 851-865.
6. Li, W., Zhao, Q., Deng, W., Chen, T., Liu, M., and Xie, J. (2014) *Mycobacterium tuberculosis* Rv3402c enhances mycobacterial survival within macrophages and modulates the host pro-inflammatory cytokine production via NF-Kappa B/ERK/p38 signaling. *PLoS One.* *9*, e94418.
7. Dunsirn, M. M., Thoden, J. B., and Holden, H. M. (2017) Biochemical investigation of Rv3404c from *Mycobacterium tuberculosis*. *Biochemistry.* *56*, 3818-3825.
8. Thoden, J. B., Goneau, M. F., Gilbert, M., and Holden, H. M. (2013) Structure of a sugar *N*-formyltransferase from *Campylobacter jejuni*. *Biochemistry.* *52*, 6114-6126.
9. Zimmer, A. L., Thoden, J. B., and Holden, H. M. (2014) Three-dimensional structure of a sugar *N*-formyltransferase from *Francisella tularensis*. *Protein Sci.* *23*, 273-283.
10. Genthe, N. A., Thoden, J. B., Benning, M. M., and Holden, H. M. (2015) Molecular structure of an *N*-formyltransferase from *Providencia alcalifaciens* O30. *Protein Sci.* *24*, 976-986.
11. Woodford, C. R., Thoden, J. B., and Holden, H. M. (2015) New role for the ankyrin repeat revealed by a study of the *N*-formyltransferase from *Providencia alcalifaciens*. *Biochemistry.* *54*, 631-638.
12. Woodford, C. R., Thoden, J. B., and Holden, H. M. (2017) Molecular architecture of an *N*-formyltransferase from *Salmonella enterica* O60. *J. Struct. Biol.* *200*, 267-278.

13. Lacerda, T. L. S., Cardoso, P. G., Augusto de Almeida, L., da Cunha Camarg, I. L. B., Afonso, D. A. F., Trant, C. C., Macedo, G. C., Campos, E., Cravero, S. L., Salcedo, S. P., Gorvel, J. P., and Oliveira, S. C. (2010) Inactivation of formyltransferase (wbkc) gene generates a *Brucella abortus* rough strain that is annuuated in macrophages and in mice. *Vaccine*. *28*, 5627-5634.
14. Riegert, A. S., Chantigian, D. P., Thoden, J. B., Tipton, P. A., and Holden, H. M. (2017) Biochemical characterization of WbkC, an *N*-formyltransferase from *Brucella melitensis*. *Biochemistry*. *56*, 3657-3668.
15. Holden, H. M., Thoden, J. B., and Gilbert, M. (2016) Enzymes required for the biosynthesis of *N*-formylated sugars. *Curr. Opin. Struct. Biol.* *41*, 1-9.
16. Ma, Y., Stern, R. J., Scherman, M. S., Vissa, V. D., Yan, W., Jones, V. C., Zhang, F., Franzblau, S. G., Lewis, W. H., and McNeil, M. R. (2001) Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. *Antimicrob. Agents. Chemother.* *45*, 1407-1416.
17. Brown, H. A., Thoden, J. B., Tipton, P. A., and Holden, H. M. (2018) The structure of glucose-1-phosphate thymidyltransferase from *Mycobacterium tuberculosis* reveals the location of an essential magnesium ion in the RmlA-type enzymes. *Protein Sci.* *27*, 441-500.
18. Toney, M. D. (2014) Aspartate aminotransferase: an old dog teaches new tricks. *Arch. Biochem. Biophys.* *544*, 119-127.
19. Burgie, E. S. and Holden, H. M. (2007) Molecular architecture of DesI: a key enzyme in the biosynthesis of desosamine. *Biochemistry*. *46*, 8999-9006.
20. Rocco, C. J., Dennison, K. L., Klenchin, V. A., Rayment, I., and Escalante-Semerena, J. C. (2008) Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. *Plasmid*. *59*, 231-237.
21. Schoenhofen, I. C., Lunin, V. V., Julien, J. P., Li, Y., Ajamian, E., Matte, A., Cygler, M., Brisson, J. R., Aubry, A., Logan, S. M., Bhatia, S., Wakarchuk, W. W., and Young, N. M. (2006) Structural and functional characterization of PseC, an aminotransferase involved in the biosynthesis of pseudaminic acid, an essential flagellar modification in *Helicobacter pylori*. *J. Biol. Chem.* *281*, 8907-8916.
22. Hwang, B. Y., Lee, H. J., Yang, Y. H., Joo, H. S., and Kim, B. G. (2004) Characterization and investigation of substrate specificity of the sugar aminotransferase WecE from *E. coli* K12. *Chem. Biol.* *11*, 915-925.
23. Sasseti, C. M., Boyd, D. H., and Rubin, E. J. (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* *48*, 77-84.

24. Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S., and Small, P. M. (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science*. 284, 1520-1523.
25. Bentley, S. D., Comas, I., Bryant, J. M., Walker, D., Smith, N. H., Harris, S. R., Thurston, S., Gagneux, S., Wood, J., Antonio, M., Quail, M. A., Gehre, F., Adegbola, R. A., Parkhill, J., and de Jong, B. C. (2012) The genome of *Mycobacterium africanum* West African 2 reveals a lineage-specific locus and genome erosion common to the *M. tuberculosis* complex. *PLoS Negl. Trop. Dis.* 6, e1552.
26. McAdam, R. A., Quan, S., Smith, D. A., Bardarov, S., Betts, J. C., Cook, F. C., Hooker, E. U., Lewis, A. P., Woollard, P., Everett, M. J., Lukey, P. T., Bancroft, G. J., Jacobs Jr., W. R., and Duncan, K. (2002) Characterization of a *Mycobacterium tuberculosis* H37Rv transposon library reveals insertions in 351 ORFs and mutants with altered virulence. *Microbiology*. 148, 2975-2986.
27. Dubnau, E., Chan, J., Mohan, V. P., and Smith, I. (2005) Responses of *Mycobacterium tuberculosis* to growth in the mouse lung. *Infect. Immun.* 73, 3754-3757.
28. Betts, J. C., McLaren, A., Lennon, M. G., Kelly, F. M., Lukey, P. T., Blakemore, S. J., and Duncan, K. (2003) Signature gene expression profiles discriminate between isoniazid-, thiolactomycin-, and triclosan-treated *Mycobacterium tuberculosis*. *Antimicrob. Agents. Chemother.* 47, 2903-2913.

Chapter 4:

Biochemical studies on WbcA, a sugar epimerase from *Yersinia enterocolitica*

This work was previously published as: Salinger, A. J., Brown, H. A., Thoden, J. B., and Holden, H. M. (2015) Biochemical studies on WbcA, a sugar epimerase from *Yersinia enterocolitica*. *Protein Sci.* 24, 1633-1639.

4.1 Abstract

Yersinia enterocolitica is a Gram-negative bacterium that causes yersiniosis, a zoonotic disease affecting the gastrointestinal tract of humans, cattle, and pigs, among others. The lipopolysaccharide of *Y. enterocolitica* O:8 contains an unusual sugar, 6-deoxy-D-gulose, which requires four enzymes for its biosynthesis. Here, we describe a combined structural and functional investigation of WbcA, which catalyzes the third step in the pathway, namely an epimerization about the C-3' carbon of a CDP-linked sugar. The structure of WbcA was determined to 1.75-Å resolution, and the model was refined to an overall *R*-factor of 19.5%. The fold of WbcA places it into the well-defined cupin superfamily of sugar epimerases. Typically, these enzymes contain both a conserved histidine and a tyrosine residue that play key roles in catalysis. On the basis of amino acid sequence alignments, it was anticipated that the “conserved” tyrosine had been replaced with a cysteine residue in WbcA (Cys 133), and indeed this was the case. However, what was not anticipated was the fact that another tyrosine residue (Tyr 50) situated on a neighboring β -strand moved into the active site. Site-directed mutant proteins were subsequently constructed, and their kinetic properties analyzed to address the roles of Cys 133 and Tyr 50 in WbcA catalysis. This study emphasizes the continuing need to experimentally verify assumptions that are based solely on bioinformatics approaches.

4.2 Introduction

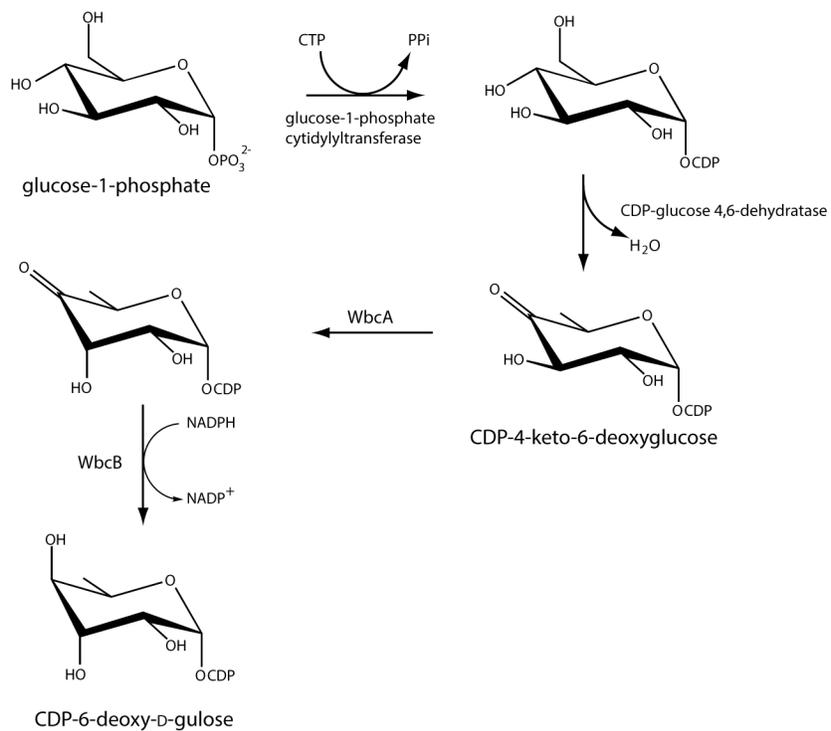
Yersinia is a genus consisting of enteric rod-shaped Gram-negative bacteria. Of the various species identified thus far, only three are known human pathogens: *Yersinia pestis*, the causative agent of bubonic plague, and *Yersinia pseudotuberculosis* and *Yersinia enterocolitica*, both of which can lead to severe gastrointestinal distress (1). Like most Gram-negative bacteria, members of this genus contain on their outer membranes complex glycoconjugates referred to as lipopolysaccharides. These remarkable molecules consist of a Lipid A component, a core oligosaccharide, and an O-antigen, which often contains highly unusual deoxysugars. Whereas the core oligosaccharides are typically conserved within a specific genus, the O-antigens are quite variable and clearly contribute to the wide species variation observed in nature (2). Much research has focused on the Lipid A component of the lipopolysaccharide and its role in bacterial virulence (3-5). Recent investigations are beginning to unravel the biological role of the O-antigen in bacterial toxicity as well. Indeed, studies suggest that the O-antigen of *Y. enterocolitica* O:8 is essential for virulence (6).

In 1997, the chemical composition of the O-antigen from *Y. enterocolitica* O:8 was reported to consist of repeat units containing *N*-acetyl-D-galactosamine, D-galactose, D-mannose, L-fucose, and the unusual sugar, 6-deoxy-D-gulose (7). Subsequently, on the basis of genome analysis, the biosynthetic pathway for the production of 6-deoxy-D-gulose was suggested as presented in Scheme 4.1 (8). The first step involves the attachment of a nucleoside monophosphate to glucose-1-phosphate, a reaction catalyzed by glucose-1-phosphate cytidyltransferase. In the next step, the C-6' hydroxyl is removed and the C-4' hydroxyl is converted to a keto moiety by the action of CDP-glucose 4,6-dehydratase. Both the cytidyltransferase and the dehydratase have been well characterized structurally and biochemically from either *Salmonella typhi* or *Y. pseudotuberculosis*

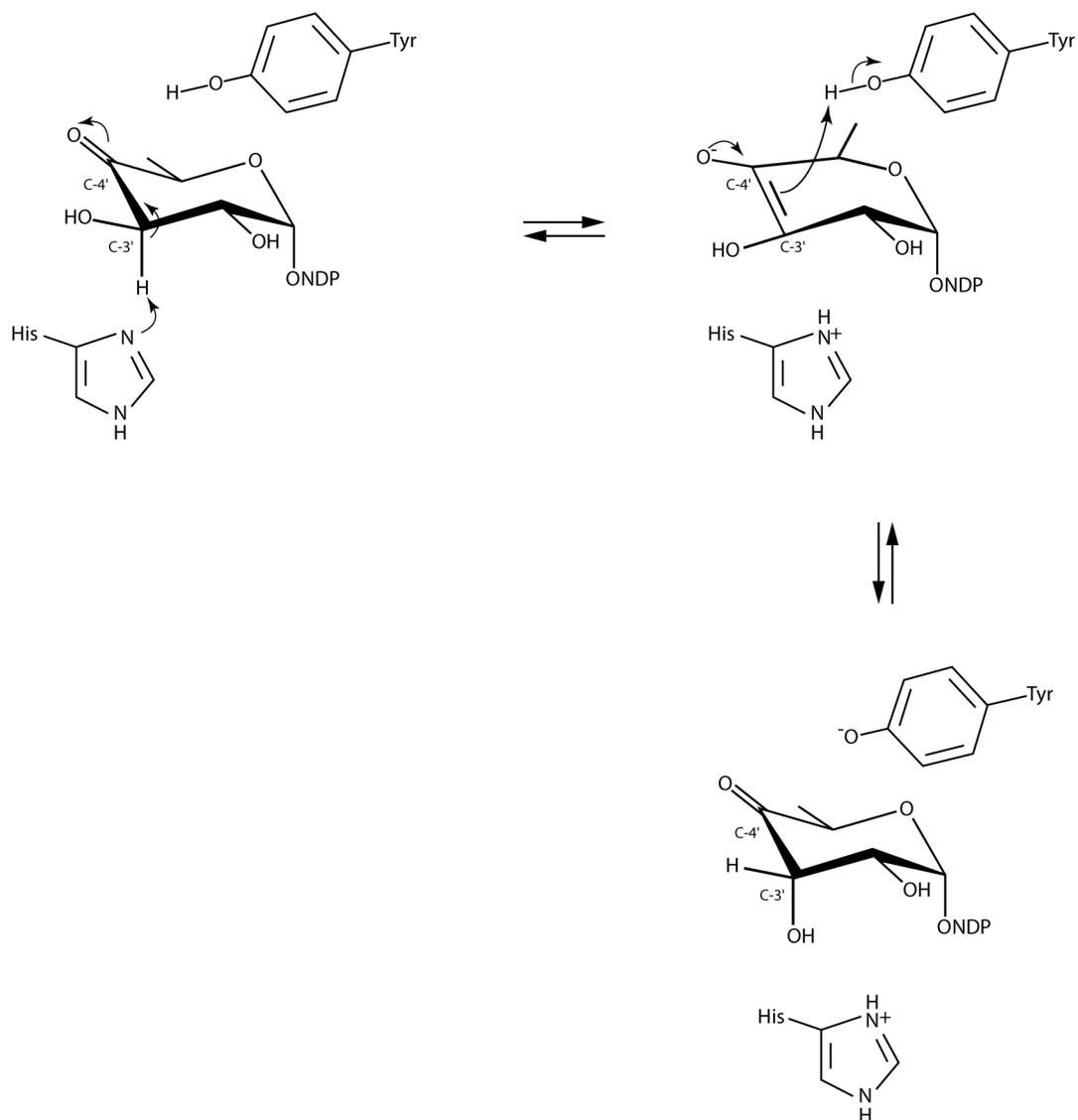
(9-12). The third step in the pathway, namely a C-3' epimerization, is reportedly catalyzed by WbcA, the focus of this investigation. The pathway is thought to be completed by the action of WbcB, a C-4' ketoreductase.

On the basis of amino acid sequence alignments, WbcA was thought to belong to the “cupin” family of sugar epimerases that catalyze reactions about the C-3' or C-5' carbons. Some function as monoepimerases whereas others catalyze epimerization reactions about both carbons. In all the enzymes that have been structurally characterized to date, two conserved residues, a histidine and a tyrosine, are thought to function as catalytic bases and acids, respectively (13-18). The reaction mechanism for a C-3' monoepimerase is outlined in Scheme 4.2. The conserved histidine abstracts the proton from the C-3' carbon resulting in an enolate anion intermediate. The conserved tyrosine provides a proton to the opposite face of the pyranosyl ring resulting in the epimerization about the C-3' carbon. It is not known whether the reaction mechanism is stepwise or concerted.

According to amino acid sequence alignments (Fig. 4.1), the conserved tyrosine in WbcA is apparently replaced with a cysteine residue (Cys 133). In order to explore the active site architecture of a sugar epimerase lacking the conserved tyrosine, we initiated a combined structural and functional investigation of WbcA as reported herein. For this analysis, the structure of WbcA was determined to 1.75-Å resolution. Strikingly, the structure revealed that whereas the conserved histidine (His 63) was present in WbcA, the cysteine residue was not in the appropriate position to function as an active site acid. Rather, a tyrosine residue (Tyr 50), located on a neighboring β -strand, moved into the proper position. Site-directed mutant proteins were subsequently constructed and their kinetic parameters measured to confirm the catalytic roles of His 63, Tyr 50, and Cys 133. Details concerning the structure and kinetic properties of WbcA are presented.



Scheme 4.1 Predicted pathway for the production of CDP-6-deoxy-D-gulose.



Scheme 4.2 Possible catalytic mechanism for a C-3' monoepimerase.

```

ChmJ      ----MHPLSIEGAWSQEPVIHSDHRGRSHEWFRGESFRQAFG--HDFPVAQVNVAVSHRGA 55
NovW      --MRLRPLGIEGVWEITPEQRADPRGVFLDWYHVDRFAEAIG--RPLRLAQANLSVSVRGV 57
EvaD      --MQARKLAVDGAIEFTPRVFADDRGLLILPYQEEAFVEAHGGPLFRVAQTIHSMKRGV 58
RmlC      -MMIVIKTAIPDVLILEPKVFGDERGFFFESYNQQTFEELIG--RKVTFVQDNHSSKKNV 58
WbcA      MIINITELNISGCYLIESPIFSDERGEFVKTHHQEIFKNFGL--EIPSAEYYRSKNNV 58

          #
ChmJ      LRGIHYTEIPPGQAKYSVCVRGAGLDVVVDVRIGSPTFGRWEIVPMDAERNTAVYLTAGL 115
NovW      VRGIHFVDVPPGQAKYVTCVRGAVFDVVVDLVRGSPTYGCWEGTRLDDVSRRAVYLSEGI 117
EvaD      VRGIHYTVTPPGTAKYVYCARGKAMDIVIDIRVGSPTFGQWDSVLMQDDPRAVYLPVGV 118
RmlC      LRGLHFQRGENAQGKLVRCVAVGEVFDVAVDIRKESPTFGQWGVNLSAENKRQLWIPEGF 118
WbcA      IRGMHFQQYPDDHNLKLVFCPEGEVLDVFLDIRKDSNTYGQFMSFILNPHNRRSIFLAKGI 118

          *
ChmJ      GRAFLSLTDDATLVYLCSSGYAPAREHSVNPLDPDLGIAWPDDIEPLLSDRDENAPTAT 175
NovW      GHGFCAISDEATLCYLSSGTYDPATEHGVHPLDPELAIDWPTG--TPLLSPRDQDALLLAE 176
EvaD      GHAFVALEDDTVMSYMLSRSYVTQDELALSALDPALGLPIDIGVEPIVSDRDRVAITLAE 178
RmlC      AHGFVTLSEYAEFLYKATNYSPSPSEGSILWNDEAIGIEWPFSQLPELSAKDAAAPLLDQ 178
WbcA      AHGFLSMKDNTLIVCKTSTVHSPSRDSGIHWN--SFGFKWPVE--NPIISDKDRNLDCFF- 174

ChmJ      AERLGLLPTYQAWQEQQAQR----- 196
NovW      ARDAGLLPTYATCQAVTVPSPAPGSVGDGP 207
EvaD      AQRQGLLPDYTTSQEIERRLTA----- 200
RmlC      ALLTE----- 183
WbcA      -----

```

Figure 4.1 Amino acid sequence alignment of WbcA and other cupin family sugar epimerases. Enzymes that have been structurally characterized were aligned and include the 3'-monoepimerases ChmJ from *Streptomyces bikiniensis* (18) and NovW from *Streptomyces spheroids* (16, 17), the 5'-monoepimerase EvaD from *Amycolatopsis orientalis* (15), and the 3',5'-diepimerase RmlC from *Salmonella enterica* (19). The conserved active site histidine and tyrosine are highlighted in yellow. Notably, the anticipated tyrosine in WbcA aligns as a cysteine (C133).

4.3 Materials and Methods

4.3.1 Cloning of *wbcA* from *Yersinia enterocolitica* (type O:8)

The *wbcA* gene from *Y. enterocolitica* ATCC 9610 served as the starting template for PCR using Platinum Pfx DNA polymerase (Invitrogen). Primers were designed that incorporated NheI and XhoI restriction sites. The PCR product was digested with NheI and XhoI and ligated into pET28T, a laboratory pET28b(+) vector that had been previously modified to incorporate a TEV protease cleavage recognition site after the N-terminal polyhistidine tag (20).

4.3.2 Protein expression and purification

The pET28T-*wbcA* plasmid was utilized to transform Rosetta2(DE3) *Escherichia coli* cells (Novagen). The cultures were grown in lysogeny broth supplemented with kanamycin (35 $\mu\text{m}/\text{mL}$) and chloramphenicol (25 $\mu\text{m}/\text{mL}$) at 37 °C with shaking until an optical density of 0.8 was reached at 600 nm. The flasks were cooled in an ice bath, and protein expression was initiated with the addition of 1 mM isopropyl β -D-1-thiogalactopyranoside.

The cells were allowed to express protein at 16 °C for 20 h and then harvested by centrifugation, flash frozen, and subsequently disrupted by sonication on ice. The lysate was cleared by centrifugation, and WbcA was purified with Ni-NTA resin (Qiagen) according to the manufacturer's instructions. The histidine tag was removed by digestion with TEV protease. The protein was dialyzed against 10 mM Tris-HCl (pH 8.0) and concentrated to 18 mg/mL based on an extinction coefficient of 1.1 (mg/mL)⁻¹cm⁻¹. The 4-ketoreductase (WbcB) required for the kinetic assays and the CDP-sugar synthesis was cloned and purified in a similar manner.

CDP-6-deoxy-D-gulose was enzymatically prepared using 30 mL reactions containing 1.2 M CDP-glucose, 0.5 mg/mL CDP-D-glucose 4-6 dehydratase (12), 1 mg/mL WbcA, 0.5 mg/mL WbcB (Scheme 4.1), 5 mM NADPH, and 50 mM HEPPS (pH 8). The reaction was incubated for

6 hours at 21 °C. Enzymes were removed by passage through an Amicon Ultra 10 kDa cutoff membrane. The reaction mixture was diluted and loaded onto a HiLoad 26/10 Q-sepharose column and purified by HPLC at a buffered pH of 4.0 with an ammonium acetate gradient of 0 – 1.2 M over 12 column volumes. The product sugar, which eluted at approximately 600 mM ammonium acetate, was pooled and lyophilized.

4.3.3 Crystallization and structural analysis

Crystallization conditions were surveyed by the hanging drop method of vapor diffusion using a laboratory based sparse matrix screen at both room temperature and 4 °C. X-ray diffraction quality crystals of the enzyme were subsequently grown from precipitant solutions containing 14 to 18% *O*-methyl-poly(ethylene glycol) 5000, 5 mM CDP-6-deoxy-D-gulose, 2% glycerol, and 100 mM MOPS (pH 7.0) at 4 °C. The crystal morphology was hexagonal with largest crystals showing deep indentations. Moderately sized crystals (~0.4 mM in length) were used for X-ray data collection. The crystals belonged to the space group *C2* with unit cell dimensions of $a = 79.2$, $b = 42.7 \text{ \AA}$, $c = 150.5 \text{ \AA}$, and $\beta = 90.4^\circ$. The asymmetric unit contained three subunits.

For X-ray data collection, the crystals were transferred gradually to a cryoprotectant solution containing 25% *O*-methyl-poly(ethylene glycol) 5000, 100 mM NaCl, 16% glycerol, 5 mM CDP-6-deoxy-D-gulose, and 100 mM MOPS (pH 7.0). X-ray data were collected at 100 K with a Bruker AXS Platinum 135 CCD detector controlled by the Proteum software suite (Bruker AXS Inc.). The X-ray source was Cu $K\alpha$ radiation from a Rigaku RU200 X-ray generator equipped with Montel optics and operated at 50 kV and 90 mA. These X-ray data were processed with SAINT version 7.06A (Bruker AXS Inc.) and internally scaled with SADABS version 2005/1 (Bruker AXS Inc.). Relevant X-ray data collection statistics are listed in Table 4.1. The structure was determined via molecular replacement with the software package PHASER and using as a search

model the X-ray coordinates 4HN0 from the Protein Data Bank (18). Iterative cycles of model building with COOT and refinement with REFMAC reduced the R_{work} and R_{free} to 19.5% and 24.8%, respectively, from 30 to 1.75 Å resolution (21-23).

Table 4.1 X-ray Data Collection Statistics and Model Refinement Statistics.

	WbcA
Resolution limits (Å)	30.0-1.75 (1.80-1.75) ^b
Number of independent reflections	50180 (3422)
Completeness (%)	98.1 (95.0)
Redundancy	2.5 (1.5)
avg I/avg $\sigma(I)$	6.9 (2.0)
R_{sym} (%) ^a	8.2 (34.1)
^c R -factor (overall)%/no. reflections	19.5/50180
R -factor (working)%/no. reflections	19.2/47639
R -factor (free)%/no. reflections	24.8/2541
number of protein atoms	4285
number of heteroatoms	445
Average B values	
protein atoms (Å ²)	20.5
solvent (Å ²)	26.9
Weighted RMS deviations from ideality	
bond lengths (Å)	0.011
bond angles (°)	2.0
planar groups (Å)	0.01
Ramachandran regions (%)^d	
most favored	92.9
additionally allowed	7.1
generously allowed	0.0

^a $R_{\text{sym}} = (\sum |I - \bar{I}| / \sum I) \times 100$.

^bStatistics for the highest resolution bin.

^c R -factor = $(\sum |F_o - F_c| / \sum |F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

^dDistribution of Ramachandran angles according to PROCHECK (24).

4.3.4 Site-directed mutagenesis

All site-directed mutant proteins were generated via the QuikChange method of Stratagene and purified in a similar manner to that of the wild-type enzyme including the removal of the N-terminal histidine tags.

4.3.5 Kinetic analyses

Steady-state kinetic parameters for WbcA, and the site-directed mutant variants were determined via a coupled spectrophotometric assay, which followed the conversion of NADPH to NADP⁺ by the action of WbcB (Scheme 4.1). The starting substrate was CDP-4-keto-6-deoxyglucose. All reaction mixtures contained 50 mM HEPES (pH 8), 0.2 mM NADPH, and 1 mg/mL WbcB.

The CDP-4-keto-6-deoxyglucose and WbcA wild-type and mutant protein concentrations varied between reactions as follows. For the wild-type enzyme, the WbcA concentration was 0.0025 mg/mL in the reaction mixture, with CDP-4-keto-6-deoxyglucose concentrations ranging between 0.001 and 8 mM. The C133S variant required a concentration of 0.018 mg/mL, and the ketosugar concentration was varied between 0.05 and 8 mM. Both the H63N and Y50F mutant proteins were assayed at 5 mg/mL, with CDP-4-keto-6-deoxyglucose concentrations ranging from 0.1 to 20 mM. However, no activity was seen above background.

All reactions were initiated by the addition of WbcA at 25 °C and followed using a Beckman Coulter DU-640 spectrophotometer for 5 min. Reduction of the CDP-4-keto-6-deoxyglucose sugar and concurrent oxidation of NADPH to NADP⁺ was monitored by a decrease in absorbance at 340 nm. Plots of concentrations versus initial rates were analyzed using PRISM (GraphPad Software, Inc.) and were fitted to the equation $v_o = (V_{max}[S]) / (K_M + [S])$. Kinetic parameters are listed in Table 4.2.

Table 4.2 Kinetic Parameters.

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1} s^{-1}$)
Wild-type	0.69 ± 0.08	7.8 ± 0.3	1.1×10^4
H63N	nd ^a	nd	nd
C133S	0.61 ± 0.12	2.0 ± 0.1	3.3×10^3
Y50F	nd	nd	nd

^aCould not be determined under the assay conditions employed.

4.4 Results and Discussion

The WbcA model reported here was solved to 1.75-Å resolution and refined to an overall *R*-factor of 19.5%. Relevant X-ray data collection and model refinement statistics are given in Table 4.1. Like other members of the cupin superfamily, WbcA functions as a dimer. The crystals utilized for this analysis belonged to the space group *C2* with three monomers in the asymmetric unit. Two of the monomers forming the dimer were related by a local twofold rotational axis whereas the third monomer packed along a crystallographic dyad. All three subunits adopted similar molecular architectures such that their α -carbons superimpose with root-mean-square deviations of between 0.20 Å and 0.25 Å. For the sake of clarity, the following discussion refers to the dimer that packed along the local twofold rotational axis.

Shown in Figure 4.2 (a) is a ribbon representation of the WbcA dimer, which has overall dimensions of ~ 50 Å x 55 Å x 60 Å. The buried surface area is 3800 Å². As observed for members of the cupin fold, domain swapping occurs whereby a β -hairpin motif from one subunit projects towards the active site of another subunit. A close-up stereo view of one subunit is displayed in Figure 4.2 (b). The polypeptide chain initiates with an extended coil and two anti-parallel β -strands. The β -hairpin motif (Ile 20 – Val 29), responsible for domain swapping, is followed by the first of two α -helices. An additional nine β -strands wrap around to form the cupin barrel. The barrel can be envisioned as two layers of anti-parallel β -sheet. There is a final small α -helix formed by Asp 165 to Arg 168. Both Met 62 and Pro 68 adopt the *cis* conformation.

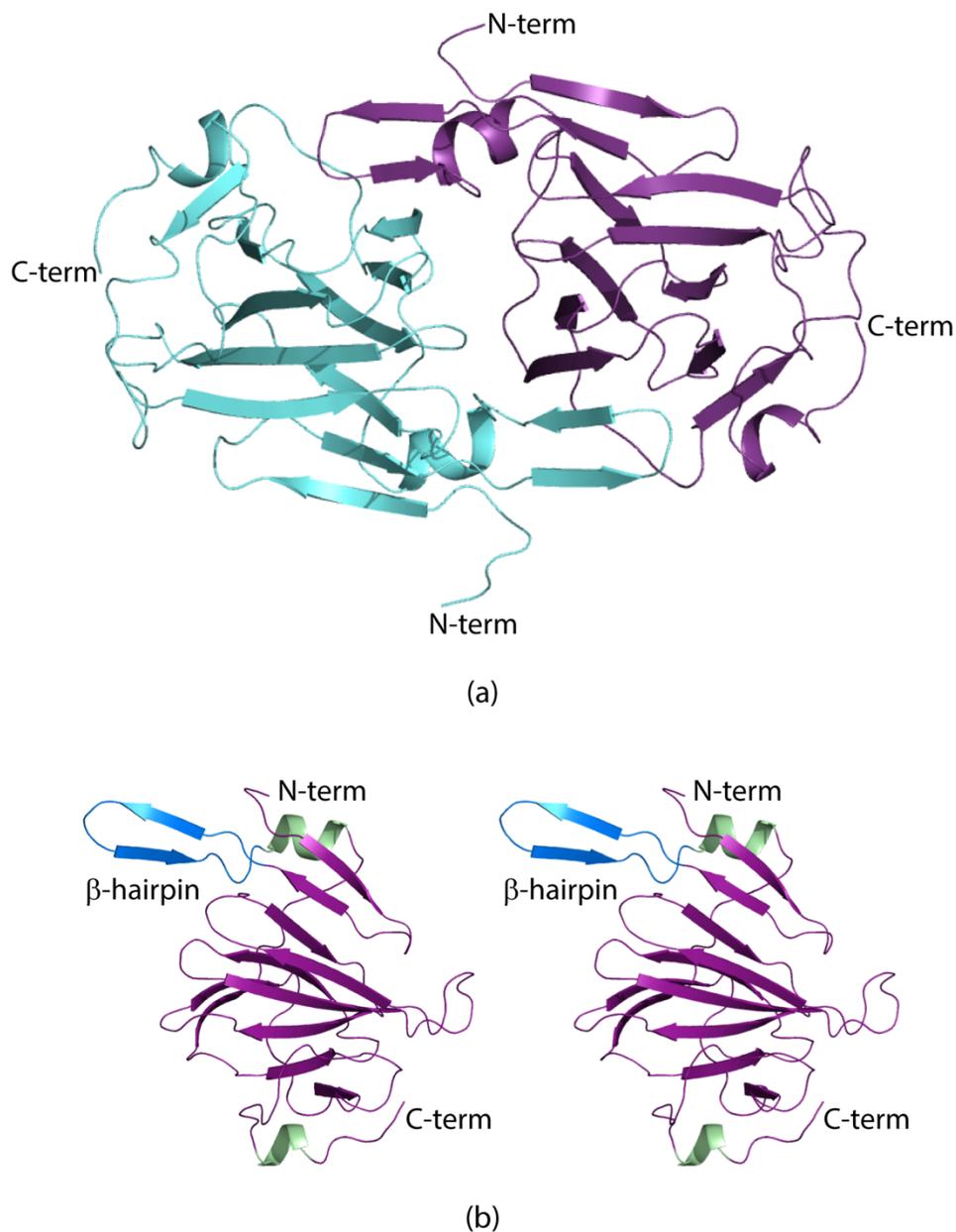


Figure 4.2 The structure of WbcA. A ribbon representation of the WbcA dimer is shown in (a) with the two subunits displayed in teal and purple, respectively. A stereo view of one subunit is presented in (b). The two α -helices and the β -hairpin motif are highlighted in green and blue, respectively. This figure and Figure 4.3 were prepared with PyMOL (25).

The crystallization conditions for WbcA included 5 mM CDP-6-deoxy-D-glucose in an attempt to trap a stable substrate analog in the active site. No clear electron density was observed for the CDP-sugar ligand. However, the structure of another C-3' epimerase from *Streptomyces bikiniensis* was determined in the laboratory in 2012 as a complex with dTDP-6-deoxy-D-glucose (18). This enzyme, referred to as ChmJ, has an amino acid sequence identity and similarity to WbcA of 30% and 50%, respectively. The α -carbons for these two enzymes correspond with a root-mean-square deviation of 1.2 Å. In ChmJ, His 60 serves as the catalytic base whereas Tyr 130 functions as the active site acid (18). Shown in Figure 4.3 is a superposition of His 60 and Tyr 130 of ChmJ onto the structural equivalent residues in WbcA. The histidine residues in WbcA (His 63) and ChmJ (His 60) align well, and in both enzymes the residue leading up to the histidine adopts the *cis* conformation (Met 62 in WbcA or Ile 59 in ChmJ). As noted in the Introduction, it was assumed that Cys 133 in WbcA would align with Tyr 130 in ChmJ, and indeed it does as can be seen in Figure 4.3. What was not expected, however, was that in WbcA Tyr 50, which is an asparagine residue in ChmJ, would project into the active site in the appropriate position to serve as the active site acid. This residue is situated on β -strand 5 whereas in ChmJ the active site tyrosine is located on β -strand 12.

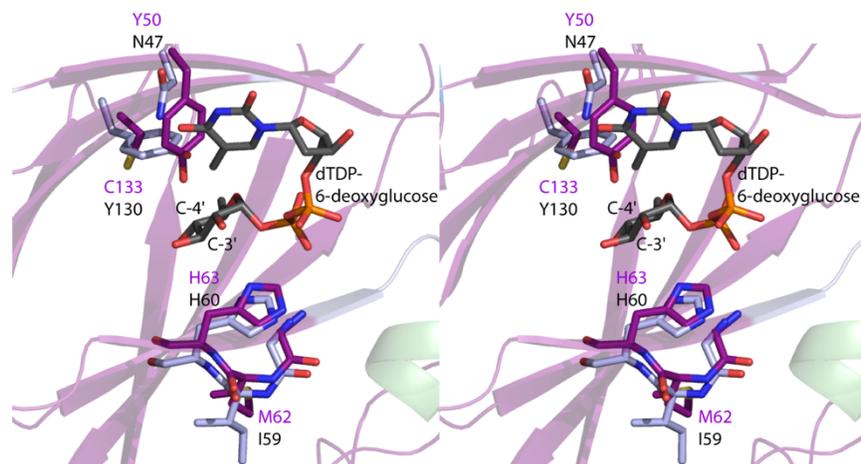
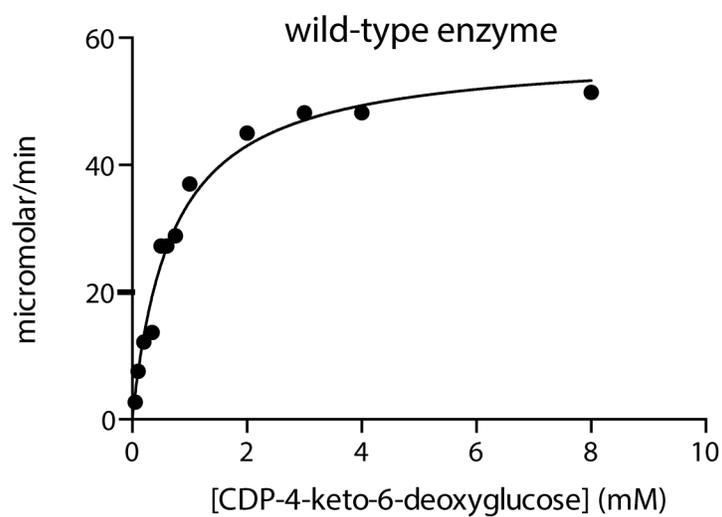


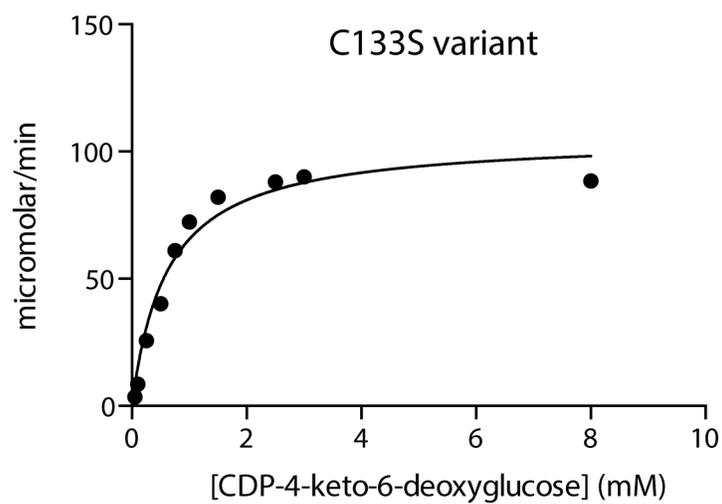
Figure 4.3 Superposition of the active sites for WbcA and ChmJ. Those residues belonging to WbcA are highlighted in purple whereas those from ChmJ are displayed in light blue. The top and bottom labels refer to amino acids in WbcA and ChmJ, respectively. The dTDP-6-deoxyglucose ligand is from the ChmJ model (PDB code 4HMZ).

To further test the roles of His 63, Tyr 50, and Cys 133 in the catalytic activity of WbcA, the following site-directed mutant variants were constructed: H63N, Y50F, and C133S. Their kinetic parameters were measured and are presented in Table 4.2. Plots of initial velocities versus substrate concentrations are presented in Figure 4.4. As expected, substituting His 63 with an asparagine residue resulted in a catalytically inert enzyme under the assay conditions employed. Likewise, replacement of Tyr 50 with a phenylalanine residue resulted in the complete abrogation of enzymatic activity. In the case of the C133S mutation, however, the K_M for the substrate was identical to that of the wild-type enzyme within experimental error. The overall k_{cat} was slightly reduced. As a consequence, the C133S mutant protein demonstrated an approximate threefold reduction in catalytic efficiency. This small reduction in catalytic efficiency is not unexpected given the change in overall hydrophobicity of the active site upon substitution of a cysteine to a serine residue.

In summary, the overall molecular architecture of WbcA has been defined. Importantly, this structural/biochemical analysis is yet another example of the perils in relying only on amino acid sequence homologies. Indeed, it was expected from such analyses that Cys 133 was the active site acid, and in fact it is a tyrosine residue from a neighboring β -strand.



(a)



(b)

Figure 4.4 Plots of initial velocities versus substrate concentration.

4.5 References

1. Brubaker, R. R. (1991) Factors promoting acute and chronic diseases caused by yersiniae. *Clin. Microbiol. Rev.* *4*, 309-324.
2. Jansson, P. E., Lindberg, B., Lindberg, A. A., and Wollin, R. (1981) Structural studies on the hexose region of the core in lipopolysaccharides from Enterobacteriaceae. *Eur. J. Biochem.* *115*, 571-577.
3. Galanos, C., Lüderitz, O., Rietschel, E. T., Westphal, O., Brade, H., Brade, L., Freudenberg, M., Schade, U., Imoto, M., Yoshimura, H., Kusumoto, S., and Shiba, T. (1985) Synthetic and natural *Escherichia coli* free lipid A express identical endotoxic activities. *Eur. J. Biochem.* *148*, 1-5.
4. Raetz, C. R. H. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* *71*, 635-700.
5. Raetz, C. R. H., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007) Lipid A modification systems in gram-negative bacteria. *Annu. Rev. Biochem.* *76*, 295-329.
6. Bengoechea, J. A., Najdenski, H., and Skurnik, M. (2004) Lipopolysaccharide O antigen status of *Yersinia enterocolitica* O:8 is essential for virulence and absence of O antigen affects the expression of other *Yersinia* virulence factors. *Mol. Microbiol.* *52*, 451-469.
7. Tomshich, S. V., Gorshkova, R. P., and Ovodov, Y. S. (1987) Structural studies on lipopolysaccharide from *Y. enterocolitica* serovar O:8. *Khimia Prirodnikh Soedinenii.* 657-664.
8. Zhang, L., Radziejewska-Lebrecht, J., Krajewska-Pietrasik, D., Toivanen, P., and Skurnik, M. (1997) Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O:8. *Mol. Microbiol.* *23*, 63-76.
9. Koropatkin, N. M. and Holden, H. M. (2004) Molecular structure of α -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *279*, 44023-44029.
10. Koropatkin, N. M., Cleland, W. W., and Holden, H. M. (2005) Kinetic and structural analysis of α -D-glucose-1-phosphate cytidyltransferase from *Salmonella typhi*. *J. Biol. Chem.* *280*, 10774-10780.
11. Vogan, E. M., Bellamacina, C., He, X., Liu, H. W., Ringe, D., and Petsko, G. A. (2004) Crystal structure at 1.8 Å Resolution of CDP-D-glucose 4,6-dehydratase from *Yersinia pseudotuberculosis*. *Biochemistry* *43*, 3057-3067.
12. Koropatkin, N. M. and Holden, H. M. (2005) Structure of CDP-D-glucose 4,6-dehydratase from *Salmonella typhi* complexed with CDP-D-xylose. *Acta. Crystallogr. D. Biol. Crystallogr.* *61*, 365-373.

13. Naismith, J. H. (2004) Chemical insights from structural studies of enzymes. *Biochem. Soc. Trans.* *32*, 647-654.
14. Dong, C., Major, L. L., Srikannathasan, V., Errey, J. C., Giraud, M. F., Lam, J. S., Graninger, M., Messner, P., McNeil, M. R., Field, R. A., Whitfield, C., and Naismith, J. H. (2007) RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J. Mol. Biol.* *365*, 146-159.
15. Merkel, A. B., Major, L. L., Errey, J. C., Burkart, M. D., Field, R. A., Walsh, C. T., and Naismith, J. H. (2004) The position of a key tyrosine in dTDP-4-keto-6-deoxy-D-glucose-5-epimerase (EvaD) alters the substrate profile for this RmlC-like enzyme. *J. Biol. Chem.* *279*, 32684-32691.
16. Jakimowicz, P., Tello, M., Meyers, C. L. F., Walsh, C. T., Buttner, M. J., Field, R. A., and Lawson, D. M. (2006) The 1.6-Å resolution crystal structure of NovW: A 4-keto-6-deoxy sugar epimerase from the novobiocin biosynthetic gene cluster of *Streptomyces spheroides*. *Proteins.* *63*, 261-265.
17. Tello, M., Jakimowicz, P., Errey, J. C., Meyers, C. L. F., Walsh, C. T., Buttner, M. J., Lawson, D. M., and Field, R. A. (2006) Characterisation of *Streptomyces spheroides* NovW and revision of its functional assignment to a dTDP-6-deoxy-D-xylo-4-hexulose 3-epimerase. *Chem. Commun. (Camb).* *14*, 1079-1081.
18. Kubiak, R. L., Phillips, R. K., Zmudka, M. W., Ahn, M. R., Maka, E. M., Pyeat, G. L., Roggensack, S. J., and Holden, H. M. (2012) Structural and functional studies on a 3'-epimerase involved in the biosynthesis of dTDP-6-deoxy-D-allose. *Biochemistry.* *51*, 9375-9383.
19. Giraud, M. F., Leonard, G. A., Field, R. A., Berlind, C., and Naismith, J. H. (2000) RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat. Struct. Biol.* *7*, 398-402.
20. Thoden, J. B. and Holden, H. M. (2005) The molecular architecture of human N-acetylgalactosamine kinase. *J. Biol. Chem.* *280*, 32784-32791.
21. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* *60*, 2126-2132.
22. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta. Crystallogr. D. Biol. Crystallogr.* *66*, 486-501.
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D. Biol. Crystallogr.* *53*, 240-255.
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* *26*, 283-291.

25. DeLano, W. L., The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. 2002.

4.6 Acknowledgement

We gratefully acknowledge Dr. Evgeny Vinogradov at Human Health Therapeutics, National Research Council, Canada for confirming by NMR the identity of the CDP-6-deoxy-D-gulose ligand.

Chapter 5:

Insights into *O*-Methyl Phosphoramidate Biosynthesis in *Campylobacter jejuni*

5.1 Abstract

Campylobacter jejuni is a Gram-negative bacterium that represents one of the most frequent causes of food-borne gastroenteritis. Intriguingly, the most common infection preceding the onset of Guillain-Barré syndrome (GBS), an autoimmune disorder leading to acute paralysis, is that caused by *C. jejuni*. Like many pathogens, *C. jejuni* produces a capsular polysaccharide (CPS) that serves as a virulence marker in the organism. The capsule is often decorated with non-stoichiometric modifications, including an *O*-methyl phosphoramidate (MeOPN) moiety that is beneficial during early stages of infection for immune evasion. A mechanistic understanding of MeOPN production is therefore merited to target the pathway and potentially control *C. jejuni* pathogenicity. Genes *cj1415c-cj1418c* are required for MeOPN display on capsules, thus their enzyme products were predicted to catalyze the formation of MeOPN. The focus of this study is on the activities of Cj1418 and Cj1417 because there were hypothesized to initiate MeOPN biosynthesis with the production of phosphoramidate. Cj1418 was predicted to use adenosine triphosphate (ATP) and pyruvate to produce phosphoenolpyruvate, which would subsequently serve as a substrate for Cj1417 to synthesize phosphoramidate in a glutamine-dependent reaction. Cj1417 was thought to harness ammonia from glutamine, but data presented here do not support glutamine-dependent Cj1417 activity. Instead, Cj1418 ATP turnover was glutamine-dependent in the absence of Cj1417. The results described herein led to the identification of Cj1418 as a glutamine kinase and Cj1417 as a γ -glutamyl-cytidine diphosphoramidate hydrolase, details of which can be found in Appendices 1 and 2.

5.2 Introduction

The Gram-negative bacteria *Campylobacter jejuni* is one of the most common causes of gastroenteritis worldwide with contaminated poultry being a common source (1). Beyond the burden of acute infection, approximately 1 in 1000 individuals who experience *C. jejuni* enteritis develop the autoimmune, polyneuropathy Guillain-Barré syndrome (GBS) or the sub-type Miler Fisher syndrome (MFS) (2, 3). Sialic acids on the surface lipooligosaccharide (LOS) of *C. jejuni* mimic host gangliosides on peripheral nerves and are responsible for disease onset (4).

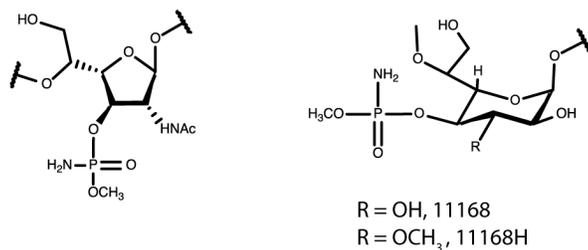
Besides making a LOS, *C. jejuni* produces a capsular polysaccharide (CPS), which is a major Penner serotyping scheme determinant and a well-established virulence marker (5-7). Non-encapsulated strains are less pathogenic in a ferret disease model and show reduced colonization in chicken, mice, and piglets (8-10). The CPS is required for adherence and invasion of human epithelial cells, resistance to complement-mediated killing, and cytokine modulation (10-13).

C. jejuni capsules harbor unusual substituents including an *O*-methyl phosphoramidate (MeOPN), an unprecedented chemical group in nature (14, 15). Remarkably, 75% of capsules are appended with MeOPN and the modification was prevalent in clinical isolates from GBS, MFS, and enteritis patients, suggesting it plays an important role in pathogenesis (15). Indeed, mutant strains that produce capsule without MeOPN show decreased resistance to serum-mediated killing but an increased ability to invade cultured epithelial cells whereas capsule mutants demonstrate reduced invasiveness (8, 10, 11, 16). MeOPN also plays a role in modulating host inflammation (8, 16). Furthermore, its presence is regulated by phase variation due to slipped strand mispairing in genes responsible for its transfer (5, 17, 18). The ability to switch MeOPN expression on and off, coupled with its immunomodulatory effects, necessitates a functional understanding of the enzymes involved in MeOPN biosynthesis as they may be suitable antibiotic targets (5, 19).

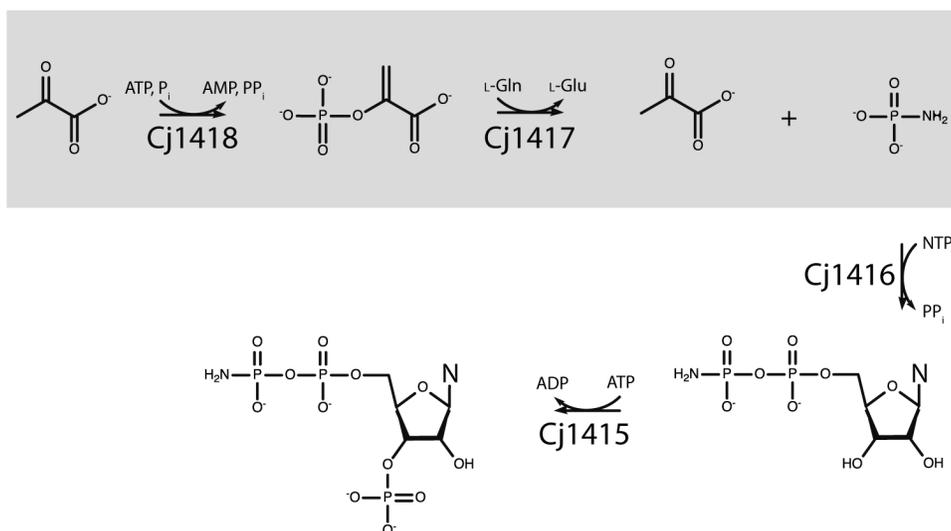
C. jejuni strain 11168 capsules contain MeOPN modified *N*-acetylgalactosamine and heptose sugars [Scheme 5.1(a)] (14). Variant strains no longer able to synthesize MeOPN led to the identification of four genes, *cj1415c-cj1418c*, whose enzyme products are implicated in MeOPN biosynthesis [Scheme 5.1 (b)] (15). Cj1418 is predicted to catalyze the formation of phosphoenolpyruvate (PEP) from pyruvate and adenosine triphosphate (ATP) due to its similarity to pyruvate, phosphate dikinases and PEP synthetases (20, 21). Cj1417 may synthesize phosphoramidate from PEP and glutamine as it is most similar to type 1 glutamine-dependent amidotransferases (22). Phosphoramidate is predicted to be activated with a nucleoside monophosphate via Cj1416, owing to its similarity to LicC, a CTP:phosphocholine cytidyltransferase (23). Finally, Cj1415 is annotated as an adenosine 5'-phosphosulfate kinase, and may therefore phosphorylate the 3' ribose hydroxyl of the Cj1416 product (24).

It was hypothesized that Cj1418 and Cj1417 initiate MeOPN biosynthesis by catalyzing the formation of phosphoramidate, so the functions of these two enzymes were the focus of this study. Data presented here suggest that Cj1418 does not use pyruvate, but instead converts ATP to adenosine monophosphate (AMP) in a glutamine-dependent manner. Cj1417 does not produce glutamate from the Cj1418 product, suggesting that it may not be a substrate for Cj1417. These data led to the eventual identification of the Cj1418 product as *L*-glutamine-phosphate and the annotation of Cj1417 as a γ -glutamyl-cytidine diphosphoramidate hydrolase (Appendices 1 and 2).

(a)



(b)



Scheme 5.1 Structures of MeOPN modified sugars and predicted enzyme activities required for MeOPN biosynthesis. The MeOPN modified 2-acetamido-2-deoxy- β -D-galactofuranose and D-glycero- α -L-gluco-heptopyranose sugars in the *C. jejuni* capsule are shown in (a). The proposed pathway for the biosynthesis of the phosphoramidate moiety in MeOPN is shown in (b), with the enzymes studied in this investigation highlighted in grey.

5.3 Materials and Methods

5.3.1 Cloning, expression, and purification

The genes encoding the enzymes Cj1417 and Cj1418 were cloned from *C. jejuni* 11168 genomic DNA by Cory Q. Wenzel (University of Alberta, Alberta, CA) in the laboratory of Dr. Christine M. Szymanski (University of Georgia, Athens, GA). *cj1417c* was cloned using NdeI and XhoI sites in pET28t, which affords a TEV protease cleavable N-terminal (His₆) tag (25). The N-terminal tag on this protein did not cleave well so one with a longer linker (two extra glycines) between the cleavage site and the start of the enzyme was generated via the QuikChange method of Stratagene. *cj1418c* was initially cloned into pET30a using NcoI and XhoI sites, which affords a 4.8 kDa N-terminal His₆-tag and S-tag. The plasmid was subsequently utilized as a template in polymerase chain reactions using the PrimeSTAR HS Polymerase system from Takara and primers that incorporate a N-terminal NheI restriction site and a C-terminal XhoI site. The insert was eventually integrated into the pET28t plasmid to yield a TEV protease cleavable N-terminal (His₆) tag construct and into the pET21b plasmid to produce a C-terminal (His₆) construct.

pET28t-*cj1417c* and *cj1418c* plasmids were used to transform Rosetta2(DE3) *Escherichia coli* cells (Novagen). The cultures were grown in lysogeny broth supplemented with kanamycin and chloramphenicol (50 mg/L and 25 mg/L concentrations, respectively) at 37 °C with shaking until an optical density of 0.9 was reached at 600 nm. The cells used for Cj1417 production were cooled in cold water, and the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and allowed to express protein at 21 °C for 20 h. The same cooling and induction was done with Cj1418-producing cells, except on ice, and the protein was expressed at 16 °C for 20 h.

The cells were harvested by centrifugation and frozen as pellets in liquid nitrogen. The Cj1417-producing cell pellets were subsequently disrupted by sonication on ice in a lysis buffer composed

of 50 mM sodium phosphate, 20 mM imidazole, 10% glycerol, and 300 mM sodium chloride, pH 8.0. The Cj1418-producing cell pellets were disrupted in the same buffer except that 500 mM sodium chloride and a Pierce Protease Inhibitor Tablet (no EDTA, Thermo-Fisher) were utilized because of the impure nature of the test expressions. The lysate was cleared by centrifugation, and both enzymes were purified at 4 °C utilizing Ni-nitrilotriacetic acid resin (Qiagen) according to the manufacturer's instructions. All buffers were adjusted to pH 8.0 and contained 50 mM sodium phosphate, 300 mM sodium chloride (500 mM for Cj1418), and imidazole concentrations of 25 mM for the wash buffer and 250 mM for the elution buffer. Following purification, the protein was pooled and mixed in a 30:1 molar ratio (His₆-protein:TEV protease) based on calculated extinction coefficients of 0.82 (mg/mL)⁻¹cm⁻¹ for Cj1417 and 1.21 (mg/mL)⁻¹cm⁻¹ for Cj1418. The recombinant proteins were allowed to digest at 4 °C for 48 h dialyzed against wash buffer containing 1 mM dithiothreitol (DTT). Uncleaved protein, the His₆ tag, and the TEV protease were removed via passage over Ni-nitrilotriacetic acid resin. Cleaved protein was collected and dialyzed against 10 mM Tris-HCl (pH 8.0) and 150 mM NaCl at 4 °C for 4 h and then dialyzed against 10 mM Tris-HCl (pH 8.0) at 4 °C overnight. Cj1417 was concentrated to 5-7 mg/mL, and Cj1418 was concentrated to 10-14 mg/mL based on the previously provided extinction coefficients and flash frozen in liquid nitrogen.

Cj1418 with either a N-terminal or a C-terminal His₆-tag construct was expressed and purified in a similar manner. Cells were transformed with the pET21b-*cj1418c* plasmid and grown in lysogeny broth supplemented with ampicillin and chloramphenicol (100 mg/L and 25 mg/L, respectively) or transformed with the aforementioned pET28t-*cj1418c* plasmid and grown in kanamycin and chloramphenicol (50 mg/L and 25 mg/L, respectively). Cells were induced, harvested, and frozen as previously described for the TEV-cleaved construct. To perform kinetic

assays with Cj1418 that was freshly purified (not frozen), N- and C-terminal tag constructs were purified as previously described using slightly modified buffers. Lysis was performed in 50 mM Tris-HCl, 20 mM imidazole, 10% glycerol, 500 mM sodium chloride, and 1 mM β -mercaptoethanol (β ME), pH 7.5. The lysate was cleared by centrifugation, and the enzymes were purified at 4 °C utilizing Ni-nitrilotriacetic acid resin. All buffers were adjusted to pH 7.5 and contained 50 mM Tris-HCl, 500 mM sodium chloride, 1 mM β ME and imidazole concentrations of 25 mM for the wash buffer and 250 mM for the elution buffer. Following purification, the protein was pooled and concentrated 15-fold and re-suspended in 15 mL of 10 mM Tris-HCl (pH 7.5) and 200 mM sodium chloride. Concentration and re-suspension were performed three times to buffer exchange, to reach a concentration of 5 mg/mL, and to subsequently assay the enzymes.

5.3.2 Enzymatic assays

Cj1418 PEP synthetase (PEPS) or pyruvate, phosphate dikinase (PPDK) activity was monitored in the forward (PEP producing) direction, and PEPS activity in the reverse direction was assayed by measuring pyruvate production. PPDK utilizes ATP, pyruvate, and phosphate to produce PEP, AMP, and pyrophosphate whereas PEPS uses ATP, pyruvate, and water to produce PEP, AMP, and phosphate (Scheme 5.2). Pyruvate production was monitored with a coupled system in which lactate dehydrogenase (LDH) reduces pyruvate to lactate and oxidizes reduced nicotinamide adenine dinucleotide (NADH) to NAD^+ with a concomitant decrease in absorbance at 340 nm. Reactions were performed at room temperature in 50 mM HEPES-KOH (pH 7.5) with 0.4 mM PEP, 1 mM AMP, 4 mM sodium phosphate, 10 mM MgCl_2 , 0.2 mM NADH, and 5 units of LDH. Reactions were initiated with the addition of Cj1418 to 1 mg/mL.

PEP production was monitored via a coupled assay using PEP carboxylase, (PEPC), which produces oxaloacetate from PEP and carbon dioxide, and malate dehydrogenase (MDH), which

reduces oxaloacetate to malate with concomitant oxidation of NADH. Reactions were performed at room temperature in 50 mM HEPES-KOH (pH 7.5) with 2 mM pyruvate, 10 mM sodium bicarbonate, 1 mM ATP, 0.2 mM NADH, 8 mM MgCl₂, 1 mM DTT, 1 mM glucose-6-phosphate (positive regulator of PEPC), 20 units of PEPC, 2 units of MDH, and ± 5 mM sodium phosphate. Reactions were initiated with the addition of Cj1418 to 1 mg/mL.

Glutamate production was monitored via a coupled assay using glutamate dehydrogenase (GDH) which oxidizes glutamate to α -ketoglutarate with concomitant reduction of NAD⁺. Reactions were performed in 50 mM HEPES-KOH (pH 7.5) with 1 mM glutamine, 0.2 mM NAD⁺, 0.4 mM adenosine diphosphate (ADP-positive GDH regulator), 1 mM DTT, and 10 units of GDH with or without 1 mM ATP, 2 mM pyruvate, 2 mM PEP, 1 mg/mL Cj1418, and 1 mg/mL Cj1417.

All reactions were monitored with a Beckman DU 640B spectrophotometer using an extinction coefficient for NADH at 340 nm of 6220 M⁻¹ cm⁻¹.

5.3.3 High-performance liquid chromatography analyses

ATP turnover was monitored via high-performance liquid chromatography (HPLC). Reactions containing 50 mM HEPPS (pH 8.0), 100 mM KCl, and various combinations of 5 mM ATP, 5 mM pyruvate, 5 mM PEP, 10 mM sodium phosphate, and 5 mM glutamine were set up with or without 1 mg/mL Cj1417 and/or 1 mg/mL Cj1418 and allowed to incubate for 1 hour at room temperature. Reactions were stopped by removing the enzyme(s) via filtration through a 10 kDa cutoff filter. The product mixtures were diluted 30x with water and examined by HPLC. Reaction samples were separated with a 10-column volume, 0.0 – 3.0 M gradient of ammonium acetate (pH 4.0) on a 1 ml Resource-Q column. Based on standards of each, ATP had a retention volume of 13.2 mL; ADP had a retention volume of 10.1 mL; AMP had a retention volume of 7.5 mL.

5.4 Results and Discussion

5.4.1 Cj1418 does not function as a PPDK or PEPS

Based upon its primary amino acid sequence, Cj1418 was predicted to function as a pyruvate, phosphate dikinase (PPDK) or a phosphoenolpyruvate (PEP) synthetase (PEPS), whose reactions are shown in Scheme 5.2 (b) (15). Cj1418, however, appears to exhibit a unique domain structure in which the putative phospho-histidine domain is situated at the C-terminus, and instead of a β/α PEP barrel, Cj1418 contains a central domain with no similarity to a characterized enzyme or domain [Scheme 5.2 (a)] (20). Additionally, Cj1418 likely possesses an N-terminal ATP-grasp domain, albeit smaller than that found in PPDK. ATP-grasp enzymes typically catalyze the reaction of a carboxylate group with a nucleophile using ATP and proceed via an alkyl-phosphate intermediate (26). PPDK and PEPS are therefore exceptions to this mechanism because they function as kinases.

Cj1418 was assayed as a PEPS and PPDK by monitoring PEP production in the forward direction using a PEP carboxylase/malate dehydrogenase coupled assay and by monitoring the reverse reaction, pyruvate production, using a lactate dehydrogenase coupled assay as these enzymes are known to catalyze reversible reactions (21, 27). Furthermore, PPDK is activated by the monovalent cations NH_4^+ or K^+ , so buffer titrated with KOH was used throughout (28, 29).

(a)

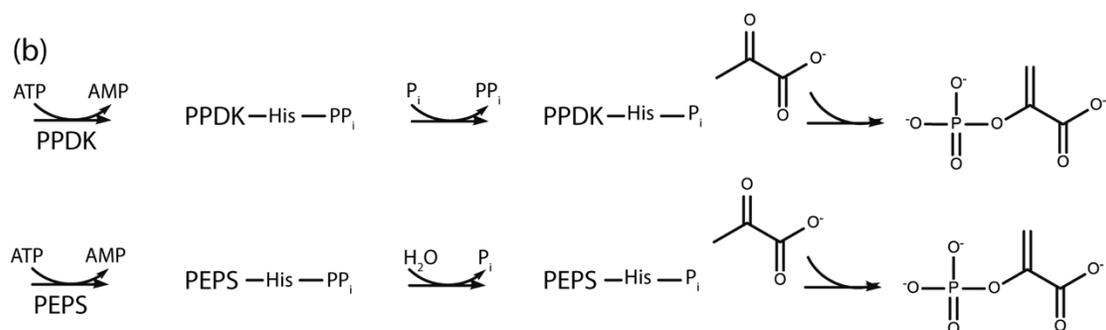
PPDK



Cj1418



(b)



Scheme 5.2 Possible Cj1418 activity based on similarity to PPDK and PEPS. Shown in (a) is a comparison of the domain structure of PPDK and predicted domain structure of Cj1418. Both harbor ATP-grasp domains, however, Cj1418 contains a central domain with no similarity to an enzyme with annotated function, and the putative phospho-His domain is predicted to be situated at the C-terminus of Cj1418. PPDK and PEPS reactions assayed in this study are shown in (b). The only difference between the two is that in PPDK, pyrophosphate is released whereas phosphate is released in the course of the PEPS reaction.

As shown in Table 5.1, the maximal rate achieved by Cj1418, either in the forward or reverse direction, was 1.8 $\mu\text{M}/\text{min}$. Furthermore, enzymes with a N- or C- terminal tag were purified and assayed on the same day to eliminate the possibility that freezing was leading to inactivity. Reactions were performed using 0.4 mM PEP or 2 mM pyruvate, well above the reported K_M values for authentic PEPS and PPDK. The K_M for pyruvate in PEPS is ~ 0.02 mM and for PPDK is ~ 0.08 mM (21, 29). In the reverse direction, the K_M for PEP in PEPS is ~ 1 mM and for PPDK is between 0.04 and 0.1 mM (21, 27). The high concentrations of potential substrates and products tested against Cj1418 yielded a maximum rate of 1.8 $\mu\text{M}/\text{min}$ or 1.8 nmol/min per mg of protein whereas authentic PEPS and PPDK function on the order of 3 to 5 $\mu\text{mol}/\text{min}$ per mg of protein (21, 29), suggesting that Cj1418 is likely not a PEPS or PPDK.

As further evidence, ATP turnover (instead of PEP production) in the presence and absence of pyruvate or phosphate was monitored via HPLC. Based upon the mechanisms of PPDK and PEPS, AMP was the expected nucleotide product (21, 27). However, in comparison to the no enzyme control [Fig. 5.1 (a)] after one hour, Cj1418 did not turn over ATP, regardless of whether or not pyruvate or phosphate was present [Fig. 5.1 (b) and (c)]. Both the coupled assays and HPLC analyses suggest that Cj1418 is not a PEPS or a PPDK.

Table 5.1 Cj1417 and Cj1418 Rates Determined via Spectrophotometric Analysis.

Enzyme(s) Tested^a	Substrates Present	Product Assayed	Rate ($\mu\text{M}/\text{min}$)
Cj1418	PEP, PO_4^{3-} , AMP	pyruvate	0.2
Cj1418	pyruvate, ATP	PEP	0
Cj1418	pyruvate, PO_4^{3-} , ATP	PEP	1.0
Cj1418 – NHT ^b	pyruvate, ATP	PEP	0.8
Cj1418 – CHT ^b	pyruvate, ATP	PEP	1.8
Cj1417	glutamine	glutamate	0.1
Cj1417	glutamine, PEP	glutamate	0.4
Cj1417/Cj1418	glutamine, ATP	glutamate	0.5
Cj1417/Cj1418	glutamine, pyruvate, ATP	glutamate	0.4

^aAll enzymes tested were TEV-cleaved unless indicated.

^bEnzymes were purified in Tris-HCl buffer and assayed the same day they were purified.

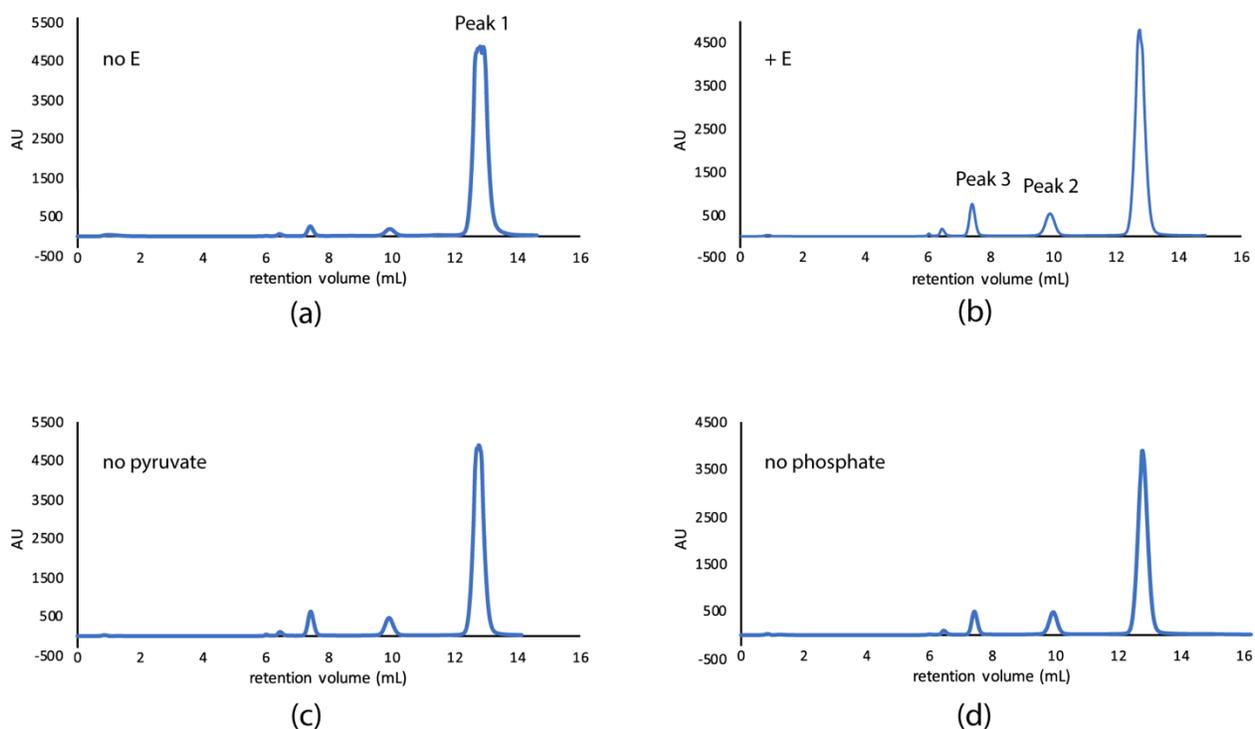


Figure 5.1 Cj1418 ATP turnover after 1 hour in the presence and absence of pyruvate or phosphate. (a) In the no enzyme control, ATP was not hydrolyzed. In the presence of Cj1418 and all potential substrates (b), no ATP turnover occurred, thus the absence of pyruvate (c) or added phosphate (d) did not affect ATP turnover. Peak 1 corresponds to ATP; peak 2 corresponds to ADP; peak 3 corresponds to AMP.

5.4.2 Cj1417 does not hydrolyze glutamine

On the basis of amino acid sequence homology, Cj1417 is annotated as a type 1 glutamine-dependent amidotransferase (GAT) and was predicted to catalyze the formation of phosphoramidate (15). Type 1 GATs use a catalytic histidine-cysteine dyad to hydrolyze glutamine, or a glutamine-containing analogue, releasing ammonia from the “glutaminase” domain in the former case (22). The ammonia is transferred to a “synthetase” domain to react with another substrate (30). While Cj1418 function was being interrogated for PEP production, Cj1417’s ability to utilize PEP and glutamine to produce phosphoramidate was assayed using a glutamate dehydrogenase coupling reaction to monitor the formation of glutamate. No appreciable glutamate production was observed when Cj1417 was incubated with 1 mM glutamine in the presence or absence of 2 mM PEP (Table 5.1). 1 mM glutamine is well above the K_M of most type 1 GATs – for example, the glutaminase subunit (small subunit) of carbamoyl phosphate synthetase displays a Michaelis constant of ~ 0.1 mM (31) and a maximal velocity with respect to glutamine of 1.1 $\mu\text{mol}/\text{hour}$ per mg of protein (32). The highest rate of 0.02 $\mu\text{mol}/\text{hour}$ per mg of Cj1417 suggests that Cj1417 does not catalyze the reaction depicted in Scheme 5.1 (b). The lack of activity with PEP was not surprising once it was determined that Cj1418 is not a PPDK or a PEPS, however, the lack of Cj1417 glutaminase activity was unexpected.

5.4.3 Cj1417 and Cj1418 may function in tandem

In type 1 GATs with glutaminase and synthetase activities on separate protein chains, the glutaminase usually contains between 200 and 300 amino acids (22). Cj1417 is 200 amino acids long, so it was hypothesized to function as a glutaminase alongside Cj1418 as the synthetase in a manner analogous to carbamoyl phosphate synthetase. The small subunit of carbamoyl phosphate synthetase hydrolyzes glutamine and the large subunit catalyzes the ATP-dependent formation of product (33). Synthetases usually do not hydrolyze ATP in the absence of their glutaminase, so the lack of demonstrable Cj1418 PEPS or PPDK activity may have been because Cj1417 was not present (30). Thus, Cj1417 and Cj1418 were incubated together with pyruvate, glutamine, phosphate, and ATP to assess ATP turnover after one hour via HPLC. In the presence of all enzymes and substrates, ATP was converted to AMP within one hour [Fig. 5.2 (a)]. The reaction was glutamine-dependent [Fig. 5.2 (b)], as expected, but not pyruvate or phosphate dependent as might be expected if Cj1418 performs a PEPS or PPDK like reaction [Fig. 5.2 (c) and (d)].

To parse out the function of the individual enzymes, reactions were performed in the absence of Cj1417 or Cj1418 (Fig. 5.3). Part (a) is repeated in Figures 5.2 and 5.3 because the reactions for both were set up and analyzed on the same day. As expected if Cj1418 were to function as a synthetase, no AMP product was observed in reactions lacking Cj1418 [Fig. 5.3 (b)]. Strikingly, ATP was just as swiftly hydrolyzed in the absence of Cj1417 [Fig. 5.3 (c)]. Furthermore, the only substrate required for Cj1418 ATP turnover was glutamine, as reactions lacking Cj1417 and glutamine did not proceed [Fig. 5.3 (d)], but those containing glutamine but lacking Cj1417 and pyruvate or added phosphate still afforded AMP [Fig. 5.3 (e) and (f)].

Before it became clear that Cj1418 was the enzyme performing a glutamine-dependent reaction, the possibility that Cj1417 glutaminase activity was stimulated by its interaction with Cj1418 in the presence of ATP and pyruvate was monitored using the GDH coupled assay described in the previous section. However, incubating Cj1417, Cj1418, glutamine, and ATP in the presence or absence of pyruvate did not lead to significant glutamate production (Table 5.1).

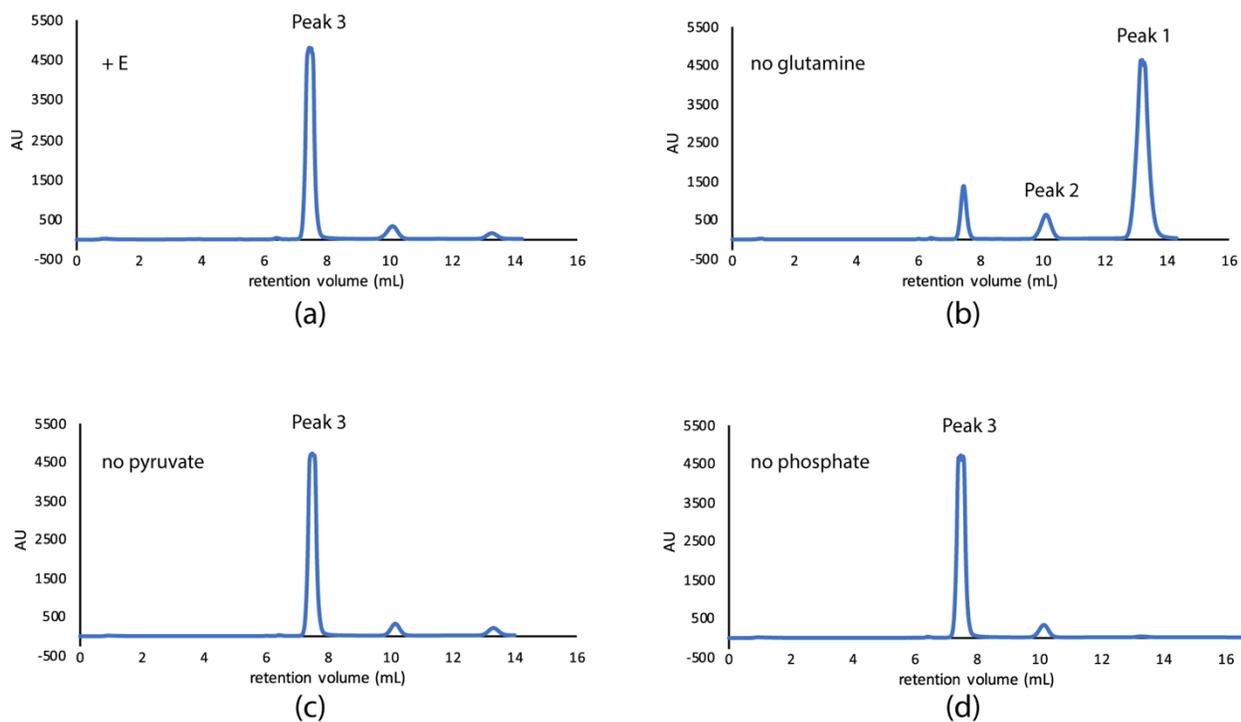


Figure 5.2 ATP turnover after 1 hour in the presence of Cj1417 and Cj1418. In panel (a), both the enzymes and all potential substrates were present, and all of the ATP was converted to AMP. The reaction course without added glutamine (b) resulted in very little ATP to AMP turnover, whereas reactions that lacked pyruvate (c) or added phosphate (d) still demonstrated an ATP to AMP conversion. Peak 1 corresponds to ATP; peak 2 corresponds to ADP; peak 3 corresponds to AMP.

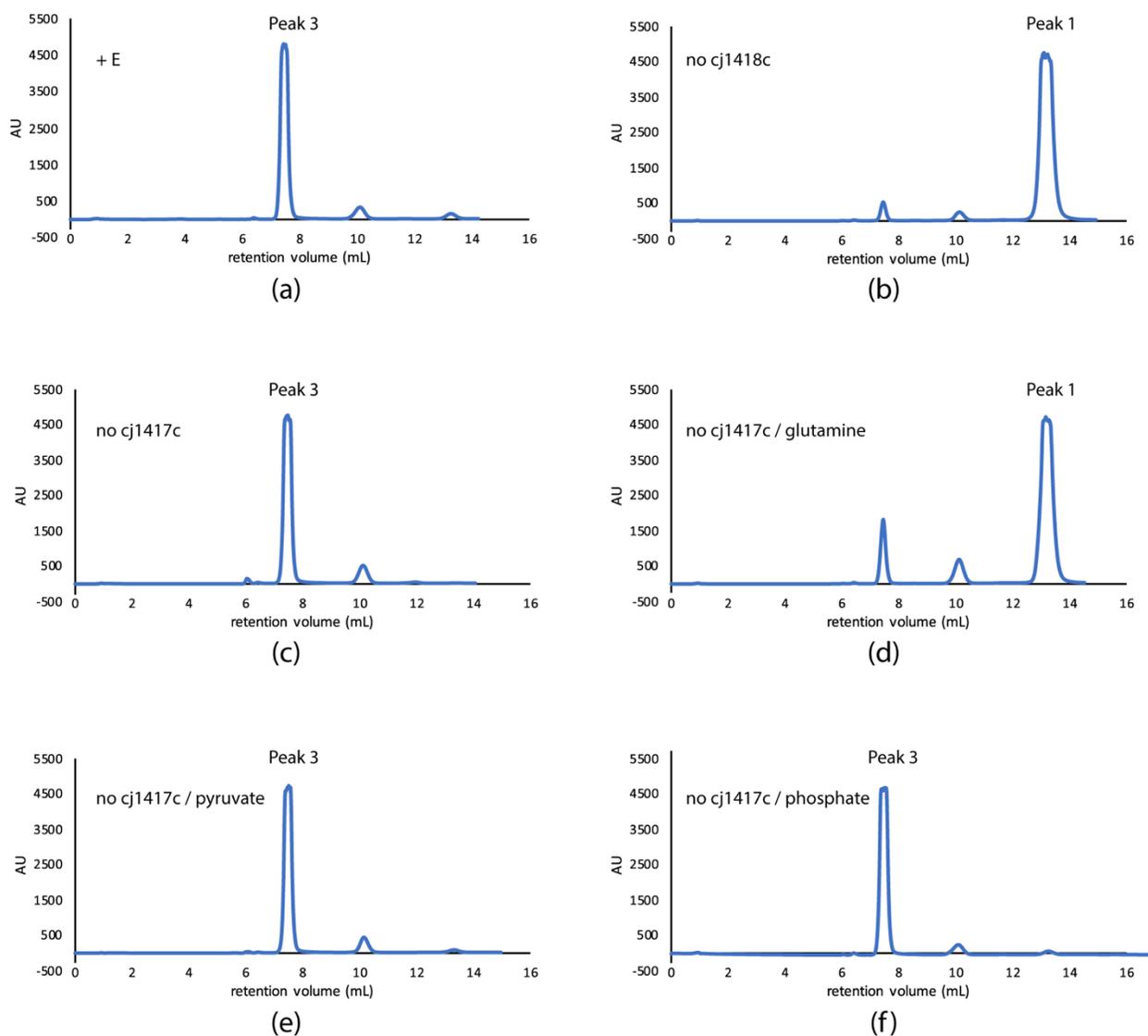


Figure 5.3 ATP turnover after 1 hour in the presence or absence of Cj1417 or Cj1418. (a) is the same as in Figure 5.2 (a) (all enzymes and potential substrates present). Without Cj1418 (b), no AMP was produced. Without Cj1417 (c), ATP was still hydrolyzed to AMP. In the absence of Cj1417 and glutamine, ATP turnover did not occur. Conversely, in the absence of Cj1417 and pyruvate (e) or added phosphate (f), full conversion to AMP occurred. Peak 1 corresponds to ATP; peak 3 corresponds to AMP.

5.4.4 Conclusions

Based on the data described here, Cj1418 turns over ATP in the presence of glutamine without Cj1417, pyruvate, or phosphate. The HPLC results suggesting that Cj1418 turns over ATP in a glutamine-dependent manner in the absence of Cj1417 led to speculation that there might be a contaminating type 1 GAT in the Cj1418 purification. However, the glutaminase assay results suggest otherwise as there would have been robust glutamate production, which was not observed (Table 5.1). Similarly, if Cj1418 is an ATP-dependent glutaminase, glutamate should have been detected, but was not. Albeit this would not have been distinguishable from the activity of a contaminating amidotransferase, the results nonetheless suggest that, together, Cj1418 and Cj1417 do not synthesize glutamate but that Cj1418 hydrolyzes ATP in a glutamine-dependent reaction. Thus, the identity of the Cj1418 product and for which enzyme it serves as a substrate, be it Cj1417 or another enzyme in the MeOPN operon, required further characterization.

The data presented herein led to the eventual identification of the Cj1418 product. Appendix 1 describes the characterization of the Cj1418 reaction, namely that it phosphorylates the side chain of L-glutamine to produce L-glutamine-phosphate. This is fully supported by the glutamine-dependent ATP turnover described in section 5.4.3. Furthermore, the data suggesting that Cj1417 does not make glutamate led to the characterization of the fate of the Cj1418 product, chronicled in Appendix 2. Indeed, the Cj1418 product is not the Cj1417 substrate, but rather the substrate for Cj1416 which appends L-glutamine-phosphate to cytidine monophosphate (CMP) to create CDP-L-glutamine. Cj1417 subsequently hydrolyzes CDP-L-glutamine to afford CDP-amidate (CDP-NH₂) and glutamate. Thus, Cj1417 does produce glutamate, but its substrate is CDP-L-glutamine, not the Cj1418 product or glutamine.

5.5 References

1. Young, K. T., Davis, L. M., and DiRita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* *5*, 665-679.
2. Prendergast, M. M. and Moran, A. P. (2000) Lipopolysaccharides in the development of the Guillain-Barré syndrome and Miller Fisher syndrome forms of acute inflammatory peripheral neuropathies. *J. Endotoxin. Res.* *6*, 341-359.
3. Jacobs, B. C., Rothbarth, P. H., van der Meché, F. G. A., Herbrink, P., Schmitz, P. I. M., de Klerk, M. A., and van Doorn, P. A. (1998) The spectrum of antecedent infections in Guillain-Barré syndrome: a case-control study. *Neurology.* *51*, 1110-1115.
4. Ang, C. W., Laman, J. D., Willison, H. J., Wagner, E. R., Endtz, H. P., de Klerk, M. A., Tio Gillen, A. P., van den Braak, N., Jacobs, B. C., and van Doorn, P. A. (2002) Structure of *Campylobacter jejuni* lipopolysaccharides determines antiganglioside specificity and clinical features of Guillain-Barré and Miller Fisher patients. *Infect. Immun.* *70*, 1202-1208.
5. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., and Monteiro, M. A. (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines *Front. Cell Infect. Microbiol.* *2*, doi: 10.3389.
6. Bachtiar, B. M., Coloe, P. J., and Fry, B. N. (2007) Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 8116 and its involvement in bacterium-host interactions. *FEMS Immunol. Med. Microbiol.* *49*, 149-154.
7. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J., and Wren, B. W. (2000) Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol. Microbiol.* *35*, 529-541.
8. Maue, A. C., Mohawk, K. L., Giles, D. K., Poly, F., Ewing, C. P., Jiao, Y., Lee, G., Ma, Z., Monteiro, M. A., Hill, C. L., Ferdberber, J. S., Porter, C. K., Trent, M. S., and Guerry, P. (2013) The polysaccharide capsule of *Campylobacter jejuni* modulates the host immune response. *Infect. Immun.* *81*, 665-672.
9. Bacon, D. J., Szymanski, C. M., Burr, D. H., Silver, R. P., Alm, R. A., and Guerry, P. (2001) A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol. Microbiol.* *40*, 769-777.
10. van Alphen, L. B., Wenzel, C. Q., Richards, M. R., Fodor, C., Ashmus, R. A., Stahl, M., Karlyshev, A. V., Wren, B. W., Stintzi, A., Miller, W. G., Lowary, T. L., and Szymanski, C. M. (2014) Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One.* *9*, e87051.
11. Wong, A., Lange, D., Houle, S., Arbatsky, N. P., Valvano, M. A., Knirel, Y. A., Dozois, C. M., and Creuzenet, C. (2015) Role of capsular modified heptose in the virulence of *Campylobacter jejuni*. *Mol. Microbiol.* *96*, 1136-1158.

12. Rose, A., Kay, E., Wren, B. W., and Dallman, M. J. (2012) The *Campylobacter jejuni* NCTC 11168 capsule prevents excessive cytokine production by dendritic cells. *Med. Microbiol. Immunol.* *201*, 137-144.
13. Stahl, M., Ries, J., Vermeulen, J., Yang, H., Sham, H. P., Crowley, S. M., Badayeva, Y., Turvey, S. E., Gaynor, E. C., Li, X., and Vallance, B. A. (2014) A novel mouse model of *Campylobacter jejuni* gastroenteritis reveals key pro-inflammatory and tissue protective roles for toll-like receptor signalling during infection. *PLoS Pathog.* *10*, e1004264.
14. Szymanski, C. M., St. Michael, F., Jarrell, H. C., Li, J., Gilbert, M., Larocque, S., Vinogradov, E., and Brisson, J. R. (2003) Detection of conserved *N*-linked glycans and phase-variable lipooligosaccharides and capsules from campylobacter cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. *J. Biol. Chem.* *278*, 24509-24520.
15. McNally, D. J., Lamoureux, M. P., Karlyshev, A. V., Fiori, L. M., Li, J., Thacker, G., Coleman, R. A., Khieu, N. H., Wren, B. W., Brisson, J. R., Jarrell, H. C., and Szymanski, C. M. (2007) Commonality and biosynthesis of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* *282*, 28566-28576.
16. Kim, S., Vela, A., Clohisey, S. M., Athanasiadou, S., Kaiser, P., Stevens, M. P., and Vervelde, L. (2018) Host-specific differences in the response of cultured macrophages to *Campylobacter jejuni* capsule and *O*-methyl phosphoramidate mutants. *Vet. Res.* *49*, s13567-017-0501-y.
17. Holst Sørensen, M. C., van Alphen, L. B., Fodor, C., Crowley, S. M., Christensen, B. B., Szymanski, C. M., and Brønsted, L. (2012) Phase variable expression of capsular polysaccharide modifications allows *Campylobacter jejuni* to avoid bacteriophage infection in chickens. *Front. Cell Infect. Microbiol.* *2*, Article 11.
18. Karlyshev, A. V., Champion, O. L., Churcher, C., Brisson, J. R., Jarrell, H. C., Gilbert, M., Brochu, D., St. Michael, F., Li, J., Wakarchuk, W. W., Goodhead, I., Sanders, M., Stevens, K., White, B., Parkhill, J., Wren, B. W., and Szymanski, C. M. (2005) Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* *55*, 90-103.
19. Guerry, P. and Szymanski, C. M. (2008) *Campylobacter* sugars sticking out. *Trends Microbiol.* *16*, 428-435.
20. Herzberg, O., Chen, C. C. H., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano, D. (1996) Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci. USA.* *93*, 2652-2657.
21. Berman, K. M. and Cohn, M. (1970) Phosphoenolpyruvate synthetase of *Escherichia coli*. *J. Biol. Chem.* *245*, 5309-5318.
22. Massière, F. and Badet-Denisot, M. A. (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol. Life Sci.* *54*, 205-222.

23. Kwak, B. Y., Zhang, Y. M., Yun, M., Heath, R. J., Rock, C. O., Jackowski, S., and Park, H. W. (2002) Structure and mechanism of CTP:phosphocholine cytidylyltransferase (LicC) from *Streptococcus pneumoniae*. *J. Biol. Chem.* *277*, 4343-4350.
24. Leyh, T. S., Taylor, J. C., and Markham, G. D. (1988) The sulfate activation locus of *Escherichia coli* K12: cloning, genetic, and enzymatic characterization. *J. Biol. Chem.* *263*, 2409-2416.
25. Thoden, J. B. and Holden, H. M. (2005) The molecular architecture of human *N*-acetylgalactosamine kinase. *J. Biol. Chem.* *280*, 32784-32791.
26. Fawaz, M. V., Topper, M. E., and Firestone, S. M. (2011) The ATP-grasp enzymes. *Bioorg. Chem.* *39*, 185-191.
27. Wang, H. C., Ciskanik, L., Dunaway-Mariano, D., von der Saal, W., and Villafranca, J. J. (1988) Investigations of the partial reactions catalyzed by pyruvate phosphate dikinase. *Biochemistry.* *27*, 625-633.
28. Reeves, R. E., Menzies, R. A., and Hsu, D. S. (1968) The pyruvate-phosphate dikinase reaction. *J. Biol. Chem.* *243*, 5486-5491.
29. Reeves, R. E. (1971) Pyruvate, phosphate dikinase from *Bacteroides symbiosus*. *Biochem. J.* *125*, 531-539.
30. Raushel, F. M., Thoden, J. B., and Holden, H. M. (2003) Enzymes with molecular tunnels. *Acc. Chem. Res.* *36*, 539-548.
31. Raushel, F. M., Anderson, P. M., and Villafranca, J. J. (1978) Kinetic mechanism of *Escherichia coli* carbamoyl-phosphate synthetase. *Biochemistry.* *17*, 5587-5591.
32. Miran, S. G., Chang, S. H., and Raushel, F. M. (1991) Role of the four conserved histidine residues in the amidotransferase domain of carbamoyl phosphate synthetase. *Biochemistry.* *30*, 7901-7907.
33. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry.* *36*, 6305-6316.

Chapter 6:

A Preliminary Structural Analysis of Cytidine Diphosphoramidate Kinase from

Campylobacter jejuni

6.1 Abstract

Campylobacter jejuni is one of the most common causes of food-borne gastroenteritis and is also the most prevalent infection preceding the onset of Guillain-Barré syndrome (GBS), an acute polyneuropathy caused by *C. jejuni* mimicry of host gangliosides. Strikingly, bacterial isolates from GBS patients frequently contain a unique *O*-methyl phosphoramidate (MeOPN) moiety on their capsular polysaccharides. Although the capsule is a virulence marker in the organism, MeOPN presence is beneficial during initial stages of infection for immune evasion. Four enzymes essential for MeOPN production have recently been characterized, however there are no structural data available for these novel enzymes that could be targeted to control *C. jejuni* infections. The focus of this structural investigation is centered around the enzyme responsible for the fourth step, Cj1415. Cj1415 catalyzes a reaction similar to the adenosine 5'-phosphosulfate kinases (APSK) and phosphorylates the 3' ribose hydroxyl of cytidine diphosphoramidate (CDP-NH₂). The data described here are of an initial Cj1415 structure refined to 2.6 Å resolution. Although there was electron density calculated with ($F_o - F_c$) coefficients corresponding to the predicted CDP-NH₂ binding site, the overall quality and resolution of the data did not allow reliable modeling of CDP-NH₂ or assignment of active site amino acids. Alignments with APSK did, however, reveal residues that are predicted to discriminate CDP-NH₂ from adenosine 5'-phosphosulfate. Higher resolution X-ray data will be necessary to more fully address the interactions required for ligand binding. Nonetheless, this is the first three-dimensional model of an enzyme necessary for MeOPN biosynthesis and will promote further structural analyses of Cj1415.

6.2 Introduction

The Gram-negative bacterium *Campylobacter jejuni* is a leading cause of food-borne gastroenteritis worldwide and is the most common infection preceding the onset of the autoimmune disorder Guillain-Barré Syndrome (GBS) (1, 2). *C. jejuni* devotes an estimated 10% of its genome to carbohydrate production, encoding systems for the synthesis of a lipooligosaccharide (LOS), *O*- and *N*-linked glycans, and a capsular polysaccharide (CPS) (3, 4). *C. jejuni* is unusual as its LOS/CPS outer membrane more closely resembles those found in Gram-negative mucosal pathogens while enteric pathogens usually produce lipopolysaccharide (5).

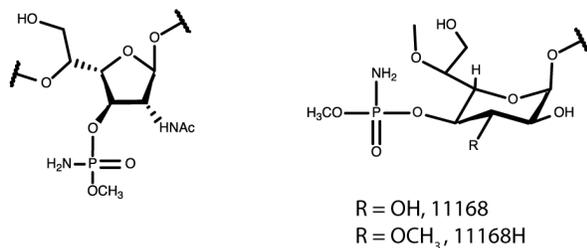
Still, the *C. jejuni* capsule is necessary for adherence, invasion, and modulation of the host immune response (6-9). Its structure is unique for the incorporation of heptoses in unusual configurations as well as its modification with an unusual *O*-methyl phosphoramidate (MeOPN) moiety [Scheme 6.1 (a)] (4, 10, 11). MeOPN presence on the capsule increases resistance to serum-mediated killing, decreases the ability of *C. jejuni* to invade cultured epithelial cells, and modulates host inflammation (8, 9, 12, 13). Targeting MeOPN biosynthetic enzymes may therefore serve as a way to alter the pathogenicity of the organism (4, 10).

Four enzymes essential for MeOPN biosynthesis have been characterized biochemically thus far. As shown in Scheme 6.1 (b), Cj1418 catalyzes the formation of the phosphorous-nitrogen bond via phosphorylation of L-glutamine (14). L-glutamine-phosphate is appended to cytidine monophosphate (CMP) by Cj1416 to form CDP-L-glutamine which is subsequently hydrolyzed by Cj1417 to make CDP-amidate (CDP-NH₂) (15). Cj1415 phosphorylates the 3' ribose hydroxyl of CDP-NH₂ using adenosine triphosphate (ATP) (16). Cj1415 is therefore a cytidine diphosphoramidate kinase, but for the sake of brevity will be referred to here as Cj1415.

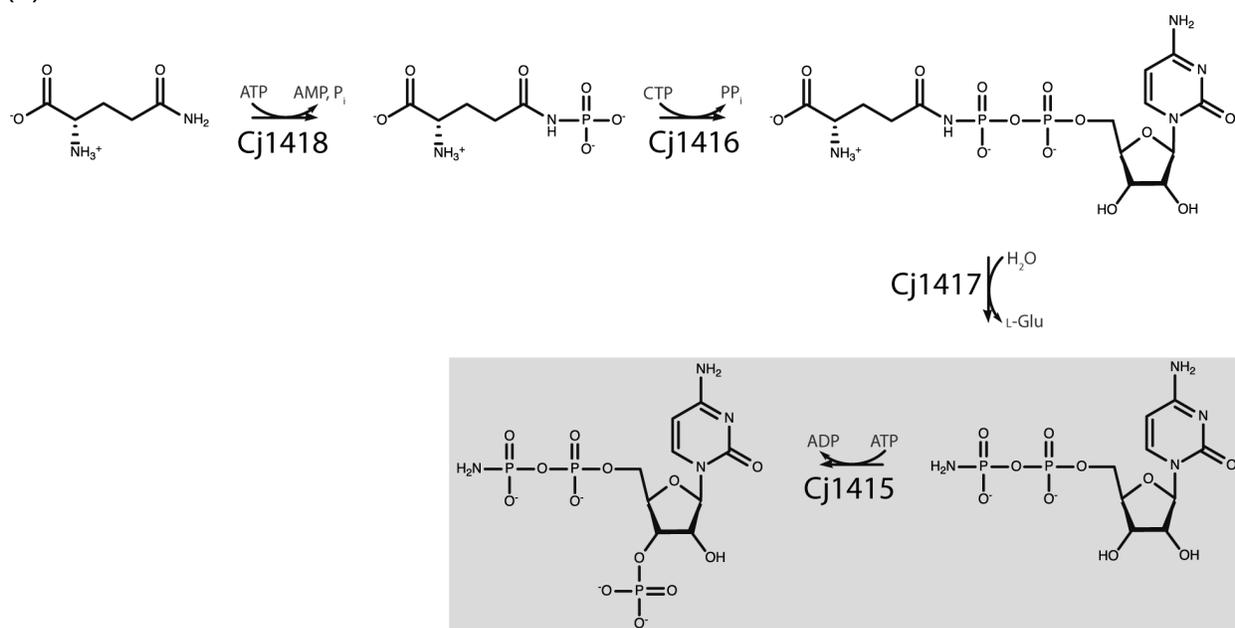
Cj1415 was initially annotated as an adenosine 5'-phosphosulfate (APS) kinase (APSK). A BLAST search against the Protein Data Bank indicates that Cj1415 is most similar to this enzyme from the cyanobacterium *Synechocystis* sp. PCC 6803, demonstrating a sequence identity and similarity of 35% and 52%, respectively (17). The question then arises as to what constitutes the structural determinants of CDP-NH₂ versus APS recognition in Cj1415 (Scheme 6.2). Indeed, Cj1415 does not phosphorylate APS, and its catalytic efficiency is significantly reduced with other nucleoside diphosphoramidates. Likewise, substitution of the substrate phosphoramidate moiety with phosphate, *O*-methyl phosphate, or methyl phosphate abrogates activity approximately 10 to 130-fold (16).

Here, the first structural model of an enzyme required for MeOPN biosynthesis is reported, namely that of Cj1415. Like the APSKs, Cj1415 functions as a dimer, displays an α/β purine nucleotide binding fold, and contains a predicted phosphate-binding loop (P-loop) to accommodate ATP. Electron density corresponding to the likely CDP-NH₂ binding site was present; however, because of the insufficient quality of the data, accurate arrangement of the ligand in this density was not possible. Nonetheless, a comparison of the Cj1415 model with the APSK from *Arabidopsis thaliana* solved in the presence of Mg²⁺, APS, and an ATP analogue led to the identification of the predicted CDP-NH₂ binding site in Cj1415. The model described herein will be invaluable for a subsequent high resolution structural analysis.

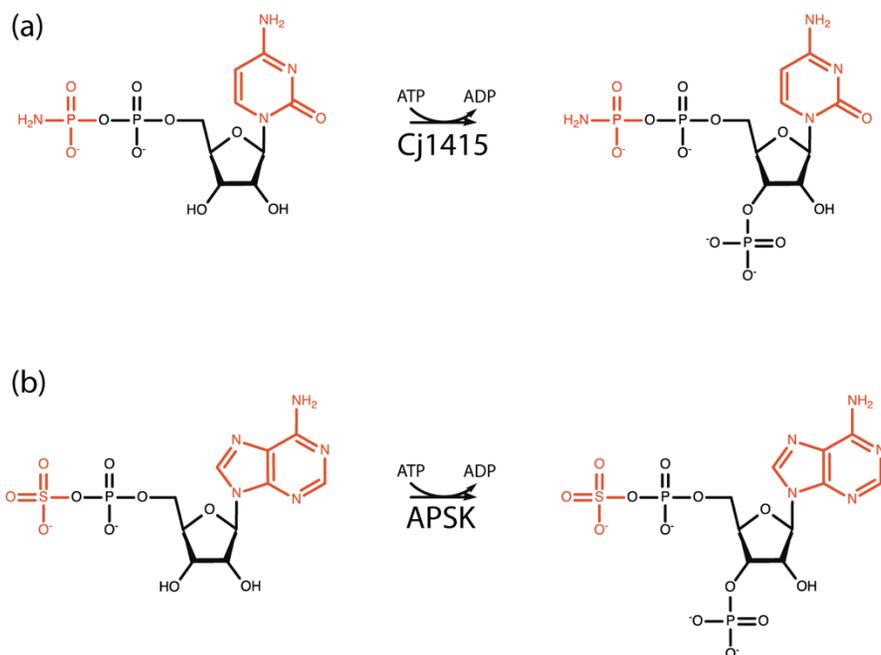
(a)



(b)



Scheme 6.1 The structures of MeOPN modified sugars and the pathway for phosphoramidate production in *C. jejuni*. The MeOPN modified 2-acetamido-2-deoxy- β -D-galactofuranose and D-glycero- α -L-gluco-heptopyranose sugars in the *C. jejuni* capsule are shown in (a). Part (b) details the enzymes necessary to make the phosphoramidate component of MeOPN. The highlighted enzyme, Cj1415, is the focus of this structural investigation.



Scheme 6.2 Comparison of Cj1415 and APSK activities. The differences between the substrates are highlighted in red. Whereas Cj1415 utilizes a cytosine-based substrate with a 5'-diphosphoramidate moiety (a), APSKs act on an adenosine-based substrate with a 5'-phosphosulfate group (b). Both enzymes, however, phosphorylate the 3'-carbons of their respective nucleoside substrates.

6.3 Materials and Methods

6.3.1 Protein expression and purification

The gene encoding Cj1415 from *C. jejuni* 11168 was cloned and donated by Cory Q. Wenzel (University of Alberta, Alberta, CA) from Dr. Christine Szymanski's laboratory (University of Georgia, Athens, GA). Briefly, the gene was cloned into the NdeI and XhoI sites of the pET28t vector, a laboratory pET28b vector that was previously modified to incorporate a TEV protease cleavage recognition sequence after the N-terminal His₆-tag (18). The pE28t-*cj1415c* plasmid was utilized to transform Rosetta2(DE3) *Escherichia coli* cells (Novagen). The cultures were grown in lysogeny broth supplemented with kanamycin and chloramphenicol (50 mg/L and 25 mg/L concentrations, respectively) at 37 °C with shaking until an optical density of 0.9 was reached at 600 nm. The flasks were cooled in ice water, and the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside and allowed to express protein at 16 °C for 20 h.

The cells were harvested by centrifugation and frozen as pellets in liquid nitrogen. The pellets were subsequently disrupted by sonication on ice in a lysis buffer composed of 50 mM sodium phosphate, 20 mM imidazole, 10% glycerol, and 300 mM sodium chloride, pH 8.0. The lysate was cleared by centrifugation, and Cj1415 with a N-terminal His₆-tag was purified at 4 °C utilizing Ni-nitrilotriacetic acid resin (Qiagen) according to the manufacturer's instructions. All buffers were adjusted to pH 8.0 and contained 50 mM sodium phosphate, 300 mM sodium chloride, and imidazole concentrations of 25 mM for the wash buffer and 250 mM for the elution buffer. The protein was pooled and mixed in a 30:1 molar ratio (Cj1415 protein:TEV protease) based on a calculated extinction coefficient for Cj1415 of 0.85 (mg/mL)⁻¹cm⁻¹. The recombinant protein was allowed to digest at 4 °C for 48 h. Uncleaved protein, His₆-tag, and the TEV protease were removed via passage over Ni-nitrilotriacetic acid resin. Cleaved protein was collected and dialyzed against

10 mM Tris-HCl (pH 8.0) and 150 mM NaCl at 4 °C for 4 h and subsequently transferred to dialysis buffer containing 10 mM Tris-HCl (pH 8.0) at 4 °C overnight and concentrated to 14 - 16 mg/mL based on the extinction coefficient provided above. Selenomethione-labeled protein was expressed using standard methods (19) and purified in the same manner as that described for the wild-type enzyme.

6.3.2 Protein crystallization and X-ray structural analyses

Crystallization conditions were surveyed by the hanging drop method of vapor diffusion using a sparse matrix screen developed in the Holden laboratory. Suitably diffracting crystals of wild-type and selenomethionine [SeMet]-labeled protein were set up at 4 °C with 10 mM MgCl₂, 5 mM adenosine 5'-(β,γ-imido) triphosphate (AMP-PNP) (Sigma-Aldrich), and 5 mM cytidine diphosphoramidate (CDP-NH₂). The CDP-NH₂ was synthesized and provided by Zane W. Taylor from the laboratory of Dr. Frank M. Raushel (Texas A&M University, College Station, TX) and made according to (16). X-ray diffraction quality crystals reached a maximal size of approximately 0.4 mm x 0.5 mm x 0.3 mm after two weeks. Specifically, wild-type crystals were grown by mixing 1:1 the protein sample at 16 mg/mL containing 10 mM MgCl₂, 5 mM AMP-PNP, and 5 mM CDP-NH₂ with a precipitant solution composed of 15 - 17% poly(ethylene glycol) 8000, 100 mM MOPS (pH 7.0), and 200 mM LiCl. The SeMet labeled crystals were grown by mixing 1:1 the protein sample at 16 mg/mL containing 10 mM MgCl₂, 5 mM AMP-PNP, and 5 mM CDP-NH₂ with a precipitant solution composed of 17 - 22% poly(ethylene glycol) 8000, 100 mM MOPS (pH 7.0), and 200 mM KCl. The crystals belonged to the space group *P1* with unit cell dimensions of $a = 46.7 \text{ \AA}$, $b = 97.3 \text{ \AA}$, $c = 110.7 \text{ \AA}$, $\alpha = 81.7^\circ$, $\beta = 78.7^\circ$, and $\gamma = 86.3^\circ$. The asymmetric unit contained eight subunits. Wild-type and SeMet crystals were both frozen for X-ray data collection, but the only suitably diffracting crystal was one labeled with SeMet that was transferred step-wise into a

solution composed of 18% poly(ethylene glycol) 8000, 250 mM KCl, 100 mM MOPS (pH 7.0), 10 mM MgCl₂, 5 mM AMP-PNP, 5 mM CDP-NH₂, and 20% trehalose.

X-ray data sets were collected at the Structural Biology Center beamline 19-ID at a wavelength of 0.979 Å (Advanced Photon Source). The X-ray data were processed and scaled with HKL3000 (20). Relevant X-ray data collection statistics are provided in Table 6.1. Selenium sites were located using SHELXD (21); phases were calculated using MLPHARE (22); density modification was performed using DM (23), and an initial model was built using Buccaneer (24). Large portions of the protein chains were missing, so the most complete subunit was subsequently used as a search model to re-solve the structure via molecular replacement with Phaser assuming ten monomers in the asymmetric unit based on a Matthews coefficient (V_M) of 2.4 Å³/Da, or 49% solvent (25, 26). The resulting map was improved via molecular averaging with DM (27). However, upon building and checking the model using COOT, it became clear that the asymmetric unit did not contain ten monomers, but instead eight (28). The structure was solved a final time via molecular replacement with Phaser using the dimer as a search model assuming eight monomers in the asymmetric unit to yield a V_M of 3.1 Å³/Da or 60% solvent. While the model was built with COOT and refined with REFMAC, it was clear that higher order crystals would be necessary to complete a sufficient structural analysis of Cj1415 (29). In an attempt to determine the location of the Cj1415 active site, the initial model was superimposed with that from the *A. thaliana* (AT) PDB ID: 3UIE (30), and it was clear that density for AMP-PNP was lacking whereas that for the predicted CDP-NH₂ was discernable.

Therefore, variants of the enzyme with altered residues at the surface of Cj1415 were generated in an attempt to modify crystal packing and reduce the contents of the asymmetric unit. Surface entropy reduction variants of Cj1415 were cloned via the QuikChange method of Stratagene.

VAR1 was constructed based on the Surface Entropy Reduction server (31, 32), whereas VAR2 and VAR3 were produced based on the initial structure of Cj1415. Variants were expressed and purified as described for the wild-type protein except that VAR2 and VAR3 were dialyzed against 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl and were concentrated to 10-12 mg/mL. VAR1-3 were screened for crystallization properties. While VAR2 crystallized in the presence of 10 mM MgCl₂, 5 mM CDP-NH₂, ± 5 mM AMP-PNP, attempts to grow diffraction quality crystals via seeding and batch methods were unsuccessful. VAR2 was a triple variant (E158A/E159A/K160A), so the single amino acid variants were produced and purified the same as wild-type protein. Crystallization trials performed in the absence of AMP-PNP were successful. X-ray diffraction quality crystals of the VAR5 protein (E159A) were grown by mixing 1:1 the protein sample at 16 mg/mL containing 10 mM MgCl₂, 5 mM CDP-NH₂, and 200 mM NaCl with a precipitant solution composed of 15 - 18% poly(ethylene glycol) 8000, 100 mM HEPES (pH 7.5), 200 mM NaCl, and 2 - 4% ethylene glycol. The crystals belonged to the space group *C2* with unit cell dimensions of $a = 216.1 \text{ \AA}$, $b = 45.3 \text{ \AA}$, $c = 92.5 \text{ \AA}$, and $\beta = 95.2^\circ$. The asymmetric unit contained four subunits. Prior to X-ray data collection, the crystals were transferred step-wise into a solution composed of 25% poly(ethylene glycol) 8000, 500 mM NaCl, 100 mM HEPES (pH 7.5), 10 mM MgCl₂, 5 mM CDP-NH₂, and 18% ethylene glycol.

X-ray data sets were collected at the Structural Biology Center beamline 19-ID at a wavelength of 0.979 Å (Advanced Photon Source). The X-ray data were processed and scaled with HKL3000 (20). Relevant X-ray data collection statistics are provided in Table 6.1. The Cj1415 VAR5 structure was solved via molecular replacement with the software package PHASER (25) using the initial model described above as a search probe. Molecular averaging was performed with DM (27). Iterative cycles of model building with COOT (28) and refinement with REFMAC (29)

reduced the R_{work} and R_{free} to 23.9% and 32.0%, respectively. Relevant model refinement statistics are provided in Table 6.1.

6.3.3 Size exclusion chromatography

A Superdex 200 10/300 GL (GE Healthcare) column was equilibrated at room temperature with a buffer solution of 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl on an ÄKTA high-performance liquid chromatography (HPLC) system. TEV-cleaved protein, at 14 mg/mL in 10 mM Tris-HCl (pH 8.0) and 200 mM NaCl was applied to the column. The retention was monitored via ultraviolet absorption at 280 nm. Results were compared to standards of known molecular masses using the equation:

$$\log(\text{MW}) = -0.1123(\text{retention time}) + 8.05$$

6.3.4 HPLC activity assays

Wild-type and VAR5 Cj1415 were assayed in 50 mM HEPES (pH 8.0) at 1 mg/mL in the presence of 1.0 mM ATP, 1.0 mM CDP-NH₂, and 2.0 mM MgCl₂ at room temperature for 10 min. The reaction was stopped via removal of the enzymes using a 10 kDa cutoff filter. Reaction products were diluted 10-fold in water and analyzed on a 1 mL Resource Q column using a 10 column volume, 0 – 3.0 M ammonium acetate (pH 4.0) gradient. Peaks were assigned based on standards of all substrates and products except the phospho-CDP-NH₂ product, with CDP-NH₂ eluting at a retention volume of 6.6 mL, adenosine diphosphate (ADP) at 8.5 mL, presumed phospho-CDP-NH₂ at 9.6 mL, and ATP at 11.4 mL.

Table 6.1 X-ray Data Collection Statistics and Model Refinement Statistics.

	SeMet-labeled protein	E159A variant/ MgCDP-NH ₂
resolution limits	50 – 3.10 (3.13 – 3.10) ^b	50.0 – 2.6 (2.63 – 2.60)
Space Group	<i>P1</i>	<i>C2</i>
Unit Cell		
<i>a</i> (Å)	46.7	216.1
<i>b</i> (Å)	97.3	45.4
<i>c</i> (Å)	110.7	92.4
α (°)	81.7	
β (°)	78.7	95.1
γ (°)	86.3	
number of independent reflections	31243 (789)	27877 (1412)
completeness (%)	90.0 (91.6)	98.9 (99.4)
redundancy	3.7 (3.6)	6.4 (6.5)
avg I/avg σ(I)	17.2 (2.9)	43.9 (4.8)
<i>R</i> _{sym} (%) ^a	5.9 (39.9)	7.7 (33.3)
^c <i>R</i> -factor (overall)%/no. reflections		24.3/26528
<i>R</i> -factor (working)%/no. reflections		23.9/25180
<i>R</i> -factor (free)%/no. reflections		32.0/1348
Number of protein atoms		5353
Number of heteroatoms		0
average B-values		
protein atoms (Å ²)		84.5
weighted RMS deviations from ideality		
bond lengths (Å ²)		0.011
bond angles (°)		1.5
general planes (°)		0.006
Ramachandran regions (%)^d		
most favored		91.2
additionally allowed		7.1
generously allowed		1.7

$$^a R_{\text{sym}} = (\sum |I - \bar{I}| / \sum I) \times 100.$$

^bStatistics for the highest resolution bin.

^c*R*-factor = $(\sum |F_o - F_c| / \sum |F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

^dDistribution of the Ramachandran angles according to PROCHECK (33)

6.4 Results and Discussion

6.4.1 Overall architecture of Cj1415

Wild-type Cj1415 crystallized in the presence of MgAMP-PNP and CDP-NH₂ at 4 °C, and the first structure determined was phased via anomalous selenomethionine diffraction. A preliminary model was refined to 3.1 Å resolution in the space group *P1*. The asymmetric unit contained eight monomers with portions of the protein chains incomplete. To determine the quaternary structure of Cj1415, the enzyme was subjected to size exclusion chromatography and ran as a species with a molecular weight of ~39,500, consistent with the presence of two monomers (individual molecular weight of ~19,900) [Fig. 6.1(a)]. The asymmetric unit thus contained four dimers. Furthermore, there was no discernable electron density for AMP-PNP, whereas density for CDP-NH₂ was present and predicted based on the location of APS in APSKs (30).

In order to improve the quality of the crystals, two approaches were undertaken. First, a surface entropy reduction variant was produced (K128A, E129A) (31, 32); however, this variant did not crystallize in the presence of MgAMP-PNP and CDP-NH₂. Next, two variant enzymes were purified and surveyed to modify crystalline contacts in the *P1* lattice and reduce the contents of the asymmetric unit. VAR2 (E158A, E159A, K160A) and VAR3 (K28A, K30A, K32A, K34A) did not display improved crystallization properties. While the VAR2 enzyme crystallized, numerous attempts to grow suitably sized crystals were unsuccessful. Thus, the individual amino acid variants were purified and indeed, diffraction quality crystals of VAR5 (E159A) were grown in the presence of MgCDP-NH₂. VAR5 activity was assessed via HPLC and found to turn over a comparable amount of CDP-NH₂ to wild-type enzyme in ten minutes (Fig. 6.2). The enzyme was deemed catalytically active and thus appropriate to use for further crystallographic analysis.

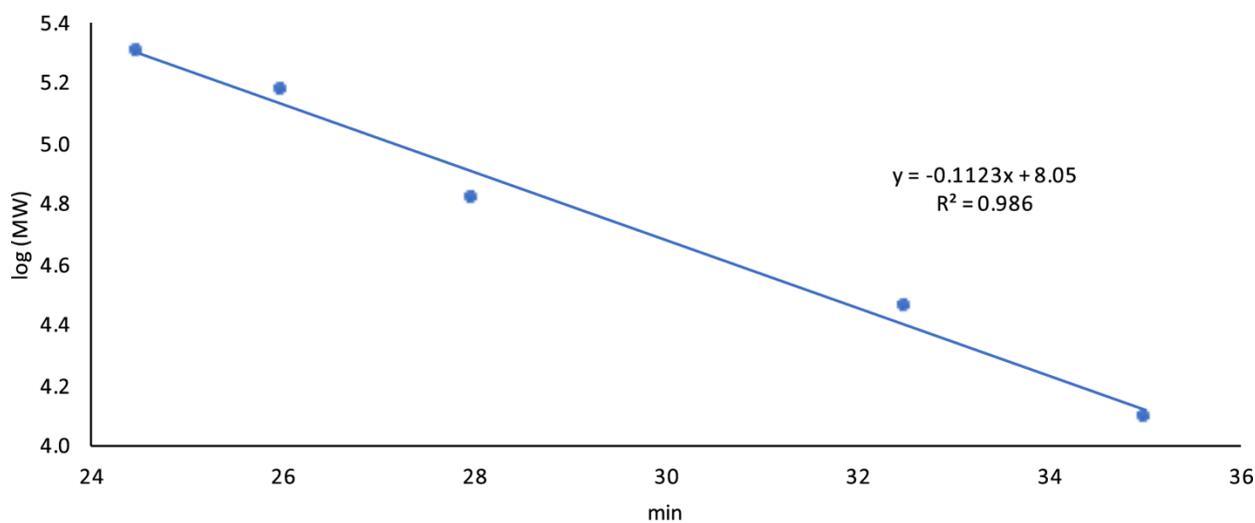
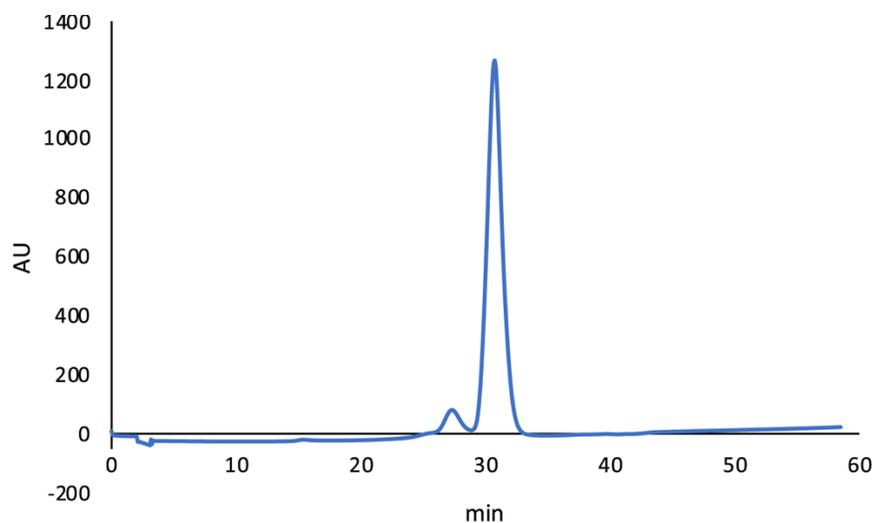


Figure 6.1 Size exclusion chromatography of Cj1415. Cj1415 was analyzed via size exclusion chromatography to determine its oligomerization state. Based on a log (MW) graph of standards of known molecular mass (b), the retention time of 30.75 min for Cj1415 (a) corresponds to a dimer species. Standards used were: cytochrome *c* (12,400 MW), carbonic anhydrase (29,000 MW), bovine serum albumin (66,000 MW), alcohol dehydrogenase (150,000 MW), and β -amylase (200,000 MW).

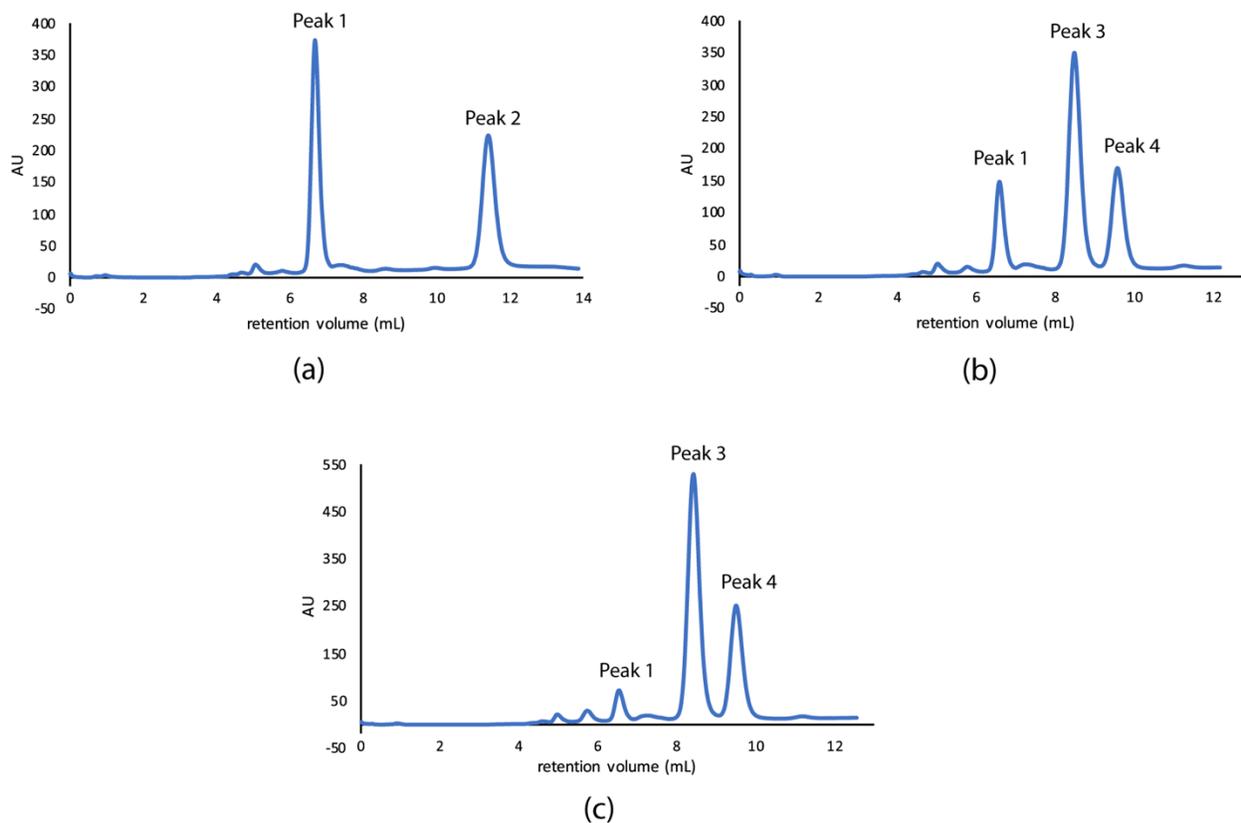


Figure 6.2 HPLC assay of wild-type and VAR5 Cj1415 activity. Cj1415 kinase activity was assessed in the absence (a) and presence (b) of Cj1415. (a) Peak 1 (6.6 mL) corresponds to CDP-NH₂, and peak 2 (11.4 mL) corresponds to ATP. (b) Peak 3 (8.5 mL) corresponds to ADP, and Peak 4 (9.6 mL) corresponds to the presumed phospho-CDP-NH₂ product. Finally, the elution profile for the activity assay of the E159A variant used in crystallization trials is presented in (c).

The VAR5 enzyme crystallized in the space group $C2$, and the contents of the asymmetric unit were reduced to two dimers. The structural model was refined to 2.6 Å resolution to an overall R -factor of 24.3%. Glu 159 is found on α -helix 7 [Fig. 6.3. (b)] and its mutation to an alanine residue did not alter the orientation of this helix. Indeed, the residue projects towards solvent in the wild-type structure, which is why it was mutated in the first place. The α -carbons for the VAR5 and wild-type enzymes superimpose with a root-mean-squared deviation (r.m.s.d.) of 0.5 to 0.6 Å. For clarity in the remainder of this discussion, the VAR5 enzyme will be referred to as Cj1415.

Shown in Figure 6.3 (a) is a ribbon representation of the Cj1415 dimer. The dimer has overall dimensions of ~ 74 Å X 36 Å X 48 Å demonstrating a total buried surface area of ~ 1600 Å². The subunit-subunit interface is composed primarily of residues from α -helices 2 and 3 [Figure 6.3. (b)]. The α -carbons for the individual subunits superimpose with a r.m.s.d. of 0.4 Å or 0.5 Å (subunits A and D). The polypeptide chains extend from Asn 4 to Leu 170 in subunits A, B, and C while the N-terminus was more ordered in subunit D, with density observed for the side chain of Asn 3 and the backbone of Lys 2. The large number of Ramachandran outliers, in addition to other refinement statistics and the resolution of the data, however, clearly demonstrates the limitations of the model, including accurate placement of backbone and side chain atoms. Indeed, a number of side chains were disordered, lacking any electron density calculated with $(2F_o - F_c)$ coefficients contoured at 1σ , and were thus modeled as alanine residues. Nonetheless, the dimer comprised of subunits A and B was marginally more ordered than that of subunits C and D and will serve as the basis for further discussion.

Shown in Figure 6.3 (b) is a ribbon representation of the Cj1415 monomer. Like the APSKs, Cj1415 contains an α/β purine nucleotide binding domain with a five-stranded parallel β -sheet flanked on either side by α -helices, displaying the topology $\beta 1-\alpha 1-\beta 2-\alpha 2-\alpha 3-\beta 3-\alpha 4-\beta 4-\beta 5-\alpha 7$. Helices 2, 3, and 4 sandwich the β -sheet on one face while helices 1 and 7 form the other face. Helices 5 and 6 cap the likely nucleotide binding site. Cj1415 possesses a P-loop ATP binding domain with the “Walker A” motif (GX_4GKS/T) comprised of $^{12}GLAGSGKT^{19}$ positioned between $\beta 1$ and $\alpha 1$ (34). The “Walker B” aspartate predicted to orient the required Mg^{2+} ion in Cj1415 (Asp 40) is situated at the end of $\beta 2$ (30). Analogous to APSKs, Cj1415 contains an extra β -strand between its P-loop and Walker B motifs, whereas in other α/β purine nucleotide binding domain proteins, such as thymidylate kinase, the Walker B aspartate is positioned at the end of the equivalent of $\beta 3$ in Cj1415 [Fig. 6.3 (c)] (35).

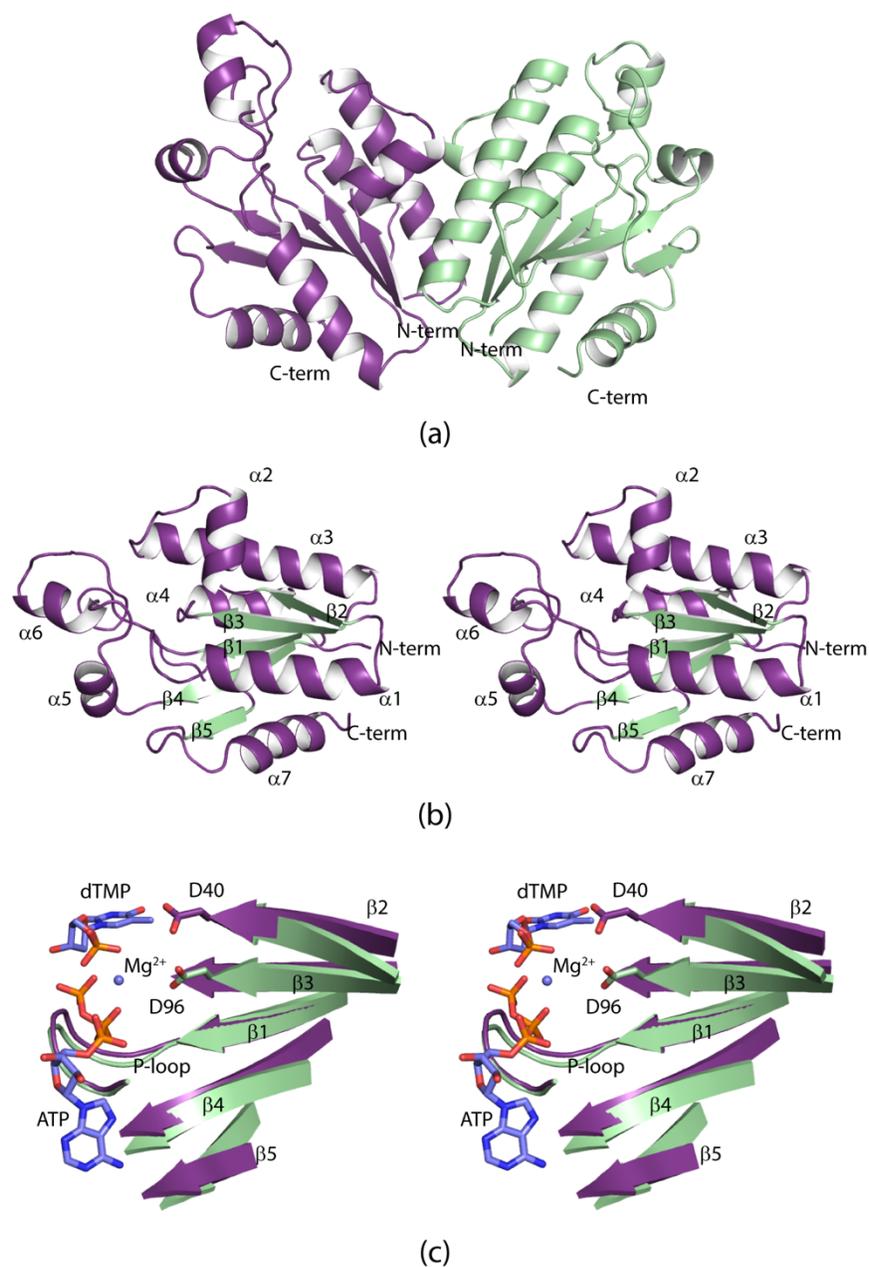


Figure 6.3 Overall structure of Cj1415. A ribbon representation of the Cj1415 dimer is shown in (a) with the two subunits displayed in purple and green, respectively. A stereo view of subunit A is provided in (b) with β -strands highlighted in green. α -helices and β -strands are numbered according to the text, from the N- to the C-terminus. Finally, shown in (c) is a stereo view of the β -sheet portion of the α/β purine nucleotide binding domain in Cj1415 (purple) and thymidylate kinase (TK) (green, 1E2Q) (35). The strands are numbered according to those in Cj1415 with the Walker B aspartates displayed as sticks. The thymidine monophosphate (dTMP), ATP, and Mg^{2+} from TK are shown in slate. This figure, and Figures 6.4, 6.5, and 6.6 were prepared with PyMOL (36).

Structural coordinates for a number of monofunctional and bifunctional APSK have been deposited in the PDB. The N-termini of monofunctional APSKs are quite diverse due to their different modes of regulation and localization sequences in certain species (17, 30). Because the core APSK domains are similar, Cj1415 was compared to a number of APSKs for which there are structural data. The DALI server was used to relate the overall three-dimensional structure of the Cj1415 monomer with those from APSK's from *A. thaliana* (AT) (DALI z-score = 22.5, r.m.s.d of 1.7 Å over 191 C α 's, PDB: 3UIE), *Penicillium chrysogenum* (PC) (DALI z-score = 23.5, r.m.s.d of 1.4 Å over 182 C α 's, PDB: 1M7G), and *Synechocystis* sp. PCC 6803 (SP) (DALI z-score = 24.5, r.m.s.d of 1.5 Å over 173 C α 's, PDB: 5CB6) (17, 30, 37, 38). The architectures of the monomers are nearly identical, as demonstrated in Figure 6.4. (a) (the N-termini of the APSKs were deleted for clarity).

However, the monomers associate in slightly different dispositions to form dimers. Subunit A from Cj1415 and an APSK were superimposed in order to observe the subsequent orientation of the B subunits in the dimer complex [Fig. 6.4 (b)-(d)]. The r.m.s.d.s between the Cj1415 dimer those from AT, PC, and SP are 2.7 Å, 2.8 Å, and 3.8 Å, respectively. Subunit B from AT, PC, and SP APSK is displaced 12° and 3.3 Å, 12° and 3.7 Å, and 27° and 4.6 Å, respectively, from the Cj1415 B subunit. The altered dimer structures have been noted previously in a comparison of the AT and SP enzymes (17). Based on the location of α -helix 3, it appears that the Cj1415 monomers do not associate as tightly as the AT and CP enzymes but interact more so than the SP subunits. Besides exhibiting variable dimer interfaces, small structural changes in APSK have been observed which promote an ordered binding mechanism (17, 37). Structures of Cj1415 solved in the presence of different combinations of ligands could shed light on the three-dimensional changes necessary for substrate binding and product release.

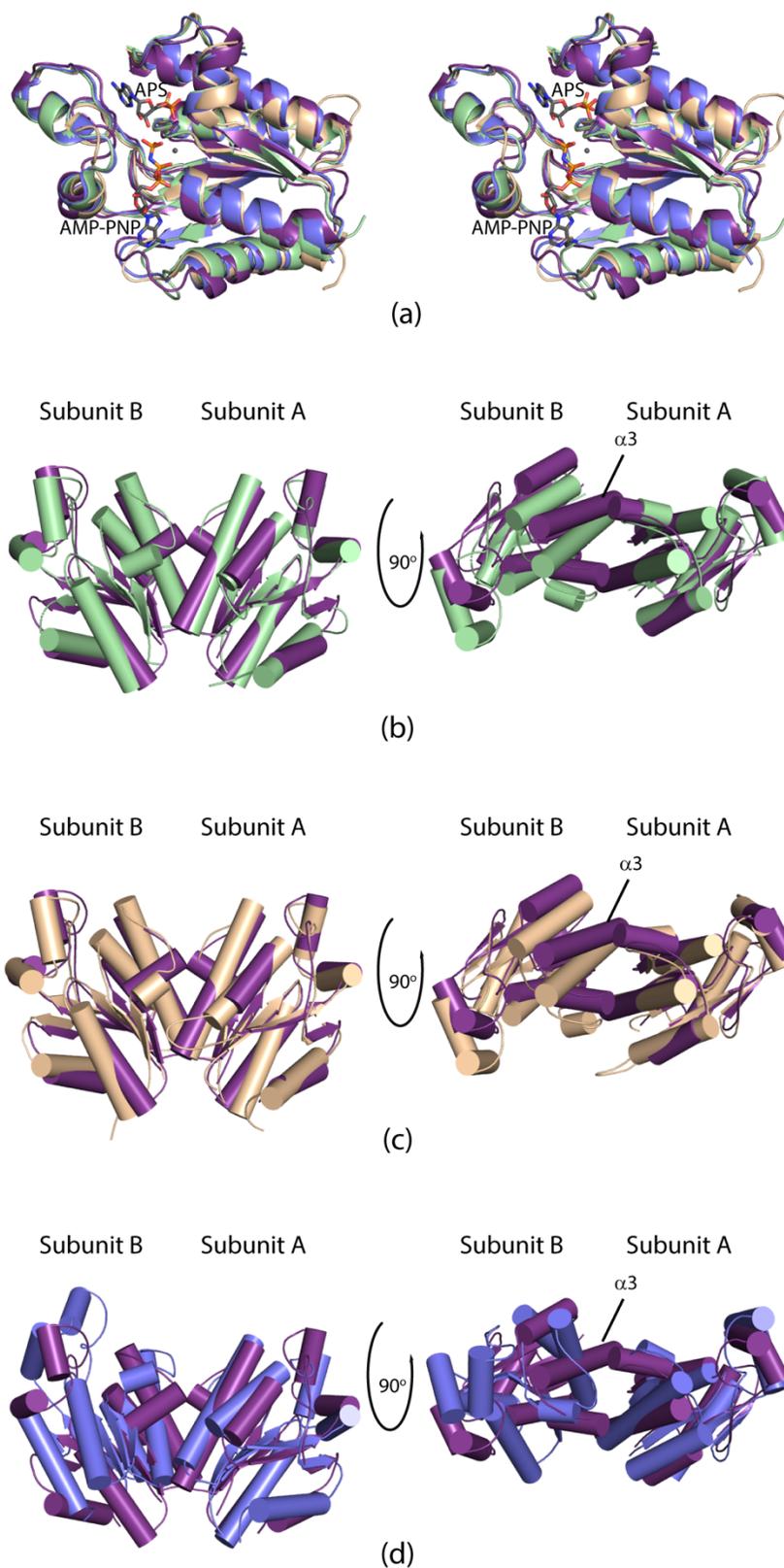


Figure 6.4 Cj1415 monomer and dimer comparisons to APSKs. Shown in part (a) is a superposition of the Cj1415 monomer (purple) and APSK monomers from AT (green), PC (wheat), and SP (slate). APS, AMP-PNP, and Mg^{2+} from the AT structure are shown in grey. In (b-d), subunit A from Cj1415 and an APSK were aligned to show the disposition of subunit B, demonstrated most clearly in the rotated dimer. The color scheme is as previously described with the AT dimer shown in (b), the PC dimer shown in (c), and the SP dimer shown in (d). α -helix 3 is highlighted to indicate its variable position between the structures.

6.4.2 Potential Cj1415 CDP-NH₂ binding site

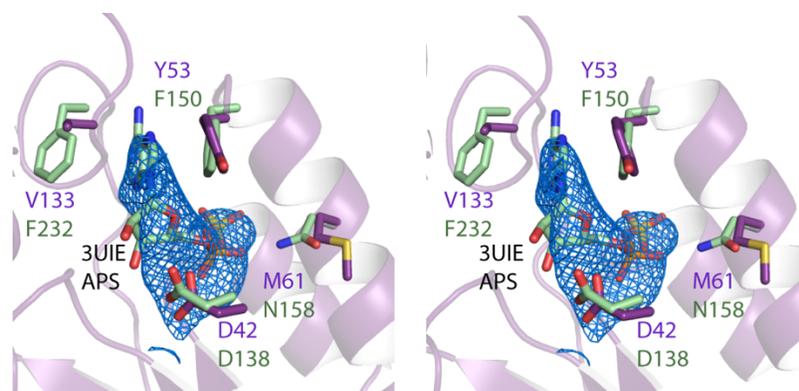
There was clear electron density calculated with (F_o-F_c) coefficients contoured at 3.0σ in the vicinity of Asp 42 in Cj1415. The equivalent aspartate in AT APSK is situated near the ribose hydroxyl groups of APS (30). Furthermore, in a superposition of the Cj1415 and APSK monomers, the APS substrate from the latter overlaps with the difference density in Cj1415 (Fig. 6.5). However, due to the insufficient resolution of the data, accurate modeling of ligand atoms in this density as well as residues involved in binding is not possible. Discussion provided here is therefore speculative until higher resolution data are collected.

A superposition of Cj1415 with the active site from AT APSK offers some insight into CDP-NH₂ binding determinants in Cj1415. Postulations are based on residues in subunit A with reasonable electron density calculated with ($2F_o-F_c$) coefficients contoured at 1.5σ . Whereas many residues in Cj1415 superimpose with equivalent amino acids implicated in APS recognition from AT (APSK/Cj1415 numbering: Asp 138/Asp 42, Arg 141/Arg 65, Arg 155/Arg 58, Ile 181/Ile 84), there are five notable exceptions [Fig. 6.5. (c)]. Phe 150 and Phe 232 in APSK stack either side of the adenine ring of the APS substrate (30); however, the residue at the 150 position in Cj1415 is Tyr 53 and the 232 position is occupied by Val 133 in Cj1415. Although the loop containing Val 133 was not well ordered in all subunits, it along with Tyr 53 may provide specificity for the cytosine ring of CDP-NH₂ (16).

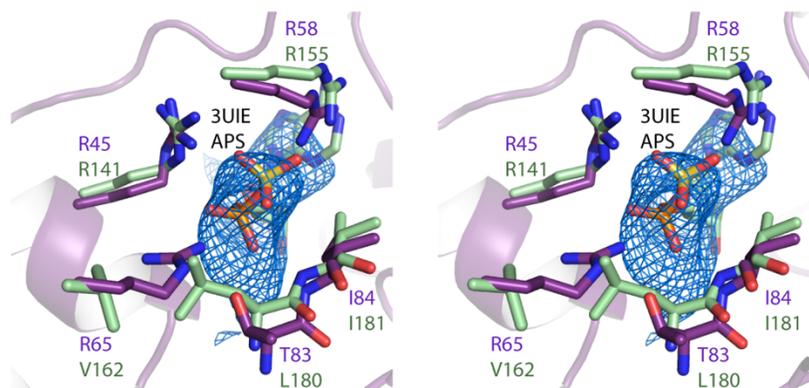
Asn 158 from APSK interacts with APS, however density for the equivalent residue in Cj1415, Met 161, is not directed towards the electron density calculated with (F_o-F_c) coefficients. Similarly, Val 161 in APSK superimposes with Arg 65 in Cj1415. The Arg 65 side chain appears to clash with the side chain of Leu 180 from APSK in the superposition [Fig. 6.5. (b)]. Intriguingly, the 180 position in Cj1415 is occupied by a threonine instead of a leucine.

6.4.3 Role of Ser 85

In the original work elucidating the function of Cj1415, Ser 85 was mutated to an alanine residue, based on studies of *E. coli* APSK, to probe the requirement of a phospho-enzyme intermediate in catalysis (16). The Cj1415 S85A variant was still active but demonstrated a reduced K_M . While the structure described here confirms that Ser 85 is near the electron density calculated with (F_o-F_c) coefficients predicted for the CDP-NH₂ ligand, rationale for the reduced K_M of the variant is not possible. In the APSK structure from AT, the backbone amide of the equivalent serine (Ser 182) donates a hydrogen bond to stabilize the APS substrate (30). While this might be possible in Cj1415, the side chain hydroxyl may also stabilize the substrate. However, the exact positions of the Ser 85 backbone and side chain were not clear in this structure. Thus, the reduced K_M of the Cj1415 S85A variant could be due to mis-positioning of the backbone amide group and/or loss of a hydrogen bond from the side-chain hydroxyl.



(a)



(b)

```

Cj1415      1 -----MKNPYIIWLTGLAGSGKTTIGQALYEKLLKLYKNLIYLD  40
3UIE       77 NSTNIKWHECSVEKVDRQRLLDQKGCVIWVTGLSGSGKSTLACALNQMLYQKGLKCYILD  136
                **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

Cj1415     41 GDELREILGH-YAYDRQGRIDMALKRAKFAKFLNDQGMVIVTTISMFNEIYDYNRKQ--  97
3UIE      137 GDNVRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGIICIASLISPVRTDRACRSLLP  196
                **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

Cj1415     98 LKNYYEIIYIECDMHელიQRDQKGLYTKALNKEIDNVVGVDFIEFDKPEADLVINNSCRNNL  157
3UIE      197 EGFVFEVFMVPLSVCEARDPKGLYKLARAGKIKGFTGIDDPYEPPLNCEISLGR--EGG  254
                *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

Cj1415     158 EEKVELIIKKLAL-----  170
3UIE      255 TSPiEMAekVVGyLDNKGyLQA  276
                *  *

```

(c)

Figure 6.5 Superposition of the potential Cj1415 CDP-NH₂ binding site with the APS binding site from APSK. Shown in (a) is a superposition of residues from APSK implicated in APS binding with the equivalent residues in Cj1415. Cj1415 and APSK residues are colored and labeled in purple and green, respectively. The Cj1415 electron density calculated with ($F_o - F_c$) coefficients contoured at 3σ is shown in blue mesh. The APS molecule is from the 3UIE structure. Shown in (b) is a rotated view of that which is shown in (a) to highlight different residues. Finally, a sequence alignment for Cj1415 and plant APSK is provided in (c) with the residues shown in parts (a) and (b) colored red. The alignment was made with Clustal Omega.

6.4.4 Potential Cj1415 ATP binding site

The likely site of ATP binding in Cj1415 is less clear because crystals were grown in the absence of AMP-PNP. However, based on the properties of P-loop containing proteins and a superposition of the Cj1415 and AT APSK monomers, some predictions are possible. The ATP molecule is likely solvent exposed in Cj1415. Indeed, the majority of contacts with AMP-PNP in AT APSK are mediated via phosphate interactions with backbone amide groups in the P-loop (30). The P-loops from each enzyme are highlighted in Figure 6.6. (a) and overlay quite well. Asp 136 in APSK, the Walker B aspartate, positions Ser 115 and one water molecule in the Mg^{2+} coordination sphere (30); reasonable density calculated with $(2F_o-F_c)$ coefficients was observed for the equivalent residue in Cj1415, Asp 40, and is therefore displayed. The ribose hydroxyls, on the other hand, are exposed to solvent in APSK.

The adenine ring of AMP-PNP is wedged between Arg 215 and Pro 257 in APSK (30). There was clear electron density calculated with $(2F_o-F_c)$ coefficients for the equivalent arginine in Cj1415 (116), displayed in Figure 6.6 (a). However, the 257 position in Cj1415 is occupied by a leucine, not a proline. Furthermore, the loops between $\beta 5$ and $\alpha 7$ in each enzyme are quite disparate in primary amino acid sequence, so it is unclear without an AMP-PNP bound Cj1415 structure if and how this loop interacts with the adenine ring of ATP [Fig. 6.6 (b)]. However, the Cj1415 loop may be orientated to occupy the potential ATP binding site in the absence of the ligand, whereas in the AT APSK structure, the loop is shifted away from the adenine ring to accommodate it.

K_M values of 0.1 mM and 1.5 mM for CDP-NH₂ and ATP, respectively, were reported for Cj1415 (16). With CDP-NH₂ present, Cj1415 crystallization trials were only conducted with 5 mM AMP-PNP. The high K_M value for ATP and the potential solvent exposed nature of the binding pocket could explain why no electron density was observed in the initial Cj1415 structure solved in the presence of 5 mM AMP-PNP. Furthermore, APSK follows an ordered mechanism wherein ATP binds first but is subject to inhibition by its own substrate, APS (39). If Cj1415 exhibits similar inhibition patterns, this may explain why the high ratio of CDP-NH₂:AMP-PNP in crystallization trials precluded AMP-PNP binding and/or why the enzyme was able to accommodate CDP-NH₂ in the absence of AMP-PNP.

6.4.5 Conclusions

An initial three-dimensional model of Cj1415 from *C. jejuni* was determined that will enable further structural analyses of the enzyme. Whereas the Cj1415 monomer aligns well with those from APSKs, the dimer structures and subunit interfaces are different. Furthermore, residues comprising the CDP-NH₂ binding site were predicted based on a superposition of the electron density calculated with (F_o-F_c) coefficients in Cj1415 with the structure of plant APSK solved in the presence of APS and MgAMP-PNP. Certain residues implicated in CDP-NH₂ specificity could be predicted from sequence alignments [Figure 6.5. (c)] (16). Tyr 53 and Val 133 do seem to be in proximity to the likely binding site of the cytosine ring of CDP-NH₂. However, amino acids such as Arg 65 and Thr 83 were not predicted because their equivalents in APSK (Val 62 and Leu 180) do not interact with APS.

The ATP binding site in Cj1415 is likely solvent exposed based on the structure of plant APSK and the disposition of the AMP-PNP molecule in this enzyme (30). Based on the properties of P-loop proteins, backbone amide groups likely interact extensively with the phosphate groups in Cj1415. Conversely, the manner in which the enzyme accommodates the adenine ring of ATP could not be deduced from a superposition of the enzyme with APSK because the loop connecting $\beta 5$ and $\alpha 7$ is not positioned in a manner conducive to ATP binding.

In summary, the likely site of CDP-NH₂ recognition in Cj1415 was predicted, along with putative amino acids implicated in its binding. Higher resolution data will illuminate the manner in which Cj1415 accommodates a cytosine-based substrate as well as one with two hydrogen bond donors (-NH₂) at the β position compared to APS which only harbors hydrogen bond acceptors at the equivalent location.

6.5 References

1. Young, K. T., Davis, L. M., and DiRita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* *5*, 665-679.
2. Nyati, K. K. and Nyati, R. (2013) Role of *Campylobacter jejuni* infection in the pathogenesis of Guillain-Barré syndrome: an update. *Biomed. Res. Int.* *2013*, doi: 10.1155.
3. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holfoyd, S., Jagels, K., Karlyshev, A. V., Moule, S., Pallen, M. J., Penn, C. W., Quail, M. A., Rajandream, M. A., Rutherford, K. M., van Vliet, A. H. M., Whitehead, S., and Barrell, B. G. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature.* *403*, 665-668.
4. Guerry, P. and Szymanski, C. M. (2008) *Campylobacter* sugars sticking out. *Trends Microbiol.* *16*, 428-435.
5. Kanipes, M. I. and Guerry, P., *Role of microbial glycosylation in host cell invasion*, in *Microbial Glycobiology: Structures, Relevance and Applications*, A.P. Moran, et al., Editors. 2009, Elsevier Inc. p. 871-890.
6. Bachtiar, B. M., Coloe, P. J., and Fry, B. N. (2007) Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 8116 and its involvement in bacterium-host interactions. *FEMS Immunol. Med. Microbiol.* *49*, 149-154.
7. Bacon, D. J., Szymanski, C. M., Burr, D. H., Silver, R. P., Alm, R. A., and Guerry, P. (2001) A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol. Microbiol.* *40*, 769-777.
8. van Alphen, L. B., Wenzel, C. Q., Richards, M. R., Fodor, C., Ashmus, R. A., Stahl, M., Karlyshev, A. V., Wren, B. W., Stintzi, A., Miller, W. G., Lowary, T. L., and Szymanski, C. M. (2014) Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One.* *9*, e87051.
9. Maue, A. C., Mohawk, K. L., Giles, D. K., Poly, F., Ewing, C. P., Jiao, Y., Lee, G., Ma, Z., Monteiro, M. A., Hill, C. L., Ferdberber, J. S., Porter, C. K., Trent, M. S., and Guerry, P. (2013) The polysaccharide capsule of *Campylobacter jejuni* modulates the host immune response. *Infect. Immun.* *81*, 665-672.
10. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., and Monteiro, M. A. (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines *Front. Cell Infect. Microbiol.* *2*, doi: 10.3389.
11. McNally, D. J., Lamoureux, M. P., Karlyshev, A. V., Fiori, L. M., Li, J., Thacker, G., Coleman, R. A., Khieu, N. H., Wren, B. W., Brisson, J. R., Jarrell, H. C., and Szymanski, C. M. (2007) Commonality and biosynthesis of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* *282*, 28566-28576.

12. Wong, A., Lange, D., Houle, S., Arbatsky, N. P., Valvano, M. A., Knirel, Y. A., Dozois, C. M., and Creuzenet, C. (2015) Role of capsular modified heptose in the virulence of *Campylobacter jejuni*. *Mol. Microbiol.* *96*, 1136-1158.
13. Kim, S., Vela, A., Clohisey, S. M., Athanasiadou, S., Kaiser, P., Stevens, M. P., and Vervelde, L. (2018) Host-specific differences in the response of cultured macrophages to *Campylobacter jejuni* capsule and *O*-methyl phosphoramidate mutants. *Vet. Res.* *49*, s13567-017-0501-y.
14. Taylor, Z. W., Brown, H. A., Narindoshvili, T., Wenzel, C. Q., Szymanski, C. M., Holden, H. M., and Raushel, F. M. (2017) Discovery of a glutamine kinase required for the biosynthesis of the *O*-methyl phosphoramidate modifications found in the capsular polysaccharides of *Campylobacter jejuni*. *J. Am. Chem. Soc.* *139*, 9463-9466.
15. Taylor, Z. W., Brown, H. A., Holden, H. M., and Raushel, F. M. (2017) Biosynthesis of nucleoside diphosphoramidates in *Campylobacter jejuni*. *Biochemistry.* *56*, 6079-6082.
16. Taylor, Z. W. and Raushel, F. M. (2018) Cytidine diphosphoramidate kinase: an enzyme required for the biosynthesis of the *O*-methyl phosphoramidate modification in the capsular polysaccharides of *Campylobacter jejuni*. *Biochemistry.* *57*, 2238-2244.
17. Herrmann, J., Nathin, D., Lee, S. G., Sun, T., and Jez, J. M. (2015) Recapitulating the structural evolution of redox regulation in adenosine 5'-phosphosulfate kinase from cyanobacteria to plants. *J. Biol. Chem.* *290*, 24705-24714.
18. Thoden, J. B. and Holden, H. M. (2005) The molecular architecture of human *N*-acetylgalactosamine kinase. *J. Biol. Chem.* *280*, 32784-32791.
19. Doublet, S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic systems. *Methods Mol. Biol.* *363*, 91-108.
20. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* *276*, 307-326.
21. Schneider, T. R. and Sheldrick, G. M. (2002) Substructure solution with *SHELXD*. *Acta Crystallogr. D. Biol. Crystallogr.* *58*, 1772-1779.
22. Otwinowski, Z., *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering.*, W. Wolf, P.R. Evans, and A.G.W. Leslie, Editors. 1991: Warrington: Daresbury Laboratory. p. 80-86.
23. Cowtan, K. D. and Zhang, K. Y. (1999) Density modification for macromolecular phase improvement. *Prog. Biophys. Mol. Bio.* *72*, 245-270.
24. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D. Biol. Crystallogr.* *62*, 1002-1111.

25. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658-674.
26. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* *44*, 491-497.
27. Cowtan, K. T. (1994) "DM:" An automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34-38.
28. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* *60*, 2126-2132.
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D. Biol. Crystallogr.* *53*, 240-255.
30. Ravilious, G. E., Nguyen, A., Francois, J. A., and Jez, J. M. (2012) Structural basis and evolution of redox regulation in plant adenosine-5'-phosphosulfate kinase. *Proc. Natl. Acad. Sci. USA.* *109*, 309-314.
31. Derewenda, Z. S. (2004) Rational protein crystallization by mutational surface engineering. *Structure.* *12*, 529-535.
32. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S., and Eisenberg, D. (2007) Toward rational protein crystallization: a web server for the design of crystallizable protein variants. *Protein Sci.* *16*, 1569-1576.
33. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* *26*, 283-291.
34. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) Distantly related sequences in the α - and β - subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* *1*, 945-951.
35. Ostermann, N., Schlichting, I., Brundiers, R., Konrad, M., Reinstein, J., Veit, T., Goody, R. S., and Lavie, A. (2000) Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate. *Structure.* *8*, 629-642.
36. DeLano, W. L., The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA. 2002.
37. Lansdon, E. B., Segel, I. H., and Fisher, A. J. (2002) Ligand-induced structural changes in adenosine 5'-phosphosulfate kinase from *Penicillium chrysogenum*. *Biochemistry.* *41*, 13672-13680.

38. Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* *38*, W545-W549.
39. MacRae, I. J. and Segel, I. H. (1999) Adenosine 5'-phosphosulfate (APS) kinase: diagnosing the mechanism of substrate inhibition. *Arch. Biochem. Biophys.* *361*, 277-282.

6.6 Acknowledgements

A portion of the research described here was performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source (U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357). I gratefully acknowledge Dr. Norma E. C. Duke for assistance during X-ray data collection and processing.

Chapter 7:

Project CRYSTAL

(Colleagues Researching with Young Scientists: Teaching and Learning)

7.1 Goals of Project CRYSTAL

Project CRYSTAL, referred to originally as Crystallographers Researching with Young Scientists: Teaching and Learning, is a middle school outreach program funded in 2009 as a collaboration between the laboratory of Professor Hazel M. Holden and Mr. Dan Toomey, a science teacher at Edgewood Middle School in Madison. The goals of the program are to instill in students an excitement for chemistry and biochemistry at a young age by performing laboratory research alongside graduate and undergraduate mentors and to inspire an appreciation for the relevance of these topics to everyday life. The objectives are motivated by two observations: 1) Despite initial interest, many students become less enthusiastic about science as they proceed through formal education and 2) Performance by students in the United States on standardized science exams consistently lags behind those in other developed nations, a trend that continues today (1, 2).

Both points underscore a consensus reached by the American Association for the Advancement of Science (AAS) that science education reform is needed from the primary to the undergraduate levels (3, 4). At times, too much emphasis seems to be placed on the pipeline effect of preparing students for particular careers to the detriment of the fundamental notion in our education system and society that science should be taught to everyone. As a result, the focus on memorization and volume of information has undermined the emphasis on thinking critically about how science impacts all aspects of our lives (5). This teaching method deters both those who were not originally interested in a scientific career and those that were, stripping science of its relevance.

Project CRYSTAL therefore teaches fundamentals in chemistry and biochemistry in an active, hands on way by involving middle schoolers in laboratory research. We believe that middle school is not too early to learn biochemistry, a class normally first encountered in college, because it is a

relevant lens through which to learn chemistry. Indeed, biochemistry is the application of chemistry to biological systems and what makes life at the molecular level possible in the first place. A commitment to a career in science is obviously not a prerequisite for Project CRYSTAL students, for reasons just described; however, at an age when their interests are being pulled in different directions, we hope that by participating in authentic laboratory work and demystifying how science is portrayed on the television or the news, that students will seek out additional experiences in the area and appreciate its importance in day-to-day life.

Furthermore, data from the National Research Council indicate that just 18.5% of K – 12 learning occurs in formal classroom settings with the rest taking place in an informal context (6). Project CRYSTAL, then, is part of the critical 81.5% of educational opportunities in students' lives, time that we hope makes them more competitive relative to their peers around the world.

The third goal of Project CRYSTAL is to provide graduate and undergraduate students with mentoring and teaching opportunities. These experiences are especially valuable given that “one-size fits all” education practices are under scrutiny. Oftentimes, activities have to be tailored to the group and academic potential of the students, and the Project CRYSTAL environment is conducive to cultivating these particular teaching skills.

7.2 Project CRYSTAL Program Structure

The original focus of the Project CRYSTAL laboratory research was centered on the three-dimensional structural determinations of deoxysugar biosynthetic enzymes. The current projects will be discussed later, but because the overall workflow of the program has largely been consistent, it will be detailed here briefly. Students purify genomic DNA from a source organism, when possible, amplify their gene of interest via polymerase chain reaction (PCR), and eventually ligate it into an appropriate expression vector during the Fall. The Spring semester is devoted to

expressing protein in *Escherichia coli*, purification using nickel-affinity chromatography, crystallization screening, and culminates in building a three-dimensional model at the graphics system. Throughout the year, classroom topics include the structure of an atom, pH, and dimensional analysis, to the DNA→RNA→protein central dogma and the four levels of protein structure.

Four to six middle school students work with graduate or undergraduate student mentors, laboratory staff, and Professor Holden every Tuesday for ninety minutes. Every session starts with snacks (hungry teenagers do not focus well) and combines lectures and worksheets with laboratory work, with some activities requiring more introduction in the classroom than others. Every week, the mentors meet with Professor Holden on Friday to plan the Tuesday sessions.

I was privileged to join the Holden laboratory when the program had been running for a number of years, so lecture material and laboratory work were already well organized. Nonetheless, during the Friday planning sessions, we reviewed and updated lecture material, organized a detailed timeframe for the mentors serving as laboratory guides on Tuesday, and identified any experiments requiring advanced preparation (i.e. large-scale growths to induce during Project CRYSTAL).

Learning was assessed at the end of each semester through a Jeopardy-style game. Additionally, in the Spring, the students would make a poster about what they learned in Project CRYSTAL, with input from the mentors, to present at the Madison Middle School Science Symposium. While the success of each year was measured by these activities, it is worth noting that in three cases before my time, middle school students were acknowledged or served as co-authors on manuscripts published in *Biochemistry* for their assistance in growing diffraction quality protein crystals (7-9).

It became clear in the Spring of 2014, however, that studying deoxysugar biosynthetic enzymes did not always guarantee that a structure could be solved in the required time-frame. Complications arose such as insoluble protein expression and a lack of X-ray diffraction quality crystals. This, coupled with our desire to make the program accessible to crystallography and non-crystallography laboratories alike, led Professor Holden, Ari Salinger (a former graduate student), and me to design a curriculum that could be implemented consistently and distributed to other universities.

Dr. Aaron Hoskins in the Department of Biochemistry donated plasmids with eGFP, mCitrine, mCherry, and mPlum genes from which Ari and I designed cloning, expression, and purification experiments along with protein-specific crystallization screens. There is no better way to teach affinity chromatography, and concepts therein like elution, than with fluorescent proteins (we did teach the students how to use ultraviolet absorption to monitor elution, that most proteins are not visible with the naked eye and especially that most are not purple!). Lecture material was updated, and the year-long plan was modified. Thus, Project CRYSTAL, *Colleagues Researching with Young Scientists: Teaching and Learning*, began in the 2014-2015 school year. The materials have since been distributed to laboratories at the University of Wisconsin-Milwaukee, Marquette University, and Indiana University, Purdue University, Indianapolis.

The program went on hiatus at the end of 2015, but Professor Holden, Dr. James Thoden, the senior scientist in the Holden laboratory, along with current and former undergraduates Tommy Anderson, Nick Bockhaus, Cory Call, Brandon Dopkins, and Murray Dunsirn resumed Project CRYSTAL in the Fall of 2017 with six new students and six new fluorescent proteins. To date, the program has impacted 26 middle school students in the Madison area.

7.3 Project CRYSTAL Mentors

The subjectively boring lectures and irrelevant laboratory activities in formal education systems may be two reasons why students lose interest in science. Both, however, are perpetuated by those doing the teaching. In 2009 former editor-in-chief of *Science* magazine and president of the National Academy of Sciences, Dr. Bruce Alberts, put it this way:

“Rather than learning how to think scientifically, students are generally being told about science and asked to remember facts. This disturbing situation must be corrected if science education is to have any hope of taking its proper place as an essential part of the education of students everywhere. Scientists may tend to blame others for the problem, but – strange as it may seem – we have done more than anyone else to create it.” (10)

Science education would greatly benefit from increased creativity and active inquiry approaches, changes that will ultimately be implemented by teachers (11). While reform is needed across the board, university teaching practices inevitably serve as examples at the primary level and should thus lead the way. With respect to Project CRYSTAL, not every mentor will go on to teach undergraduates, high school, or middle school students. However, every scientist should be able to clearly articulate why their research (or lecture) is important and relevant, in addition to clearly disseminating their work or teaching material so that non-scientists can understand it.

Thus, Project CRYSTAL has been a wonderful training ground for young teachers and science advocates because of the opportunities to hone communication skills, participate in curriculum development, and implement alternative learning strategies in a less threatening environment than a classroom of undergraduates one might encounter as a teaching assistant. Successfully conveying the basics of college level biochemistry to young students is invaluable and encourages the mentors to use the proper language to communicate complex topics, but not to dilute or distort them. I am grateful for the lessons I learned from Project CRYSTAL researchers, my colleagues, and Professor Holden as I hope to continue working with students in the future.

7.4 References

1. Osborne, J., Simon, S., and Collins, S. (2003) Attitudes towards science: a review of the literature and its implications. *Int. J. Sci. Edu.* 25, 1049-1079.
2. Desilver, D. *U.S. students' academic achievement still lags that of their peers in many other countries.* 2017; [Pew Research Center]. Available from: <http://www.pewresearch.org/fact-tank/2017/02/15/u-s-students-internationally-math-science/>.
3. Funk, C. and Rainie, L. *Public and scientists' views on science and society.* 2015; [Pew Research Center]. Available from: <http://www.pewinternet.org/2015/01/29/public-and-scientists-views-on-science-and-society/>.
4. *Vision and change in undergraduate biology education: a call to action*, C.A. Brewer and D. Smith, Editors. 2011, American Association for the Advancement of Science: Washington, DC.
5. Feinstein, N. W., Allen, S., and Jenkins, E. (2013) Outside the pipeline: reimagining science education for nonscientists. *Science.* 340, 314-317.
6. National Research Council, *Learning Science in Informal Environments: People, Places, and Pursuits.* 2009, Washington, D.C.: The National Academies Press.
7. Bruender, N. A., Thoden, J. B., Kaur, M., Avey, M. K., and Holden, H. M. (2010) Molecular architecture of a C-3'-methyltransferase involved in the biosynthesis of D-tetronitose. *Biochemistry.* 49, 5891-5898.
8. Kubiak, R. L., Phillips, R. K., Zmudka, M. W., Ahn, M. R., Maka, E. M., Pyeat, G. L., Roggensack, S. J., and Holden, H. M. (2012) Structural and functional studies on a 3'-epimerase involved in the biosynthesis of dTDP-6-deoxy-D-allose. *Biochemistry.* 51, 9375-9383.
9. Kubiak, R. L. and Holden, H. M. (2012) Structural studies of AntD: an N-acyltransferase involved in the biosynthesis of D-Anthrose. *Biochemistry.* 51, 867-878.
10. Alberts, B. (2009) Redefining science education. *Science.* 323, 437.
11. Alberts, B. (2013) Prioritizing science education. *Science.* 340, 249.

Appendix 1:

Discovery of a Glutamine Kinase Required for the Biosynthesis of the *O*-Methyl Phosphoramidate Modifications Found in the Capsular Polysaccharides of *Campylobacter jejuni*

This work was previously published as: Taylor, Z. W., Brown, H. A., Narindoshvili, T., Wenzel, C. Q., Szymanski, C. M., Holden, H. M., and Raushel, F. M. (2017) Discovery of a Glutamine Kinase Required for the Biosynthesis of the *O*-Methyl Phosphoramidate Modifications Found in the Capsular Polysaccharides of *Campylobacter jejuni*. *J. Am. Chem. Soc.* *139*, 9463-9466.

A1.1 Abstract

Bacterial capsular polysaccharides (CPS) are complex carbohydrate structures that play a role in the overall fitness of the organism. *Campylobacter jejuni*, known for being a major cause of bacterial gastroenteritis worldwide, produces a CPS with a unique *O*-methyl phosphoramidate (MeOPN) modification on specific sugar residues. The formation of P-N bonds in nature is relatively rare, and the pathway for the assembly of the phosphoramidate moiety in the CPS of *C. jejuni* is unknown. In this investigation we discovered that the initial transformation in the biosynthetic pathway for the MeOPN modification of the CPS involves the direct phosphorylation of the amide nitrogen of L-glutamine with ATP by the catalytic activity of Cj1418. The other two products are AMP and inorganic phosphate. The L-glutamine-phosphate product was characterized using ^{31}P NMR spectroscopy and mass spectrometry. We suggest that this newly discovered enzyme be named L-glutamine kinase.

A1.2 Introduction

Campylobacter jejuni is a Gram-negative bacterium that causes foodborne gastroenteritis in humans worldwide (1). It is commonly found in chickens, and as a consequence contaminated poultry are a significant reservoir for human disease. Whereas infection with *C. jejuni* is typically self-limiting, in rare cases it can lead to the subsequent development of Guillain-Barré syndrome, a devastating acute polyneuropathy (2). Like many Gram-positive and Gram-negative organisms, *C. jejuni* produces capsular polysaccharides, which are composed of chains of sugars that form extensive layers surrounding the outer surface of the bacterium. In some cases, these chains can be composed of more than 200 sugars (3). The capsular polysaccharides, hereafter referred to as CPS, protect the organism from the environment and from complement-mediated phagocytosis and killing (4). It is now well documented that, in *C. jejuni*, the CPS is important for colonization and invasion of the host organism (5). More than 40 serological strains of *C. jejuni* have been identified, and each strain is likely to produce structural variations to the CPS (6, 7). These modifications are involved in a complex strategy for evasion of both bacteriophage predation and host defense systems (4, 8). In *C. jejuni* strain NCTC11168, a cluster of 35 genes has been identified as being responsible for the synthesis and export of the CPS (9).

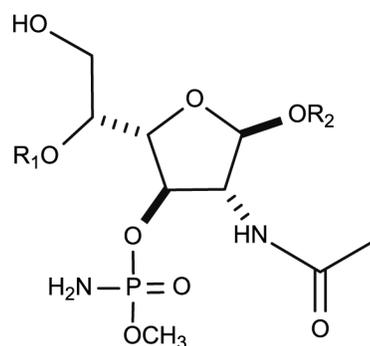
By far the most unusual modification to the CPS of *C. jejuni* is the addition of *O*-methyl phosphoramidate groups (MeOPN) attached to the polysaccharide backbone. For example, in *C. jejuni* strain NCTC11168, C3 of a 2-acetamido-2-deoxy- β -D-galactofuranose (**1**) moiety is decorated with an *O*-methyl phosphoramidate group, and the CPS of the hypermotile variant of this strain (11168H) has an additional MeOPN modification at C4 of a derivative of D-*glycero*- α -L-*gluco*-heptopyranose (**2**) as illustrated in Scheme A1.1 (6, 7). The occurrence of P–N bonds in biological systems is relatively rare (creatine phosphate and arginine phosphate are notable

exceptions), and the presence of the *O*-methyl phosphoramidate groups in the capsular polysaccharides of *C. jejuni* plays a significant role in its pathogenicity (5). In *C. jejuni* 11168H, genes with the locus tags *cj1418c*, *cj1417c*, *cj1416c*, and *cj1415c* have been implicated in the biosynthesis of the phosphoramidate moiety (3) of the CPS but the pathway leading to the formation of the P–N bond in this organism has not been elucidated (10).

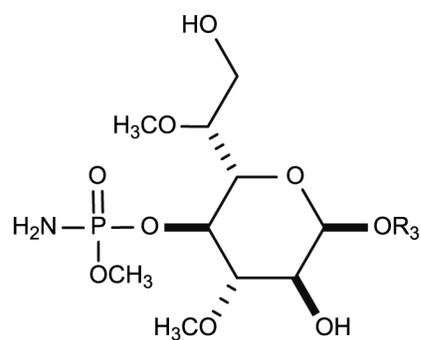
The focus of this investigation is on the catalytic functions of Cj1418 and Cj1417. Cj1418 is a member of cog0574, and this enzyme is currently annotated as a putative PEP synthase or a pyruvate phosphate dikinase (11, 12). Structurally characterized enzymes in this family are composed of three distinct protein segments, including an ATP-grasp domain, a PEP/pyruvate binding region, and a phosphohistidine domain (12). The amino acid sequence of Cj1418 suggests that it has an *N*-terminal ATP-grasp domain (residues 1–219) and a *C*-terminal phosphohistidine domain (residues 694–767). However, its central domain (residues 220–693) does not appear to be homologous to any of the known PEP/pyruvate binding regions. The closest structurally characterized homologue to Cj1418 with a known catalytic activity is rifampin phosphotransferase (23% sequence identity) from *Listeria monocytogenes* (PDB ID: 5FBS, 5FBT, and 5FBU) (13). This enzyme catalyzes the ATP-dependent phosphorylation of the antibiotic rifampin via a mechanism that involves the pyrophosphorylation of His-825, hydrolysis of this intermediate to generate a phosphorylated histidine intermediate, and subsequent phosphoryl transfer to rifampin (13, 14).

Cj1417 is a member of cog2071 and is annotated as a type I glutamine amidotransferase (11, 15). This class of enzymes catalyzes the hydrolysis of glutamine (or structurally similar glutamine analogs) via the formation of a thioester intermediate with an active site cysteine residue (15). In many cases, such as in carbamoyl phosphate synthetase, the hydrolysis of glutamine is coupled to

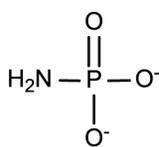
an ATP-dependent phosphorylation of a second substrate by an associated synthetase domain/subunit (16). We thus initially proposed that the combined activities of Cj1418 and Cj1417 would likely be required for the *in vivo* formation of the putative phosphoramidate intermediate (3) during the biosynthesis of the *O*-methyl phosphoramidate groups. In our proposed mechanism, the first reaction is initiated by the Cj1417 dependent hydrolysis of glutamine to form glutamate and ammonia. This step is subsequently followed by the phosphorylation of ammonia via the catalytic activity of Cj1418 (Scheme A1.2).



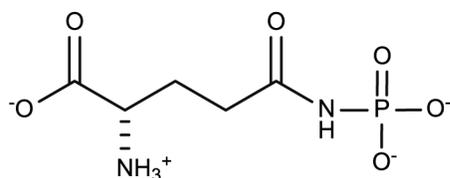
(1)



(2)

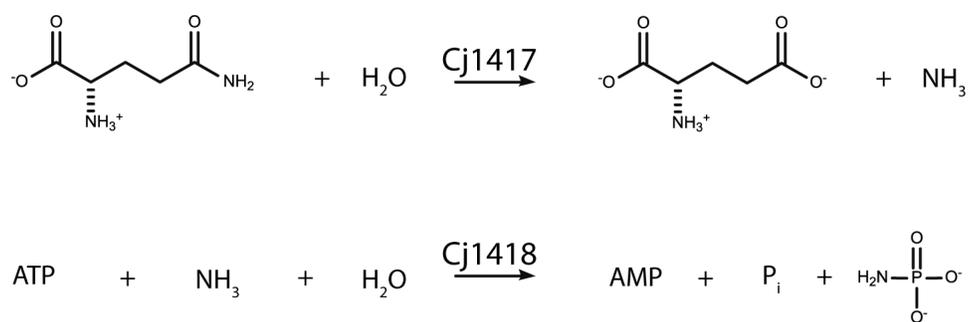


(3)



(4)

Scheme A1.1 Molecules relevant to the characterization of Cj1417 and Cj1418. (1) 2-acetamido-2-deoxy- β -D-galactofuranose with MeOPN on C3 (2) D-glycero- α -L-gluco-heptopyranose with MeOPN on C4 (3) phosphoramidate (4) L-glutamine phosphate



Scheme A1.2 Predicted functions of Cj1417 and Cj1418.

A1.3 Materials and Methods

A1.3.1 Protein expression and purification

The gene for the expression of Cj1418 with an *N*-terminal hexa-histidine purification tag was cloned from the genomic DNA of *C. jejuni* 11168H into a modified form of the pET-28b expression vector (17). This vector was subsequently used to transform Rosetta(DE3) *Escherichia coli*, and the cells were subsequently grown in a medium of lysogeny broth at 30 °C. Following induction with 1.0 mM isopropyl β -D-1-thiogalactopyranoside (IPTG), the cells were allowed to grow at 16 °C for 16 h. After cell lysis and centrifugation, Cj1418 was purified using Ni-affinity chromatography and the excess imidazole was removed by dialysis. The purified protein was concentrated and stored at -80 °C. Approximately 4 mg of Cj1418 were purified from 1.0 L of the original cell culture.

A1.3.2 Synthesis of L-glutamine phosphate (VII)

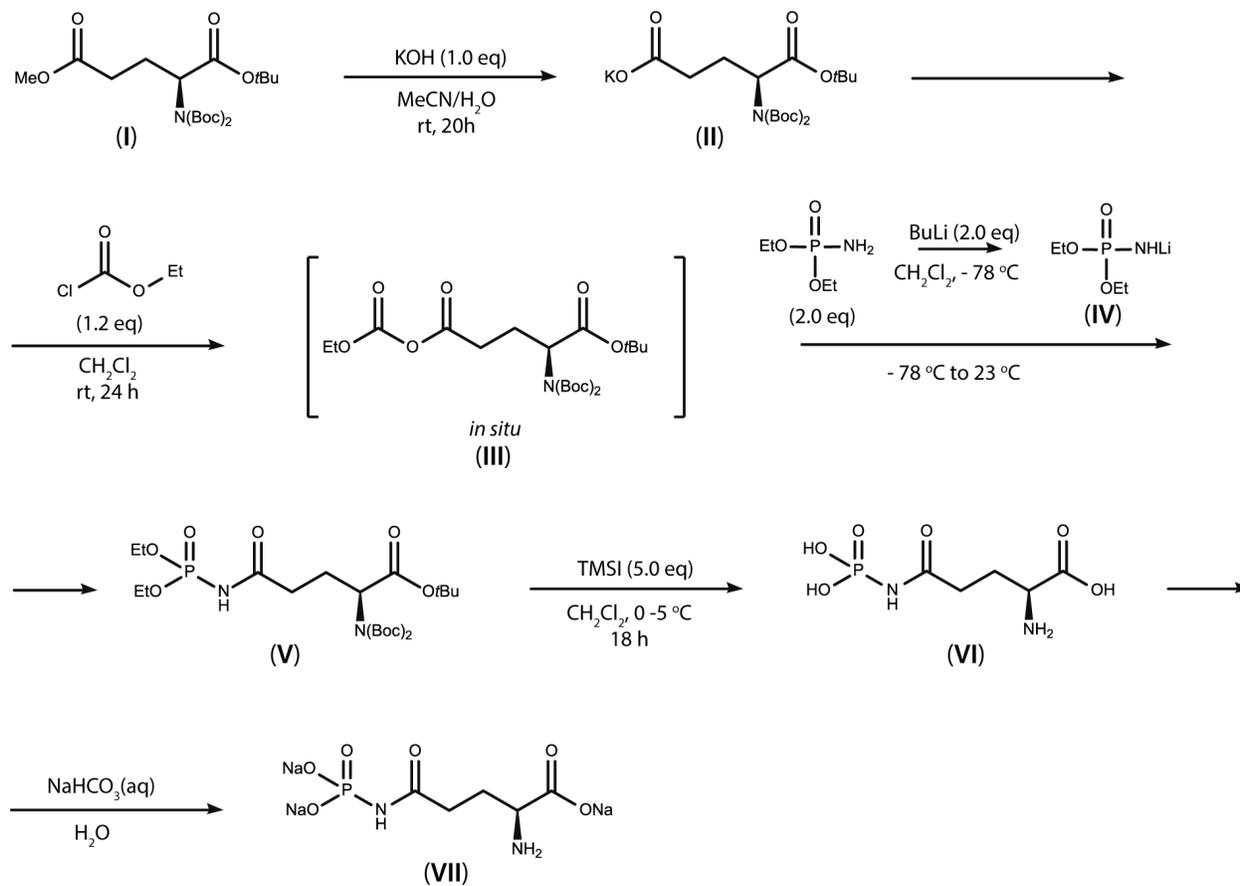
L-glutamine phosphate (VII) was synthesized by modification of a method described previously for the synthesis of L-asparagine phosphate (18, 19). The transformations are summarized in Scheme A1.3. Fully protected L-glutamic acid was synthesized as described previously (20). Compound I was quantitatively converted to the potassium salt (II) by the addition of a 5 mL solution of KOH (4.5 mmol, 1.0 equivalents) to a solution of compound I (4.5 mmol, 1.0 equivalents) in 5 mL of acetonitrile and the emulsion was stirred for 20 hours at 23 °C. The solvent was removed and the potassium salt (II) was dried under high vacuum for 24 to 48 hours. The potassium salt was converted into the mixed anhydride carbonate (III) by the addition of 1.2 equivalents of ethyl chloroformate in anhydrous dichloromethane for 24 hours at 23 °C and then cooled to -78 °C.

Diethyl phosphoramidate (2.0 equivalents) in anhydrous dichloromethane (6 mL) was treated with 2.5 M n-BuLi (2.0 equivalents) at -78 °C under argon. The suspension of the lithium salt of diethyl phosphoramidate (**IV**) was stirred for 15 minutes and then slowly added via syringe to the mixed anhydride carbonate (**III**) at -78 °C under argon. The reaction was allowed to stir until the temperature rose to 23 °C. The mixture was quenched with 5 mL of water and glacial acetic acid (1.5 mL) until mildly acidic. The organic layer was separated, and the water solution was extracted with ethyl acetate (2 x 15 mL). The combined organic solution was dried over MgSO₄ and then concentrated in vacuo. The crude oil was subjected to flash column chromatography (hexanes: ethyl acetate, 1:2) and the product (**V**) was obtained in 40 % yield. Removal of the protecting groups was performed with 5.0 equivalents of trimethylsilyl iodide in dichloromethane at 0-5 °C for 18 hours to obtain the free acid of L-glutamine phosphate (**VI**). Compound **VI** was dissolved in 5 mL of water and neutralized with a saturated solution of NaHCO₃. The solvent was evaporated under vacuum and the sodium salt (**VII**) was obtained as a colorless solid. The ¹H NMR spectrum of the free acid form of L-glutamine phosphate is presented in Figure A1.1 and the ³¹P NMR spectrum is presented in Figure A1.2. The ³¹P NMR spectrum of the sodium salt of L-glutamine phosphate is presented in Figure A1.3.

Compound **V**: ¹H NMR (CDCl₃, 400 MHz): δ = 8.32 - 8.10 (broad singlet, 1H), 4.79 - 4.75 (multiplet, 1H), 4.26 - 4.12 (multiplet, 4H), 2.48 - 2.39 (multiplet, 3H), 2.17 - 2.07 (multiplet, 1H), 1.50 (singlet, 18H), 1.44 (singlet, 9H), 1.34 ppm (doublet of triplets, J₁ = 7.2 Hz, J₂ = 2.4 Hz, 6H). ³¹P NMR (CDCl₃, 160 MHz): δ = -1.94 ppm; HRMS (ESI⁺): [M+Na]⁺ calculated for C₂₃H₄₃N₂O₁₀PNa: 561.2553, found: 561.2567.

Compound **VI**: ¹H NMR (D₂O, 400 MHz): δ = 4.10 (triplet, 1H, J = 6.42 Hz), 2.67 - 2.49 (broad multiplet, 2H), 2.33 - 3.14 (multiplet, 2H). ³¹P NMR (D₂O, 160 MHz): δ = -5.10 ppm.

Compound **VII**: (colorless solid). ^1H NMR and ^{31}P NMR spectra for the sodium salt of L-glutamine phosphate in D_2O indicated the existence of two species, which are likely due to *syn* and *anti* conformations of the amide bond. ^1H NMR (D_2O , 400 MHz): $\delta = 3.68 - 3.61$ (multiplet, 1H), 3.01 - 2.90 (multiplet, 0.5H), 2.70 - 2.59 (multiplet, 0.5H), 2.44 - 2.23 (multiplet, 1.5H), 2.18 - 1.93 ppm (multiplet, 1.5H). ^{31}P NMR (D_2O , 160 MHz): $\delta = -3.49$ and -3.69 ppm; HRMS (ESI $^-$): calculated for the (M-H) $^-$ species ($\text{C}_5\text{H}_{10}\text{N}_2\text{O}_6\text{P}$): 225.03, found: 225.02.



Scheme A1.3 Synthetic transformations for the synthesis of L-glutamine phosphate.

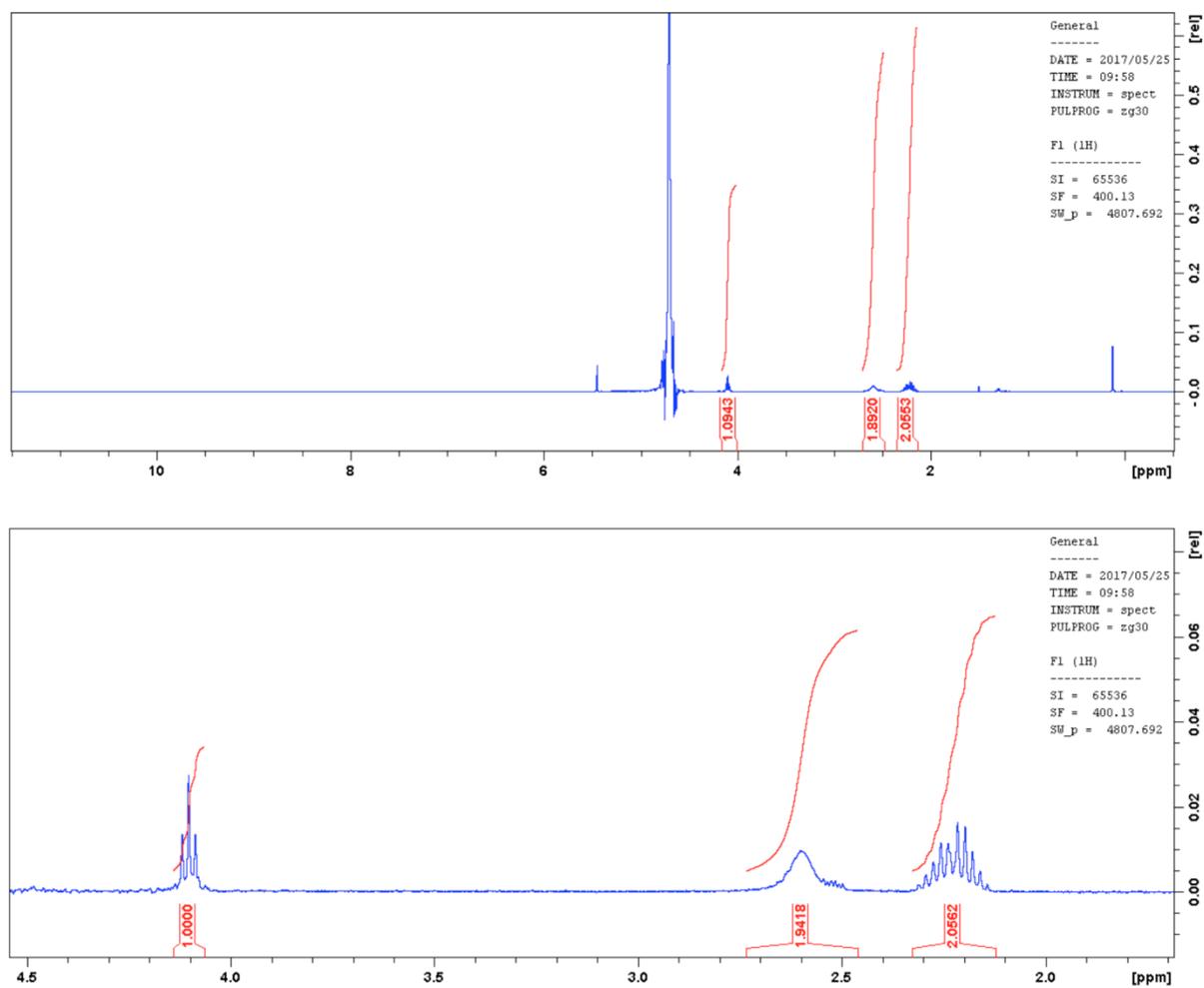


Figure A1.1 ^1H NMR spectrum of the free acid form of L-glutamine phosphate in D_2O .

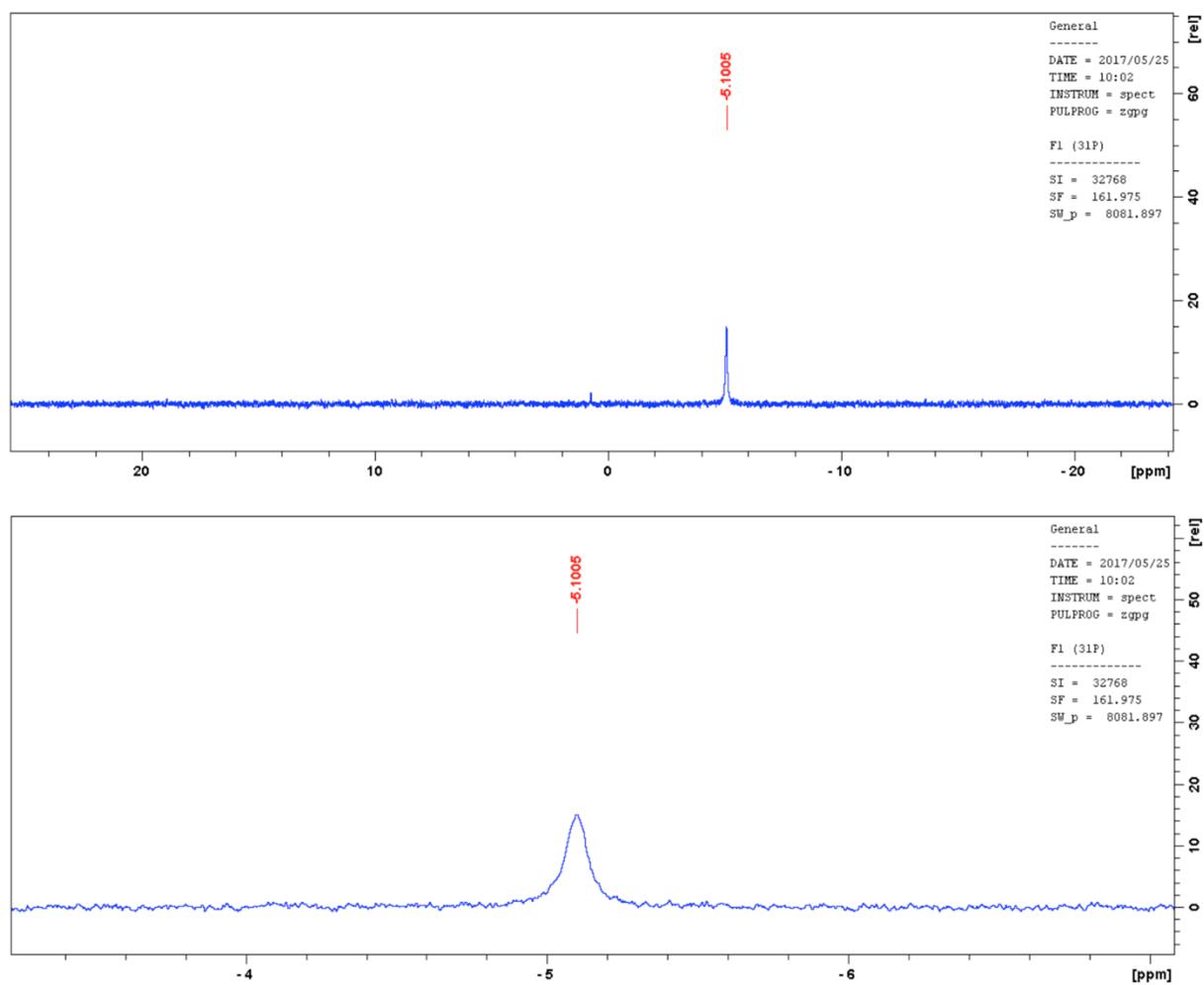


Figure A1.2 ^{31}P NMR spectrum of the free acid form of L-glutamine phosphate in D_2O .

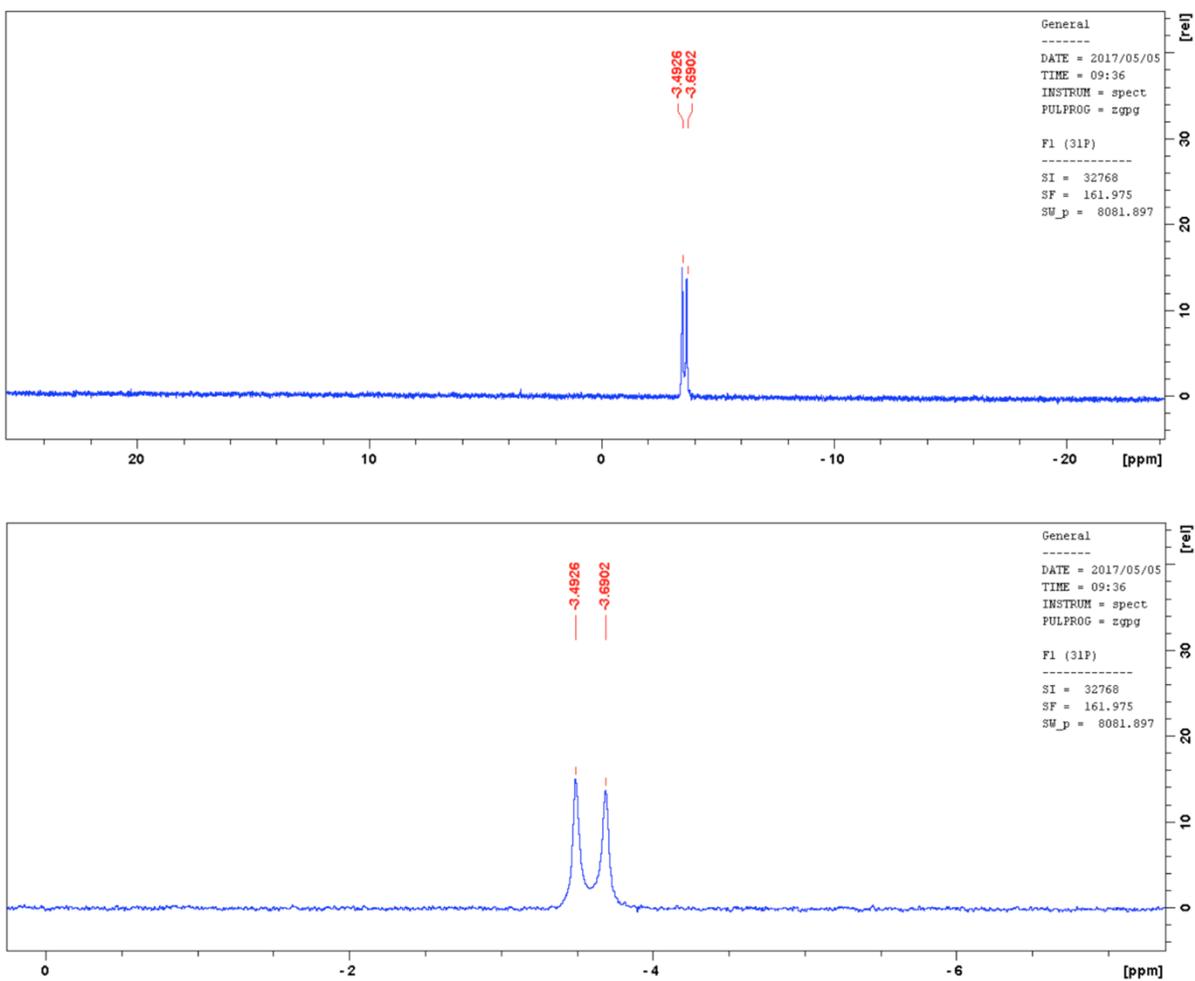


Figure A1.3 ^{31}P NMR spectrum of the sodium salt of L-glutamine phosphate in D_2O .

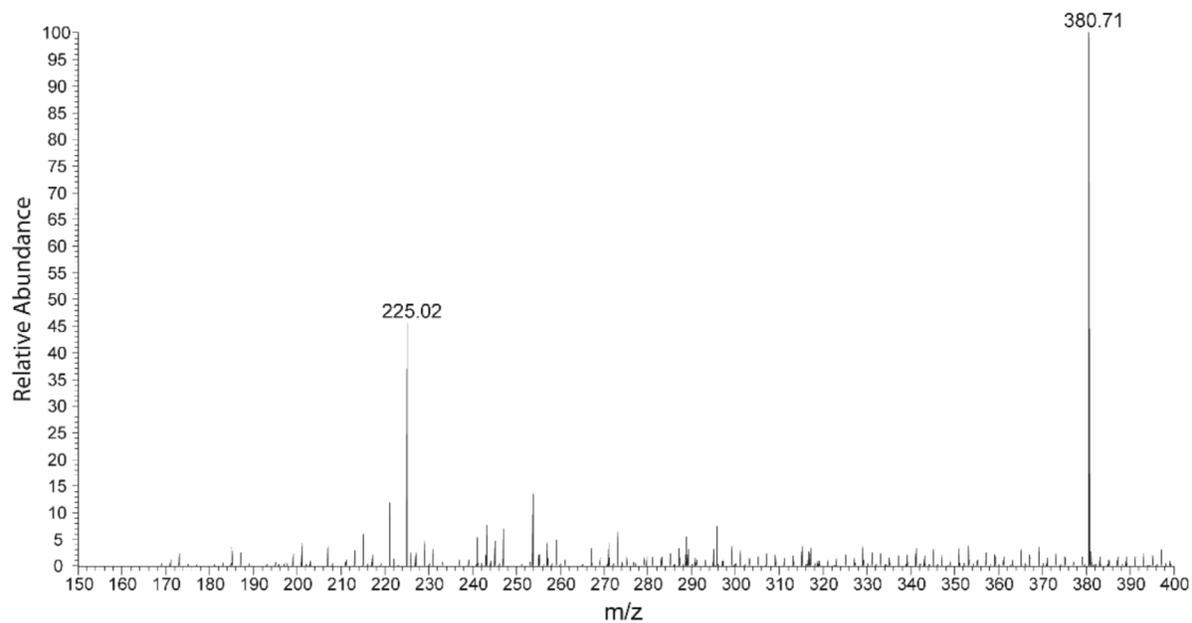


Figure A1.4 ESI negative ion mass spectrum of the sodium salt of L-glutamine phosphate prepared chemically. The peak at an m/z of 225.02 is for the $(M-H)^-$ ion ($C_5H_{10}N_2O_6P$) of L-glutamine phosphate. The peak at an m/z of 380.71 is for a contaminant of tri-iodide anion (I_3^-).

A1.4 Results and Discussion

To test our initial prediction that Cj1418 was required for the ATP-dependent phosphorylation of ammonia, we first incubated the enzyme (5.0 μ M) in the presence of 2.0 mM MgCl₂ and 1.0 mM ATP in 100 mM HEPES buffer (pH 8.0) at 30 °C. This control experiment was monitored using anion exchange chromatography by measuring the changes in the concentration of ATP at 255 nm. After an incubation period of ~60 min, the concentration of ATP (retention time of 8.2 min) did not change significantly but relatively small amounts of AMP (retention time of 5.3 min) and ADP (retention time of 7.1 min) could be detected [Fig. A1.5 (a)]. The addition of 100 mM NH₄Cl did not change the amounts of AMP or ADP that were produced [Fig. A1.5 (b)]. Since catalytic activity was not observed with ammonia, Cj1418 was next assayed in the presence of 5.0 mM L-glutamine. After an incubation period of ~60 min, all of the ATP was converted to AMP [Fig. A1.5 (c)]. The reaction mixture was subsequently examined by ³¹P NMR spectroscopy, and resonances were observed for AMP at 4.36 ppm and at 3.03 ppm for inorganic phosphate [Fig. A1.6 (a)]. Two additional resonances were observed at -3.57 and -4.06 ppm. Integration of the signal intensities for the sum of these two resonances equaled those observed for either AMP or P_i. The observed chemical shifts (-3.57 and -4.06 ppm) for the new phosphate containing product(s) did not match the ³¹P NMR spectrum for authentic phosphoramidate (**3**) at 1.3 ppm (21).

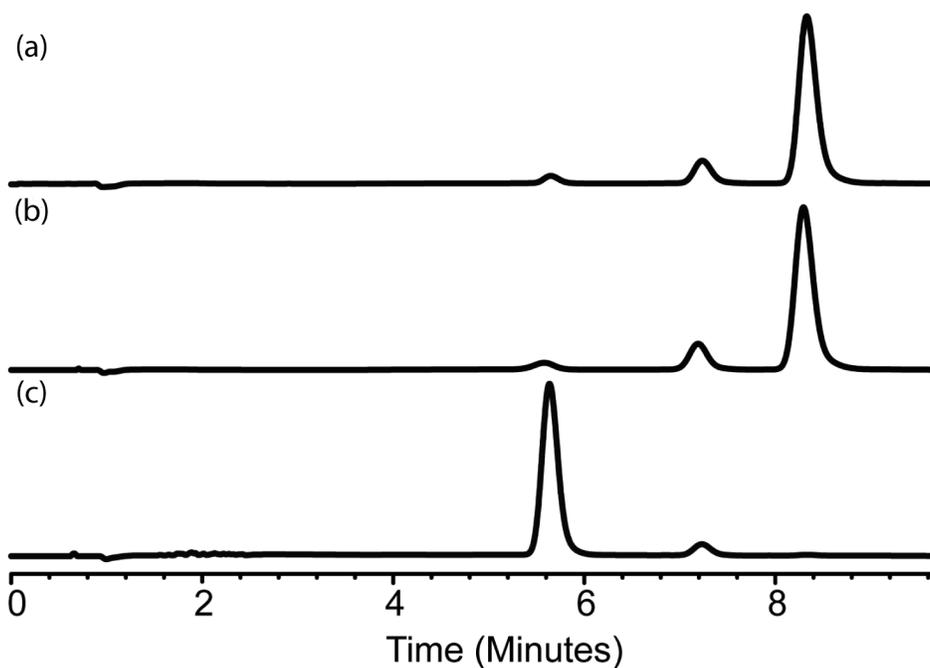


Figure A1.5 Anion exchange chromatograms of Cj1418 reaction products. Cj1418 (5.0 μM) and 100 mM HEPES buffer (pH 8.0) were incubated for 60 min at room temperature with (a) 1.0 mM ATP and 2.0 mM MgCl_2 . (b) 1.0 mM ATP, 2.0 mM MgCl_2 , and 100 mM NH_4Cl . (c) 1.0 mM ATP, 2.0 mM MgCl_2 , and 5.0 mM L-glutamine. Peak retention times correspond to the following: AMP (5.3 min), ADP (7.1 min), and ATP (8.2 min).

The most likely (but initially unexpected) product to form from the reaction catalyzed by Cj1418 is L-glutamine-phosphate (**4**) where the amide nitrogen is phosphorylated. To test this conjecture, the reaction was repeated using L-glutamine with an ^{15}N -label exclusively at the amide nitrogen. The two ^{31}P resonances of the reaction product now appear as doublets, due to the apparent spin coupling with the ^{15}N -labeled amide nitrogen [Fig. A1.6 (b)]. The observed coupling constants $J(^{15}\text{N}-^{31}\text{P})$ are 18 Hz for the phosphorus resonance at -3.57 ppm and 21 Hz for the resonance at -4.06 ppm. The magnitude of this coupling constant is consistent with that previously observed for phosphocreatine, which exhibits a $J(^{15}\text{N}-^{31}\text{P})$ coupling constant of 18–20 Hz (22). The most likely explanation for the observation of two distinct ^{31}P NMR signals for this compound is the restricted rotation of the amide functional group thereby giving rise to separate resonances for the *syn* and *anti* conformations of the L-glutamine-phosphate product. This conclusion is further supported by the direct chemical synthesis of L-glutamine phosphate (**18**). Two phosphorus resonances for the sodium salt of this compound are observed in D_2O at -3.5 and -3.7 ppm. A single resonance is observed at -5.10 ppm for the free acid where the rate of rotation about the amide bond is expected to increase. The chemical protocol for the synthesis of L-glutamine phosphate and the associated NMR (Fig. A1.1-A1.3) and mass spectra (Fig. A1.4) for the chemically synthesized compound are provided in the Materials and Method section.

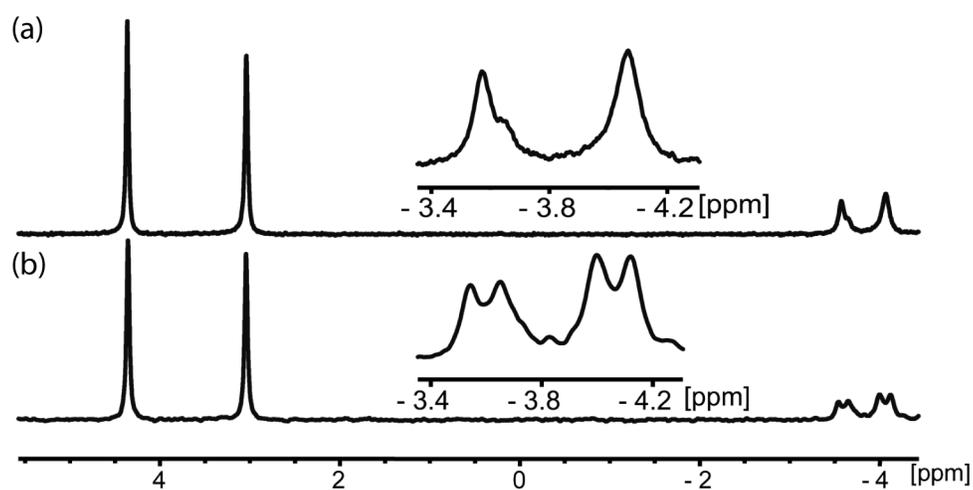


Figure A1.6 ^{31}P NMR spectra of Cj1418 reaction products. (a) ^{31}P NMR spectrum of the reaction products when Cj1418 was mixed with MgATP and L-glutamine. The resonance at 4.36 ppm is from AMP, and the resonance at 3.03 ppm is from inorganic phosphate. The resonances at -3.57 and -4.06 ppm correspond to L-glutamine phosphate (**4**). (b) ^{31}P NMR spectrum of the reaction products when Cj1418 was mixed with MgATP and L-glutamine-(amide- ^{15}N). The phosphorous resonances at -3.57 and -4.06 are now doublets due to the apparent spin coupling with the adjacent ^{15}N -nucleus.

The formation of L-glutamine-phosphate after incubation of Cj1418, ATP, and L-glutamine is further supported by the mass spectrum (ESI negative mode) of the unfractionated reaction mixture. A peak that corresponds to the mass of the expected L-glutamine phosphate is observed with an m/z of 225.03 for the $(M-H)^-$ species and at an m/z of 247.01 $(M-2H+Na)^-$ for the sodium adduct (Fig. A1.7). Several other major peaks are observed that correspond to the known compounds in the unfractionated reaction mixture including phosphate ($m/z = 96.96$), HEPES ($m/z = 237.09$), and AMP ($m/z = 346.05$). The full-width mass spectrum is presented in Figure A1.8.

The kinetic parameters for the phosphorylation of L-glutamine by ATP as catalyzed by Cj1418 at pH 8.0 and 30 °C were determined spectrophotometrically at 340 nm using a coupled enzyme assay that measures the formation of AMP. The assay contained adenylate kinase (8 units/mL), pyruvate kinase (8 units/mL), and lactate dehydrogenase (8 units/mL) in the presence of 11 mM $MgCl_2$, 0.40 mM NADH, and 2.0 mM PEP (23). Under these conditions the apparent kinetic constants for Cj1418 are $k_{cat} = 2.5 \pm 0.3 \text{ s}^{-1}$, $K_{ATP} = 340 \pm 70 \text{ }\mu\text{M}$, $k_{cat}/K_{ATP} = 7400 \pm 1700 \text{ M}^{-1} \text{ s}^{-1}$, $K_{Gln} = 640 \pm 60 \text{ }\mu\text{M}$, and $k_{cat}/K_{Gln} = 3900 \pm 800 \text{ M}^{-1} \text{ s}^{-1}$. No catalytic activity was observed (<1% of the rate with L-glutamine) in the presence of either L-glutamate (10 mM) or L-asparagine (10 mM).

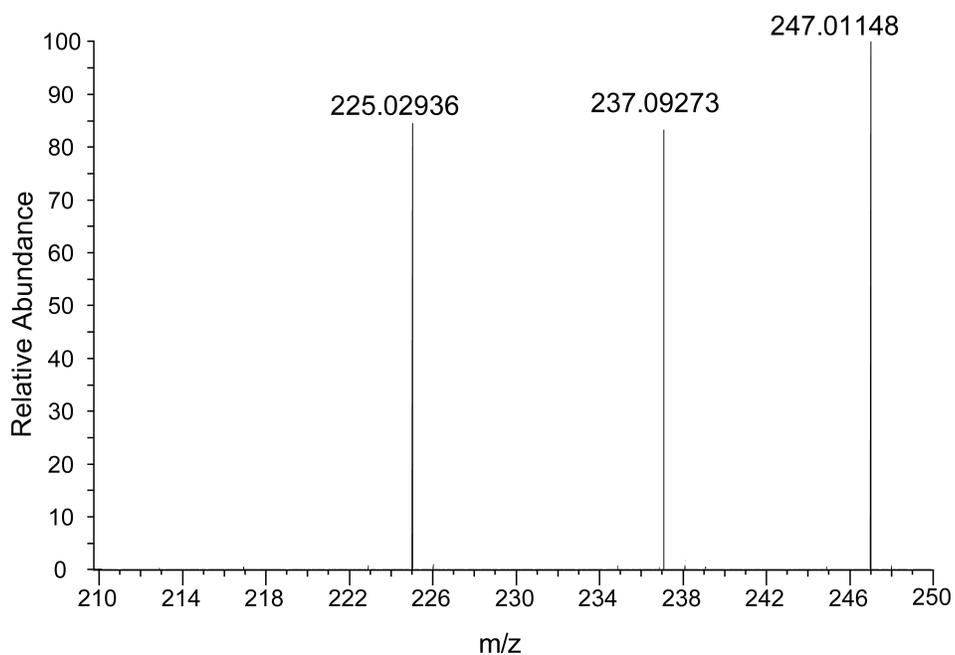


Figure A1.7 Mass spectrum of Cj1418 reaction mixture. Negative ESI mass spectrum of the reaction mixture when Cj1418 was mixed with 2.0 mM ATP and 5.0 mM L-glutamine at pH 8.0 in 100 mM sodium bicarbonate buffer (pH 8.0). The identified ions correspond to L-glutamine phosphate ($m/z = 225.03$ for M-H and $m/z = 247.01$ for M-2H+Na), and HEPES ($m/z = 237.09$ for M-H). The HEPES buffer was introduced with the preparation of Cj1418.

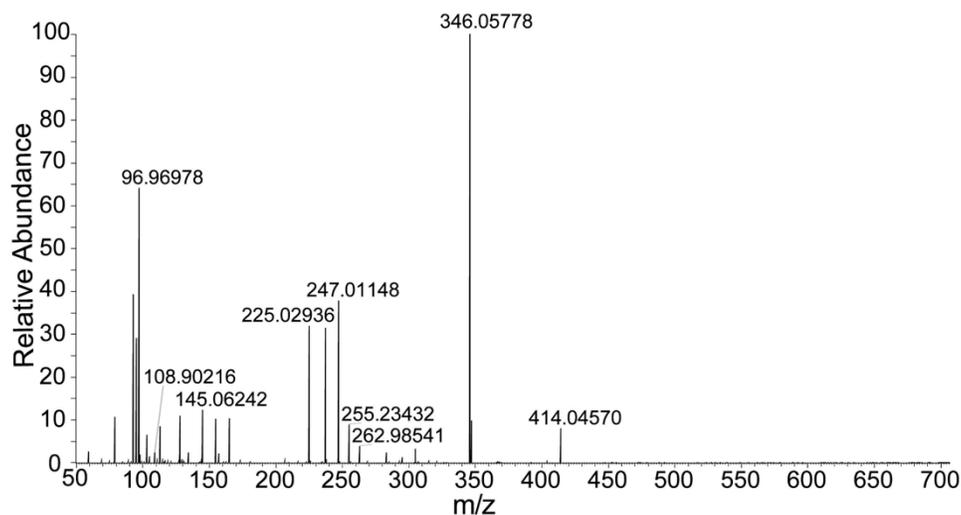
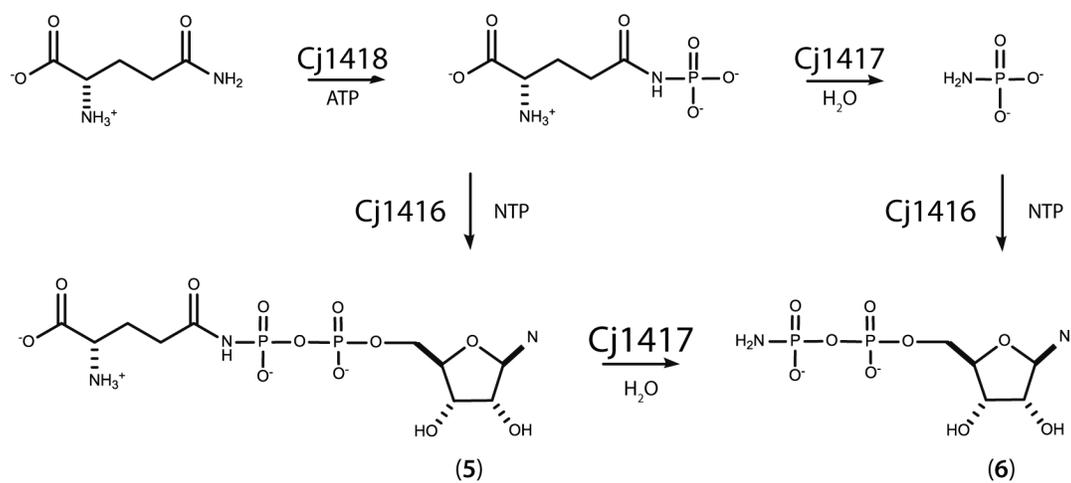


Figure A1.8 ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1418, ATP, and L-glutamine. Peaks that correspond to the mass of L-glutamine phosphate are observed with an m/z of 225.03 (M-H)⁻ species (C₅H₁₀O₆N₂P) and at an m/z of 247.01 (M-2H+Na)⁻ for the sodium adduct. Several other peaks are observed that correspond to the known compounds in the unfractionated reaction mixture including phosphate ($m/z = 96.96$), HEPES ($m/z = 237.09$), and AMP ($m/z = 346.05$).

Quite unexpectedly, we have shown that Cj1418, an enzyme involved in the biosynthesis of the *O*-methyl phosphoramidate groups in *C. jejuni*, catalyzes the phosphorylation of the amide nitrogen of L-glutamine, rather than ammonia. However, it has been shown previously that utilization of $^{15}\text{NH}_4\text{Cl}$ in the medium for growth of *C. jejuni* results in ^{15}N -labeling of the MeOPN groups in whole cells (24). Our current results suggest that the ammonia must first be transformed to L-glutamine, presumably by the action of L-glutamine synthetase. To the best of our knowledge our results represent the first documented case of an enzyme-catalyzed phosphorylation of a simple amide nitrogen. However, similar compounds have been chemically synthesized as potential inhibitors of D-alanine:D-alanine ligase (18, 25) and aspartate semialdehyde dehydrogenase (26). Glutamine synthetase has also been demonstrated to catalyze the phosphorylation of L-methionine-*S*-sulfoximine on nitrogen (27). The identity of L-glutamine phosphate was confirmed by ^{31}P NMR experiments, ^{15}N -labeling, and mass spectrometry. These results have further demonstrated that the initial series of steps as proposed in Scheme A1.2 for the biosynthesis of the *O*-methyl phosphoramidate capsule modification in *C. jejuni* is incorrect.

A more likely scenario for phosphoramidate biosynthesis is illustrated in Scheme A1.3. In this modified pathway, L-glutamine phosphate is hydrolyzed by Cj1417 to generate phosphoramidate (**3**). The closest functionally characterized homologue of Cj1417 is γ -L-glutamyl- γ -aminobutyrate hydrolase (PuuD) from *E. coli*. This protein has a 23% sequence identity with Cj1417, and thus homologous amidotransferase enzymes can catalyze the hydrolysis of substrates other than L-glutamine (28). In the next step we postulate that Cj1416 catalyzes the displacement of pyrophosphate by phosphoramidate (**3**) from a nucleotide triphosphate (NTP) to form the phosphoramidate of NDP (**5**). Cj1416 is a member of cog1213, and homologous enzymes have been shown to catalyze similar reactions. For example, CTP:phosphocholine

cytidylyltransferase from *Streptococcus pneumonia* (LicC) catalyzes the formation of CDP-choline from CTP and choline phosphate (29). Alternatively, Cj1416 may catalyze the formation of NDP-glutamine (6) through the displacement of pyrophosphate from NTP by L-glutamine-phosphate (4). The NDP phosphoramidate (5) would then be formed by the catalytic activity of Cj1417. Experiments are currently underway to firmly establish the catalytic activities of Cj1417, Cj1416, and the remaining transformations that lead to the biosynthesis of this fascinating modification to the capsular polysaccharides of *C. jejuni*.



Scheme A1.4 Experimentally determined function of Cj1418 and possible catalytic functions of Cj1417 and Cj1416.

A1.5 References

1. Young, K. T., Davis, L. M., and DiRita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* *5*, 665-679.
2. Jacobs, B. C., Rothbarth, P. H., van der Meché, F. G. A., Herbrink, P., Schmitz, P. I. M., de Klerk, M. A., and van Doorn, P. A. (1998) The spectrum of antecedent infections in Guillain-Barré syndrome: a case-control study. *Neurology.* *51*, 1110-1115.
3. Pelkonen, S., Hayrinen, J., and Finne, J. (1988) Polyacrylamide gel electrophoresis of the capsular polysaccharides of *Escherichia coli* K1 and other bacteria. *J. Bacter.* *170*, 2646-2653.
4. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., and Monteiro, M. A. (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines *Front. Cell Infect. Microbiol.* *2*, doi: 10.3389.
5. van Alphen, L. B., Wenzel, C. Q., Richards, M. R., Fodor, C., Ashmus, R. A., Stahl, M., Karlyshev, A. V., Wren, B. W., Stintzi, A., Miller, W. G., Lowary, T. L., and Szymanski, C. M. (2014) Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One.* *9*, e87051.
6. Karlyshev, A. V., Champion, O. L., Churcher, C., Brisson, J. R., Jarrell, H. C., Gilbert, M., Brochu, D., St. Michael, F., Li, J., Wakarchuk, W. W., Goodhead, I., Sanders, M., Stevens, K., White, B., Parkhill, J., Wren, B. W., and Szymanski, C. M. (2005) Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* *55*, 90-103.
7. St. Michael, F., Szymanski, C. M., Li, J., Chan, K. H., Khieu, N. H., Larocque, S., Wakarchuk, W. W., Brisson, J. R., and Monteiro, M. A. (2002) The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* *269*, 5119-5136.
8. Holst Sørensen, M. C., van Alphen, L. B., Harboe, A., Li, J., Christensen, B. B., Szymanski, C. M., and Brønsted, L. (2011) Bacteriophage F336 recognizes the capsular phosphoramidate modification of *Campylobacter jejuni* NCTC11168. *J. Bacteriol.* *193*, 6742-6749.
9. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J., and Wren, B. W. (2000) Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol. Microbiol.* *35*, 529-541.
10. McNally, D. J., Lamoureux, M. P., Karlyshev, A. V., Fiori, L. M., Li, J., Thacker, G., Coleman, R. A., Khieu, N. H., Wren, B. W., Brisson, J. R., Jarrell, H. C., and Szymanski, C. M. (2007) Commonality and biosynthesis of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* *282*, 28566-28576.

11. Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H. Y., Dosztanyi, Z., El Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D. A., Necci, M., Nuka, G., Orengo, C. A., Park, Y., Pesseat, S., Piovesan, D., Potter, S. C., Rawlings, N. D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Xenarios, I., yeh, L. S., Young, S. Y., and Mitchell, A. L. (2017) InterPro in 2017 - beyond protein family and domain notations. *Nucleic Acids Res.* *45*, D190-D199.
12. Herzberg, O., Chen, C. C. H., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano, D. (1996) Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites. *Proc. Natl. Acad. Sci. USA.* *93*, 2652-2657.
13. Stogios, P. J., Cox, G., Spanogiannopoulos, P., Pillon, M. C., Waglechner, N., Skarina, T., Koteva, K., Guarne, A., Savchenko, A., and Wright, G. D. (2016) Rifampin phosphotransferase is an unusual antibiotic resistance kinase. *Nat. Commun.* *7*, 11343.
14. Qi, X., Lin, W., Ma, M., Wang, C., He, Y., He, N., Gao, J., Zhou, H., Xiao, Y., Wang, Y., and Zhang, P. (2016) Structural basis of rifampin inactivation by rifampin phosphotransferase. *Proc. Natl. Acad. Sci. USA.* *113*, 3803-3808.
15. Massière, F. and Badet-Denisot, M. A. (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol. Life Sci.* *54*, 205-222.
16. Thoden, J. B., Holden, H. M., Wesenberg, G., Raushel, F. M., and Rayment, I. (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product. *Biochemistry.* *36*, 6305-6316.
17. Thoden, J. B. and Holden, H. M. (2005) The molecular architecture of human *N*-acetylgalactosamine kinase. *J. Biol. Chem.* *280*, 32784-32791.
18. Adams, L. A., Cox, R. J., Gibson, J. S., Mayo-Martin, M. B., Walter, M., and Whittingham, W. (2002) A new synthesis of phosphoramidates: inhibitors of the key bacterial enzyme aspartate semi-aldehyde dehydrogenase. *Chem. Commun. (Camb).* *18*, 2004-2005.
19. Grimes, K. D., Lu, Y. J., Zhang, Y. M., Luna, V. A., Hurdle, J. G., Carson, E. I., Qi, J., Kudrimoti, S., Rock, C. O., and Lee, R. E. (2008) Novel acyl phosphate mimics that target PlsY, an essential acyltransferase in gram-positive bacteria. *ChemMedChem.* *3*, 1936-1945.
20. Buckingham, F., Kirjavainen, A. K., Forsback, S., Krzyczmonik, A., Keller, T., Newington, I. M., Glaser, M., Luthra, S. K., Solin, O., and Gouverneur, V. (2015) Organomediated enantioselective ¹⁸F fluorination for PET applications. *Angew. Chem. Int. Ed. Engl.* *54*, 13366-13369.

21. Watanabe, M., Sato, S., and Wakasugi, K. (1990) The hydrolytic property of imidodiphosphate in solid and in an aqueous medium. *Bull. Chem. Soc. Jpn.* *63*, 1243-1245.
22. Reddick, R. E. and Kenyon, G. L. (1987) Syntheses and NMR studies of specifically labeled [2-¹⁵N]phosphocreatine, [2-¹⁵N]creatinine, and related ¹⁵N-labeled compounds. *J. Am. Chem. Soc.* *109*, 4380-4387.
23. Blondin, C., Serina, L., Wiesmuller, L., Gilles, A. M., and Barzu, O. (1994) Improved spectrophotometric assay of nucleoside monophosphate kinase activity using the pyruvate kinase/lactate dehydrogenase coupling system. *Anal. Biochem.* *220*, 219-221.
24. McNally, D. J., Lamoureaux, M., Li, J., Kelly, J., Brisson, J. R., Szymanski, C. M., and Jarrell, H. C. (2006) HR-MAS NMR studies of ¹⁵N-labeled cells confirm the structure of the *O*-methyl phosphoramidate CPS modification in *Campylobacter jejuni* and provide insight into its biosynthesis. *Can. J. Chem.* *84*, 676-684.
25. Parsons, W. H., Patchett, A. A., Bull, H. G., Schoen, W. R., Taub, D., Davidson, J., Combs, P. L., Springer, J. P., Gadebusch, H., Weissberger, B., Valiant, M. E., Mellin, T. N., and Busch, R. D. (1988) Phosphinic acid inhibitors of D-alanyl-D-alanine ligase. *J. Med. Chem.* *31*, 1772-1778.
26. Cox, R. J., Gibson, J. S., and Mayo Martin, M. B. (2002) Aspartyl phosphonates and phosphoramidates: the first synthetic inhibitors of bacterial aspartate-semialdehyde dehydrogenase. *ChemBioChem.* *3*, 874-886.
27. Ronzio, R. A. and Meister, A. (1968) Phosphorylation of methionine sulfoximine by glutamine synthetase. *Proc. Natl. Acad. Sci. USA.* *59*, 164-170.
28. Kurihara, S., Oda, S., Kato, K., Kim, H. G., Koyanagi, T., Kumagai, H., and Suzuki, H. (2005) A novel putrescine utilization pathway involves γ -glutamylated intermediates of *Escherichia coli* K-12. *J. Biol. Chem.* *280*, 4602-4608.
29. Kwak, B. Y., Zhang, Y. M., Yun, M., Heath, R. J., Rock, C. O., Jackowski, S., and Park, H. W. (2002) Structure and mechanism of CTP:phosphocholine cytidyltransferase (LicC) from *Streptococcus pneumoniae*. *J. Biol. Chem.* *277*, 4343-4350.

Appendix 2:

Biosynthesis of Nucleoside Diphosphoramidates in *Campylobacter jejuni*

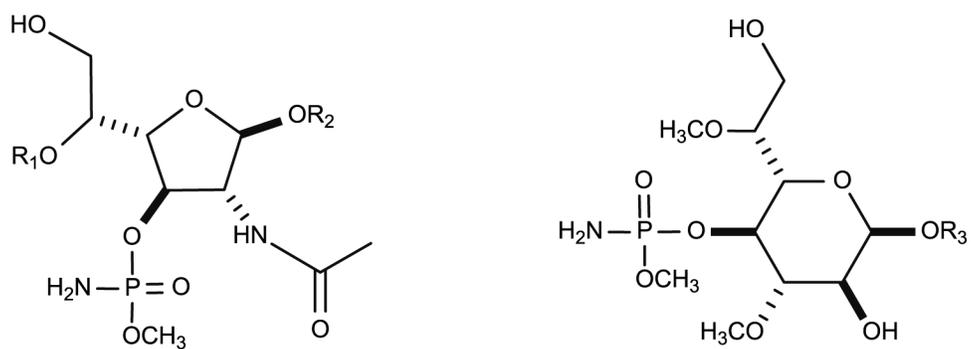
This work was previously published as: Taylor, Z. W., Brown, H. A., Holden, H. M., and Raushel, F. M. (2017) Biosynthesis of Nucleoside Diphosphoramidates in *Campylobacter jejuni*. *Biochemistry*. 56, 6079-6082.

A2.1 Abstract

Campylobacter jejuni is a pathogenic Gram-negative bacterium and a leading cause of food-borne gastroenteritis. *C. jejuni* produces a capsular polysaccharide (CPS) that contains a unique *O*-methyl phosphoramidate modification (MeOPN). Recently, the first step in the biosynthetic pathway for the assembly of the MeOPN modification to the CPS was elucidated. It was shown that the enzyme Cj1418 catalyzes the phosphorylation of the amide nitrogen of L-glutamine to form L-glutamine phosphate. In this investigation, the metabolic fate of L-glutamine phosphate was determined. The enzyme Cj1416 catalyzes the displacement of pyrophosphate from MgCTP by L-glutamine to form CDP-L-glutamine. The enzyme Cj1417 subsequently catalyzes the hydrolysis of CDP-L-glutamine to generate cytidine diphosphoramidate and L-glutamate. The structures of the two novel intermediates, CDP-L-glutamine and cytidine diphosphoramidate, were confirmed by ³¹P nuclear magnetic resonance spectroscopy and mass spectrometry. It is proposed that the enzyme Cj1416 be named CTP:phosphoglutamine cytidylyltransferase and that the enzyme Cj1417 be named γ -glutamyl-CDP-amidate hydrolase.

A2.2 Introduction

The pathogenic Gram-negative bacterium *Campylobacter jejuni* is a leading cause of food-borne gastroenteritis (1). While pathogenic to humans, *C. jejuni* is a commensal organism in chickens, and as a result, contaminated poultry serves as a common route of human infection. While most *C. jejuni* infections cause a case of gastroenteritis, approximately 1 in 1000 infections results in the autoimmune disease Guillain-Barré syndrome (2, 3). Like many other organisms, *C. jejuni* uses a capsular polysaccharide (CPS) to improve fitness. The capsular polysaccharide of *C. jejuni* protects the organism from bacteriophages and shields it from the host immune response (4, 5). More than 40 different strains of *C. jejuni* have been identified to date, and each strain is believed to produce a unique CPS variant (6, 7). In *C. jejuni* strain NCTC 11168, a cluster of approximately 35 genes is responsible for the synthesis and export of the CPS (8). Moreover, *C. jejuni* has evolved the ability to synthesize a unique *O*-methyl phosphoramidate (MeOPN) modification found on the CPS that improves the pathogenicity of the bacterium and promotes evasion of the host immune response (9). The structures of the MeOPN modification to the CPS in *C. jejuni* strain NCTC 11168 and hypermotile strain 11168H are illustrated in Scheme A2.1.

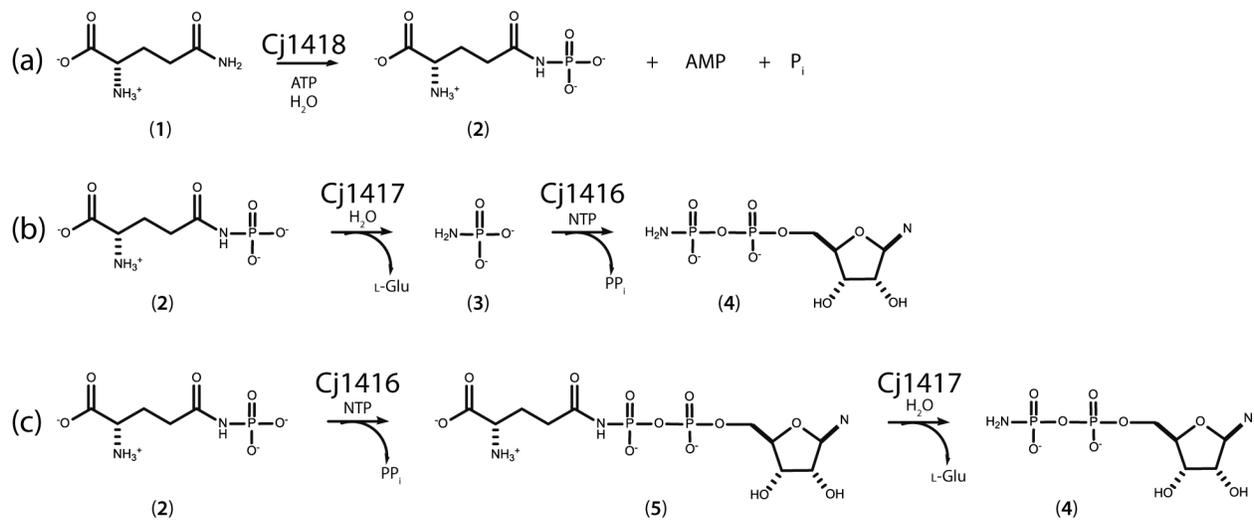


Scheme A2.1 Structures of the *O*-methyl phosphoramidate modifications to the CPS in *C. jejuni*

We have recently shown that the enzyme denoted with the locus tag Cj1418 from *C. jejuni* strain 11168H catalyzes the first committed step in the biosynthesis of the MeOPN (10). This enzyme, L-glutamine kinase, catalyzes the unprecedented ATP-dependent phosphorylation of the amide nitrogen of L-glutamine (**1**) to form L-glutamine phosphate (**2**) as shown in Scheme A2.2 (a). However, the subsequent metabolic fate of this novel enzyme intermediate has not been addressed. The primary focus of this investigation is to identify those enzymes that can harness the phosphoramidate moiety contained within L-glutamine phosphate for the ultimate construction of the *O*-methyl phosphoramidate modification of the CPS.

The two most likely enzymes of unknown function from *C. jejuni* to utilize L-glutamine phosphate as a substrate during the biosynthesis of the *O*-methyl phosphoramidate modification of the CPS are Cj1417 and Cj1416. Cj1417 is functionally annotated as a Type I amidotransferase from cog2071. Members of the Type I amidotransferase family of enzymes typically catalyze the hydrolysis of L-glutamine or amide-substituted derivatives of this amino acid (11, 12). Currently, the closest functionally characterized homologue to Cj1417 is the enzyme PuuD (23% identical sequence) from *Escherichia coli*, an enzyme that catalyzes the hydrolysis of 4-(γ -L-glutamylamino)butanoate to 4-aminobutanoate and L-glutamate (12). Cj1416 is currently annotated as a nucleotidyltransferase from cog1213. The closest functionally characterized homologue of Cj1416 is CTP:phosphocholine cytidyltransferase from *Streptococcus pneumoniae* with a sequence identity of 28% (13). This enzyme catalyzes the formation of CDP-choline and pyrophosphate from CTP and choline phosphate. We therefore predict that the combination of Cj1417 and Cj1416 will catalyze the synthesis of a nucleoside diphosphoramidate (**4**) using L-glutamine phosphate (**2**) and a nucleoside triphosphate as substrates.

The biosynthesis of a nucleoside diphosphoramidate (**4**) by the consecutive reactions catalyzed by Cj1417 and Cj1416 can be envisioned to occur via one of two possible reaction schemes. In the first scenario, Cj1417 catalyzes the hydrolysis of L-glutamine phosphate (**2**) to L-glutamate and phosphoramidate (**3**). This reaction is followed by the displacement of pyrophosphate from a nucleoside triphosphate by phosphoramidate (**3**) to generate the nucleoside diphosphoramidate (**4**) in a reaction catalyzed by Cj1416 as illustrated in Scheme A2.2 (b). Alternatively, Cj1416 catalyzes the displacement of pyrophosphate from a nucleoside triphosphate by L-glutamine phosphate to form NDP-L-glutamine (**5**). This reaction is subsequently followed by the hydrolysis of this intermediate by Cj1417 to form L-glutamate and the nucleoside diphosphoramidate (**4**) as presented in Scheme A2.2 (c).



Scheme A2.2 Reaction catalyzed by Cj1418 and the predicted functions of Cj1417 and Cj1416.

A2.3 Materials and Methods

A2.3.1 Gene expression and enzyme purification

The genes used for the expression of Cj1416 and Cj1417 were cloned from the genomic DNA of *C. jejuni* NCTC11168 into a pet-30b vector with a C-terminal hexahistidine tag. The expression plasmids for Cj1416 and Cj1417 were used to transform Rosetta(DE3) *E. coli* cells. Cells containing Cj1416 were grown in LB at 30 °C and were then induced with 1.0 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) and grown at 16 °C for 16 h. Cells containing Cj1417 were grown in TB in a 5.0 mL starter culture for 8 h at 37 °C, and then the 5.0 mL cultures were used to inoculate 1.0 L cultures. The 1.0 L cultures of Cj1417 were allowed to grow for 16 h at 25 °C. The enzymes Cj1416 and Cj1417 were purified using Ni-affinity chromatography, and excess imidazole was removed by dialysis. The purified proteins were concentrated and stored at -80 °C. Cj1416 yielded 25 mg of purified protein per liter of cell culture, and Cj1417 yielded 4 mg of protein per liter of cell culture.

A2.3.2 Enzyme assays and determination of kinetic constants

The kinetic constants for Cj1416 were determined at pH 8.0 in 100 mM HEPES/KOH and 100 mM KCl using anion exchange chromatography to monitor the change in concentration of the substrates and products as a function of time. The samples were first loaded onto the ion exchange column and then washed with 10 mM triethanolamine (pH 8.0). The products were eluted with a linear gradient of 10 mM triethanolamine (pH 8.0), and 2.0 M KCl. Each reaction was carried out at 25 °C, with a minimum of five time points collected for each concentration. The kinetic constants for Cj1417 were determined using a glutamate dehydrogenase (8 units/mL) coupled enzyme assay that monitors the reduction of NAD⁺ (0.5 mM) at 340 nm. These reactions were done using 100 mM HEPES/KOH (pH 8.0), 100 mM KCl, at 25 °C. The CDP-L-glutamine (**5**) for

these assays was synthesized using a mixture of Cj1416, 10 mM CTP, 10 mM L-glutamine phosphate, 12 mM MgCl₂, and 10 units/ml of inorganic pyrophosphatase. The reaction was allowed to go to completion, the enzymes were removed, and the CDP-L-glutamine was quantified by high-performance liquid chromatography and nuclear magnetic resonance spectroscopy.

A2.3.3 Chemical synthesis

L-Glutamine phosphate (**2**), phosphoramidate (**3**), and cytidine diphosphoramidate (**4**) were chemically synthesized as previously described (10, 14, 15).

A2.4 Results and Discussion

To test our initial prediction that the combination of Cj1416 and Cj1417 catalyzes the formation of a nucleoside diphosphoramidate, these two enzymes were incubated together in the presence of MgCTP and an excess of L-glutamine phosphate (**2**) at pH 8.0. After 45 min, all of the CTP {retention time of 8.7 min [Fig. A2.1 (a)]} was converted to a new product with a retention time of 5.9 min [Fig. A2.1 (b)]. The retention time of the new reaction product formed in the presence of Cj1417 and Cj1416 is identical to that of authentic cytidine diphosphoramidate [Fig. A2.1 (e)] (15).

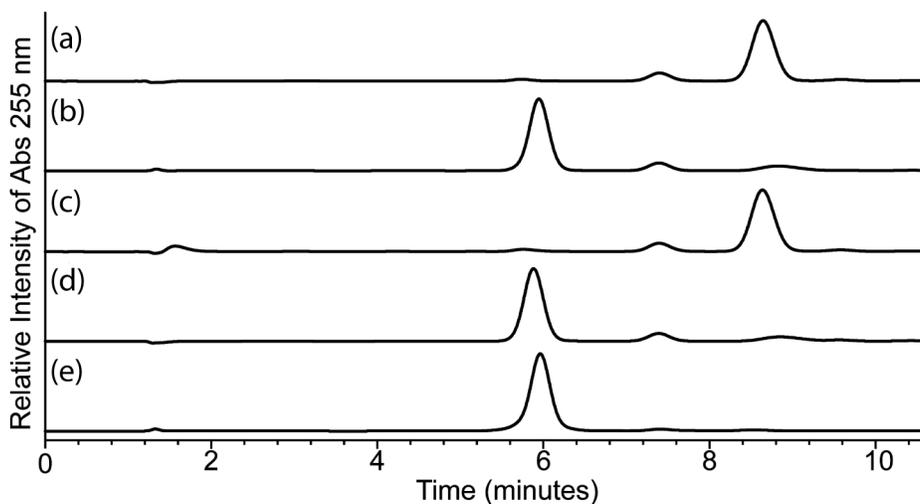


Figure A2.1 Anion exchange chromatograms of nucleotide standards and enzyme-catalyzed reaction products. Reactions were carried out in 100 mM HEPES (pH 8.0) at 25 °C with an incubation time of 45 min. The elution profiles were monitored at 255 nm. The nucleotides were separated using a 0 to 17% gradient of 10 mM triethanolamine (pH 8.0) and 2 M KCl over 17 column volumes on a 1 mL Resource Q column. (a) Control sample of 1.0 mM CTP and 2.0 mM MgCl₂ in the absence of any added enzyme. (b) Sample containing 1.0 mM CTP, 2.0 mM MgCl₂, 5.0 mM L-glutamine phosphate, Cj1416 (5 μM), and Cj 1417 (5 μM). (c) Sample containing 1.0 mM CTP, 2.0 mM MgCl₂, 5.0 mM phosphoramidate (**3**), 5 μM Cj1416, and 5 units/mL pyrophosphatase. (d) Sample containing 1.0 mM CTP, 2.0 mM MgCl₂, 5.0 mM L-glutamine phosphate, 5 μM Cj1416, and 5 units/mL pyrophosphatase. (e) Control sample of chemically synthesized CDP-phosphoramidate.

The identity of the new reaction product as cytidine diphosphoramidate {**4** [Scheme A2.2 (c)]} was confirmed by ^{31}P nuclear magnetic resonance (NMR) spectroscopy. A reaction mixture containing CTP, MgCl_2 , and an excess of L-glutamine phosphate (**2**) was incubated at pH 8.0 for 90 min in the presence of Cj1416 and Cj1417 until the reaction was quenched by the addition of 10 mM EDTA. The ^{31}P NMR spectrum of the control reaction [Fig. A2.2 (a)], obtained in the absence of Cj1416 and Cj1417, showed the expected resonances for CTP [-20.99 ppm ($\beta\text{-P}$), -10.33 ppm ($\alpha\text{-P}$), and -5.51 ppm ($\gamma\text{-P}$)] and L-glutamine phosphate (-3.57 and -3.83 ppm). In the presence of Cj1416 and Cj1417, the resonances for CTP and L-glutamine phosphate (**2**) essentially disappear and are replaced by new resonances for pyrophosphate (-6.63 ppm) and a pair of doublets at -0.42 ppm ($\beta\text{-P}$) and -10.23 ppm ($\alpha\text{-P}$) for cytidine diphosphoramidate [Fig. A2.2 (b)]. The ^{31}P NMR spectrum for authentic cytidine diphosphoramidate is presented in Figure A2.2 (c). The formation of cytidine diphosphoramidate was further supported by the acquisition of the negative ion electrospray ionization (ESI) mass spectrum of the unfractionated reaction mixture upon incubation of MgCTP, L-glutamine phosphate, Cj1417, and Cj1416 at pH 8.0 in ammonium bicarbonate buffer. A peak at m/z 401.03 was observed that is consistent with that expected for cytidine diphosphoramidate (Fig. A2.3). These experiments demonstrate that Cj1417 and Cj1416 use MgCTP and L-glutamine phosphate (**2**) to catalyze the formation of CDP-phosphoramidate, pyrophosphate, and L-glutamate.

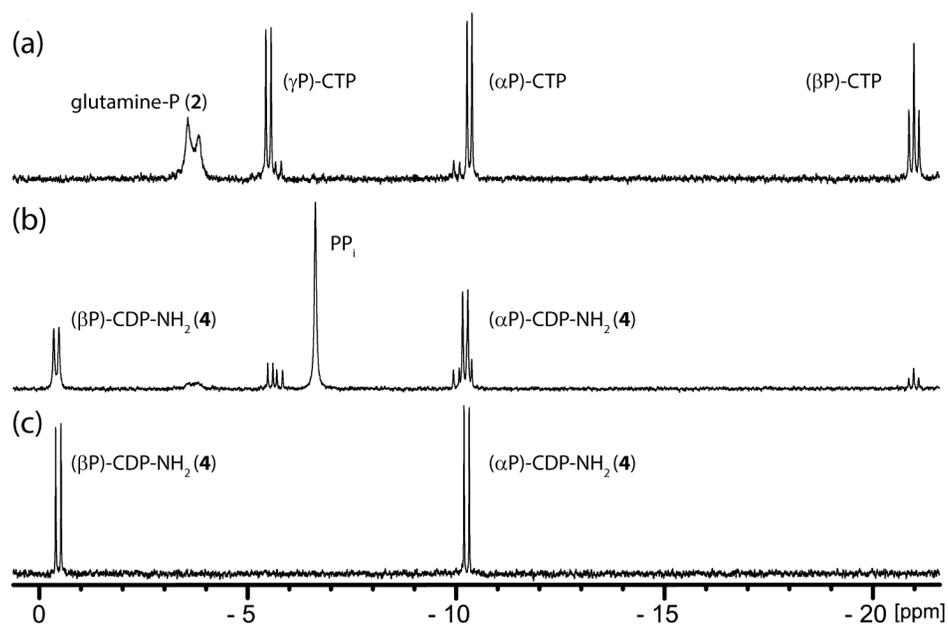


Figure A2.2 ^{31}P NMR spectra of nucleotide standards and enzyme-catalyzed reaction products. Reactions were carried out in 100 mM HEPES (pH 8.0) at 25 °C with an incubation time of 90 min before the reaction was quenched with 10 mM EDTA. (a) Control sample containing 5.0 mM CTP, 5.0 mM MgCl_2 , and 6.0 mM L-glutamine phosphate. (b) Sample containing 5.0 mM CTP, 5.0 mM MgCl_2 , 6.0 mM L-glutamine phosphate, 20 μM Cj1416, and 20 μM Cj1417. (c) Control sample of 5.0 mM CDP-phosphoramidate.

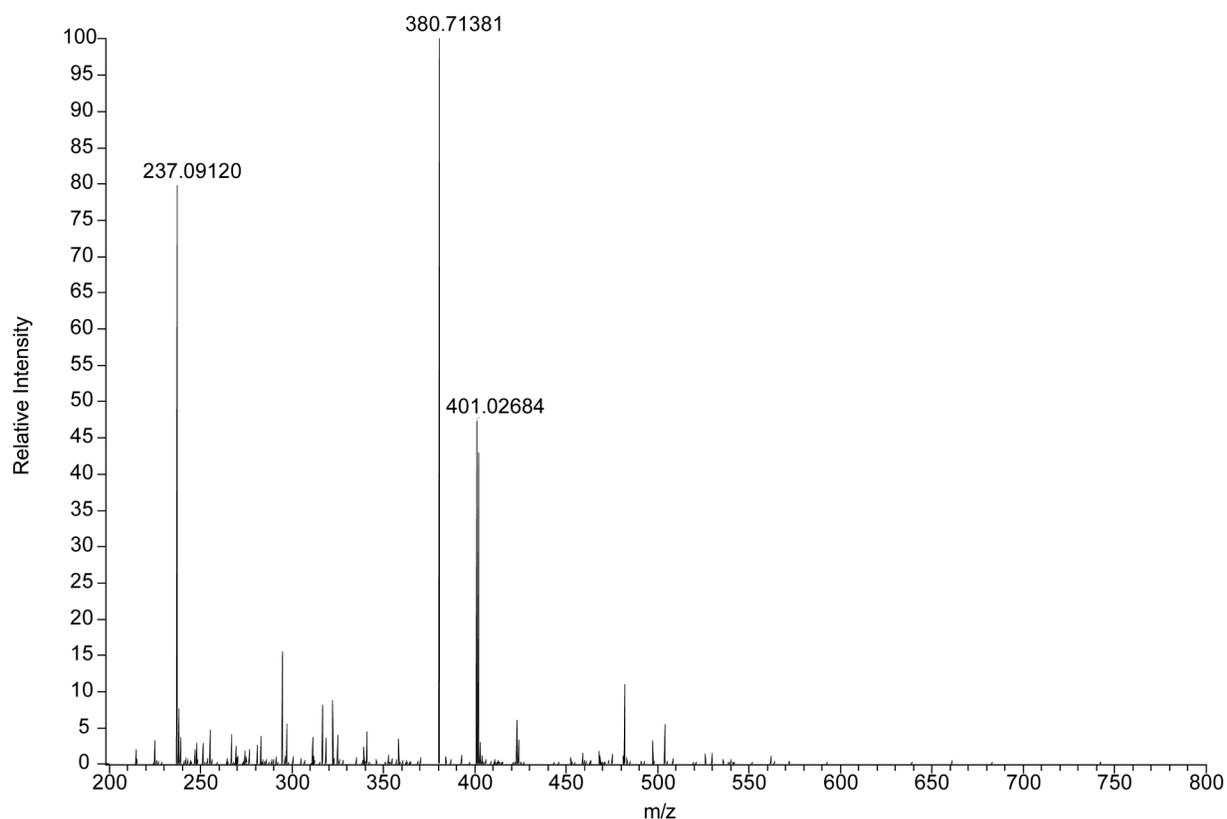


Figure A2.3 ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1416, Cj1417, MgCTP, and L-glutamine phosphate. The peak that corresponds to the mass of CDP-diphosphoramidate can be observed with an m/z of 401.03 for the $(M-H)^-$ species ($C_9H_{15}O_{10}N_4P_2$). Several other peaks are observed that correspond to known compounds in the unfractionated reaction mixture including HEPES ($m/z = 237.09$), and the peak at an m/z of 380.71 is a contaminant from the chemical synthesis of L-glutamine phosphate that corresponds to triiodide anion (I_3^-).

In parts (b) and (c) of Scheme A2.2, we have proposed that either phosphoramidate (**3**) or L-glutamine phosphate (**2**) is used to displace pyrophosphate from a nucleoside triphosphate to form either a nucleoside diphosphoramidate (**4**) or NDP-L-glutamine (**5**) as an intermediate. The reactivity of Cj1416 with each of these potential substrates was tested with MgCTP as the acceptor nucleotide, and the reaction was monitored by ion exchange chromatography at 255 nm. In separate experiments, either phosphoramidate (**3**) or L-glutamine phosphate (**2**) was incubated with CTP, MgCl₂, and Cj1416 at pH 8.0 for 45 min. Utilizing the chemically synthesized phosphoramidate (**3**) as a potential substrate, there was no change in the high-performance liquid chromatography chromatogram when compared to that of CTP alone [Fig. A2.1 (c)] (14). However, when Cj1416 was incubated with L-glutamine phosphate (**2**) and MgCTP, all of the CTP was converted to a new product that corresponds to a molecule with a net charge of approximately -2 [Fig. A2.1 (d)] (10). This result demonstrates that Cj1416 is fully capable of using L-glutamine phosphate (**2**) to displace pyrophosphate from CTP.

To provide further spectroscopic support for the Cj1416-catalyzed formation of CDP-L-glutamine, the reaction products were analyzed by ³¹P NMR spectroscopy. In this experiment, Cj1416 was incubated with CTP, MgCl₂, L-glutamine phosphate (**2**), and inorganic pyrophosphatase until the reaction was quenched with EDTA. After an incubation period of 90 min, essentially all of the CTP and L-glutamine phosphate (**2**) were converted to products [Fig. A2.4 (a)]. A new resonance is observed at 3.02 ppm for phosphate (from the hydrolysis of pyrophosphate), and two new doublets are observed at -10.98 ppm (α -P) and -16.13 ppm (β -P) for CDP-L-glutamine. The formation of CDP-L-glutamine was further supported by the acquisition of the negative ion ESI mass spectrum of the unfractionated reaction mixture upon incubation of

MgCTP, L-glutamine phosphate, and Cj1416 at pH 8.0 in ammonium bicarbonate buffer. A peak at m/z 530.07 was observed that is consistent with that expected of CDP-L-glutamine (Fig. A2.5).

When Cj1417 is subsequently added to the reaction mixture containing CDP-L-glutamine, the ^{31}P NMR resonances for CDP-L-glutamine disappear and are replaced by two new pairs of doublets at -0.43 ppm ($\beta\text{-P}$) and -10.24 ppm ($\alpha\text{-P}$) that can be assigned to cytidine diphosphoramidate [Fig. A2.4 (b) and Fig. A2.2 (c)]. These experiments demonstrate that Cj1416 catalyzes the formation of CDP-L-glutamine from CTP and L-glutamine phosphate (**2**) and that Cj1417 catalyzes the hydrolysis of CDP-L-glutamine to L-glutamate and cytidine diphosphoramidate.

The remaining issue to address for the functional characterization of Cj1417 is whether this enzyme is capable of catalyzing the hydrolysis of L-glutamine phosphate to L-glutamate and phosphoramidate (**2**). Cj1417 was therefore incubated with L-glutamine phosphate (**2**) at pH 8.0 for 90 min. The ^{31}P NMR spectrum [Fig. A2.4 (c)] of the reaction mixture demonstrated that $\sim 93\%$ of the original L-glutamine phosphate (**2**) remained intact. Two other resonances that are consistent with the presence of a small amount of phosphate (3.08 ppm, 1.2%) and phosphoramidate (1.21 ppm, 5.5%) are observed. The ^{31}P NMR spectrum for chemically synthesized phosphoramidate (**2**) is shown in Figure A2.4 (d), where a resonance is observed at 1.19 ppm. On the basis of these results, it is clear that the preferred pathway for the synthesis of cytidine diphosphoramidate (**4**) is for Cj1416 to catalyze the displacement of pyrophosphate from CTP to form CDP-L-glutamine (**5**) and for Cj1417 to catalyze the hydrolysis of this intermediate to generate cytidine diphosphoramidate (**4**) as shown in Scheme A2.2 (c).

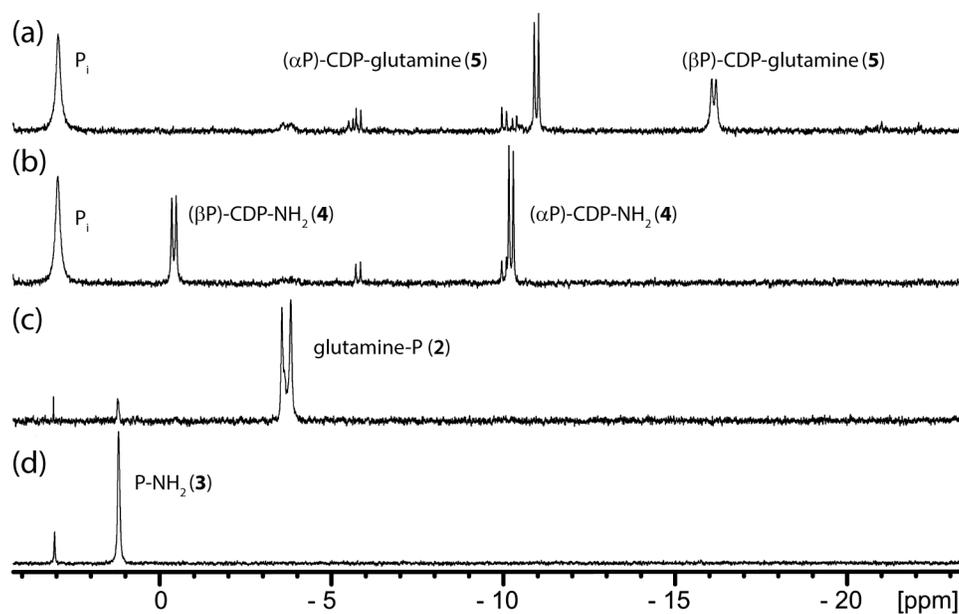


Figure A2.4 ^{31}P NMR spectra of nucleotide standards and enzyme-catalyzed reaction products. Reactions were carried out in 100 mM HEPES (pH 8.0) at 25 °C with an incubation time of 90 min before the reaction was quenched with 10 mM EDTA. (a) Sample containing 5.0 mM CTP, 5.0 mM L-glutamine phosphate, 5.0 mM MgCl_2 , 5 units/mL pyrophosphatase, and 20 μM Cj1416. (b) Cj1417 (20 μM) was added to the reaction mixture shown in panel (a) and the mixture was allowed to react for an additional 90 min. (c) Sample containing 20 μM Cj1417 and 5.0 mM L-glutamine phosphate. (d) Control sample of 5.0 mM phosphoramidate.

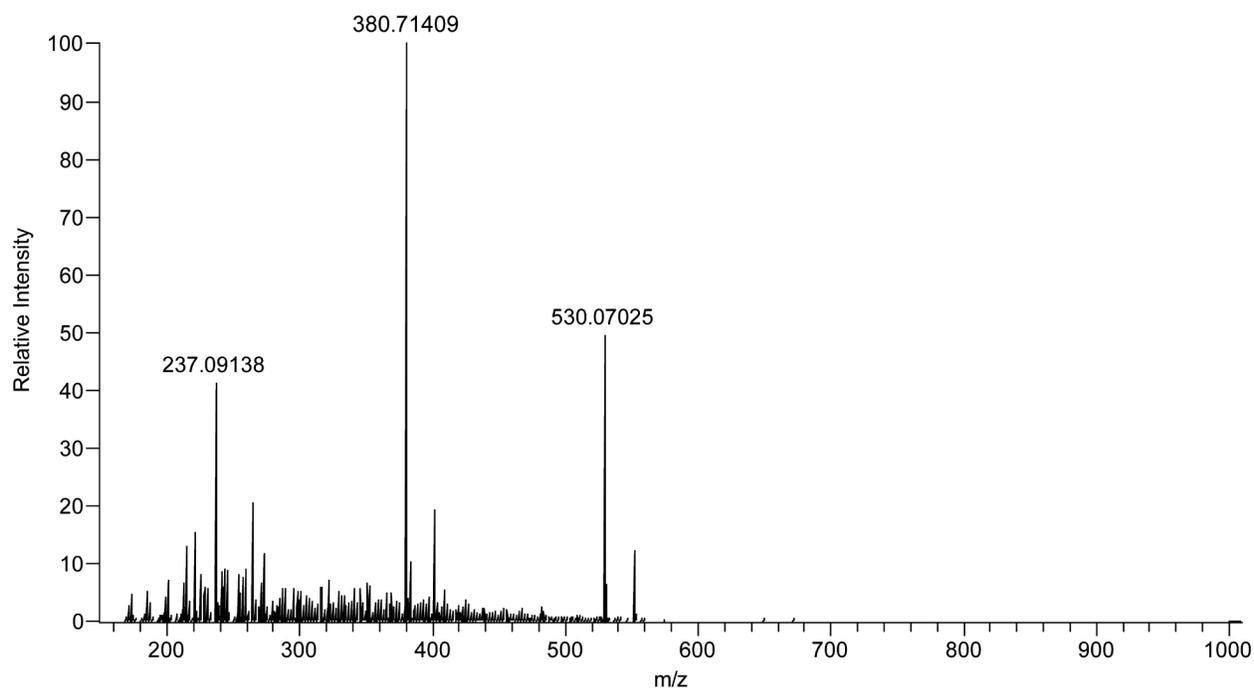


Figure A2.5 ESI negative ion mass spectrum of the unfractionated reaction mixture formed after incubation of Cj1416, MgCTP, and L-glutamine phosphate. The peak that corresponds to the mass of CDP-L-glutamine phosphate can be observed with an m/z of 530.07 for the $(M-H)^-$ species ($C_{14}H_{22}O_{13}N_5P_2$). Two other peaks are observed that correspond to the known compounds in the unfractionated reaction mixture, HEPES ($m/z = 237.09$) and triiodide anion from the synthesis of L-glutamine phosphate ($m/z = 380.71$).

The kinetic constants for the catalytic activity of Cj1417 and Cj1416 were determined. At a fixed concentration of either 5.0 mM MgCTP or 2.0 mM L-glutamine phosphate (**2**), the rates of the reaction catalyzed by Cj1416 were determined at 25 °C and pH 8.0 by monitoring the formation of products via anion exchange chromatography at 255 nm. The observed kinetic constants using L-glutamine phosphate (**2**) as the variable substrate are as follows: $K_M = 120 \pm 30 \mu\text{M}$, $k_{\text{cat}} = 57 \pm 6 \text{ min}^{-1}$, and $k_{\text{cat}}/K_M = (4.8 \pm 1.3) \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$. The observed kinetic constants using MgCTP as the variable substrate are as follows: $K_M = 170 \pm 35 \mu\text{M}$, $k_{\text{cat}} = 57 \pm 6 \text{ min}^{-1}$, and $k_{\text{cat}}/K_M = (3.4 \pm 0.8) \times 10^5 \text{ M}^{-1} \text{ min}^{-1}$.

The kinetic constants for the hydrolysis of CDP-L-glutamine (**5**) catalyzed by Cj1417 were determined at 25 °C and pH 8.0 using a glutamate dehydrogenase coupled assay that monitors the formation of NADH at 340 nm. The kinetic constants were determined to be as follows: $K_M = 28 \pm 3 \mu\text{M}$, $k_{\text{cat}} = 34 \pm 1.2 \text{ min}^{-1}$, and $k_{\text{cat}}/K_M = (1.2 \pm 0.2) \times 10^6 \text{ M}^{-1} \text{ min}^{-1}$. In an effort to determine the upper limit of the rate constant for the hydrolysis of L-glutamine phosphate (**2**) by Cj1417, 10 mM L-glutamine phosphate (**2**) was incubated with 20 μM Cj1417, and the reaction was monitored by ^{31}P NMR for 13 h. From this experiment, an upper limit of 1.6 h^{-1} was obtained for the hydrolysis of L-glutamine phosphate by Cj1417.

Previously, we have demonstrated that the first step in the biosynthesis of the MeOPN modification to the CPS of *C. jejuni* is catalyzed by Cj1418, where ATP is utilized to phosphorylate the amide nitrogen of L-glutamine (**10**). Here we have shown that Cj1416 catalyzes the displacement of pyrophosphate from MgCTP by L-glutamine phosphate (**2**), yielding CDP-L-glutamine (**5**). We have also established that the catalytic function of Cj1417 is to hydrolyze CDP-L-glutamine to L-glutamate and cytidine diphosphoramidate (**4**). This investigation has thus unveiled the identities of two new nucleoside diphosphoramidate derivatives that are involved in

the biosynthesis of MeOPN. It is likely that cytidine diphosphoramidate (**4**) will be subsequently phosphorylated at the hydroxyl group attached to C3 of the ribose ring prior to transfer of the phosphoramidate group to various carbohydrates of the CPS. On the basis of sequence similarity network analysis of the enzymes of unknown function contained within *C. jejuni*, we predict that Cj1415 will catalyze this reaction. Cj1415 is a member of cog0529, and the closest functionally characterized enzyme is CysC, an adenylyl-sulfate kinase, from *E. coli* (26% identity), an enzyme that catalyzes the ATP-dependent phosphorylation of the 3'-hydroxyl of adenylyl sulfate (16). The unique biosynthetic pathway for the assembly of the phosphoramidate functionality found in the CPS of *C. jejuni* offers many opportunities for the development of potent inhibitors that may ultimately be useful in the therapeutic control of this pathogenic organism.

A2.5 References

1. Young, K. T., Davis, L. M., and DiRita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* *5*, 665-679.
2. Jacobs, B. C., Rothbarth, P. H., van der Meché, F. G. A., Herbrink, P., Schmitz, P. I. M., de Klerk, M. A., and van Doorn, P. A. (1998) The spectrum of antecedent infections in Guillain-Barré syndrome: a case-control study. *Neurology.* *51*, 1110-1115.
3. Acheson, D. and Allos, B. M. (2001) *Campylobacter jejuni* infections: update on emerging issues and trends. *Clin. Infect. Dis.* *32*, 1201-1206.
4. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., and Monteiro, M. A. (2012) *Campylobacter* polysaccharide capsules: virulence and vaccines *Front. Cell Infect. Microbiol.* *2*, doi: 10.3389.
5. Holst Sørensen, M. C., van Alphen, L. B., Harboe, A., Li, J., Christensen, B. B., Szymanski, C. M., and Brønsted, L. (2011) Bacteriophage F336 recognizes the capsular phosphoramidate modification of *Campylobacter jejuni* NCTC11168. *J. Bacteriol.* *193*, 6742-6749.
6. Karlyshev, A. V., Champion, O. L., Churcher, C., Brisson, J. R., Jarrell, H. C., Gilbert, M., Brochu, D., St. Michael, F., Li, J., Wakarchuk, W. W., Goodhead, I., Sanders, M., Stevens, K., White, B., Parkhill, J., Wren, B. W., and Szymanski, C. M. (2005) Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* *55*, 90-103.
7. St. Michael, F., Szymanski, C. M., Li, J., Chan, K. H., Khieu, N. H., Larocque, S., Wakarchuk, W. W., Brisson, J. R., and Monteiro, M. A. (2002) The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* *269*, 5119-5136.
8. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J., and Wren, B. W. (2000) Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. *Mol. Microbiol.* *35*, 529-541.
9. van Alphen, L. B., Wenzel, C. Q., Richards, M. R., Fodor, C., Ashmus, R. A., Stahl, M., Karlyshev, A. V., Wren, B. W., Stintzi, A., Miller, W. G., Lowary, T. L., and Szymanski, C. M. (2014) Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One.* *9*, e87051.
10. Taylor, Z. W., Brown, H. A., Narindoshvili, T., Wenzel, C. Q., Szymanski, C. M., Holden, H. M., and Raushel, F. M. (2017) Discovery of a glutamine kinase required for the biosynthesis of the *O*-methyl phosphoramidate modifications found in the capsular polysaccharides of *Campylobacter jejuni*. *J. Am. Chem. Soc.* *139*, 9463-9466.
11. Massière, F. and Badet-Denisot, M. A. (1998) The mechanism of glutamine-dependent amidotransferases. *Cell Mol. Life Sci.* *54*, 205-222.

12. Kurihara, S., Oda, S., Kato, K., Kim, H. G., Koyanagi, T., Kumagai, H., and Suzuki, H. (2005) A novel putrescine utilization pathway involves γ -glutamylated intermediates of *Escherichia coli* K-12. *J. Biol. Chem.* *280*, 4602-4608.
13. Kwak, B. Y., Zhang, Y. M., Yun, M., Heath, R. J., Rock, C. O., Jackowski, S., and Park, H. W. (2002) Structure and mechanism of CTP:phosphocholine cytidyltransferase (LicC) from *Streptococcus pneumoniae*. *J. Biol. Chem.* *277*, 4343-4350.
14. Watanabe, M., Sato, S., and Wakasugi, K. (1990) The hydrolytic property of imidodiphosphate in solid and in an aqueous medium. *Bull. Chem. Soc. Jpn.* *63*, 1243-1245.
15. Wehrli, W. E., Verheyden, D. L. M., and Moffatt, J. G. (1965) Dismutation reactions of nucleoside polyphosphates. II. specific chemical synthesis of α -, β -, and γ - P^{32} -nucleoside 5'-triphosphates. *J. Am. Chem. Soc.* *87*, 2265-2277.
16. Leyh, T. S., Taylor, J. C., and Markham, G. D. (1988) The sulfate activation locus of *Escherichia coli* K12: cloning, genetic, and enzymatic characterization. *J. Biol. Chem.* *263*, 2409-2416.

A2.6 Acknowledgements

We thank Professor C. Szymanski and Mr. C. Wenzel for the plasmids that were used to express Cj1417 and Cj1416.