BIOSYNTHESIS OF THE GLUCURONAMIDE MOIETY OF THE CAPSULAR POLYSACCHARIDE OF *CAMPYLOBACTER JEJUNI*

A Dissertation

by

ALEXANDER STERLING RIEGERT

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Tatyana I. Igumenova
Committee Members,	David P. Barondeau
	Jennifer K. Herman
	Frank M. Raushel
Head of Department,	Josh Wand

December 2022

Major Subject: Biochemistry

ABSTRACT

Campylobacter jejuni is a pathogenic organism that can cause campylobacteriosis in children and adults. Most commonly, *Campylobacter* infection is brought on by consumption of raw or undercooked poultry, unsanitary drinking water, or pet feces. Surrounding the *C. jejuni* bacterium is a coat of sugar molecules known as the capsular polysaccharide (CPS). The capsular polysaccharide can be very diverse among the different strains of *C. jejuni*, and this diversity is considered important for evading the host immune system. Modifications to the CPS of *C. jejuni* NCTC 11168 include *O*-methylation, phosphoramidylation, and amidation of glucuronate with either serinol or ethanolamine. The enzymes responsible for amidation of glucuronate are currently unknown. We functionally characterized five enzymes that are proposed to be required for glucuronamide biosynthesis in *Campylobacter jejuni*. Cj1441 was shown to catalyze the oxidation of UDP-α-D-glucose to UDP-α-D-glucuronic acid with NAD^+ as the cofactor. The three-dimensional crystal structure of Cj1441 was determined in the presence of NAD^+ and UDP-glucose bound in the active site of the enzyme. The amines required for amidation were synthesized from two PLP containing enzymes, Cj1436 and Cj1437. Decarboxylation of L-serine phosphate to ethanolamine phosphate is catalyzed by Cj1436. Cj1437 catalyzed the stereoselective transamination of dihydroxyacetone phosphate to (*S*)-serinol phosphate. Cj1438 was shown to catalyze amide bond formation using ATP, glucuronic acid, and either serinol phosphate or ethanolamine phosphate. Finally, Cj1435 catalyzed the hydrolysis of the phosphodiester of the glucuronamide of

ethanolamine phosphate or the glucuronamide of (*S*)-serinol phosphate. This work represents the first known biosynthetic route for amide bond formation in capsular polysaccharides of *C. jejuni*.

DEDICATION

I dedicate this dissertation to my parents, my brother, and my wife. They've always believed in my dream of becoming a scientist and have given me endless support and love over the years. I could not have done this without them.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Frank Raushel, for his mentorship over the last five years. Dr. Raushel taught me how to be an effective scientist. His close eye for details and relentlessly curious mind are qualities I strive for. I've never met a more caring and understanding mentor. I also would like to thank my committee members, Dr. Tatyana I. Igumenova, Dr. David P. Barondeau, and Dr. Jennifer K. Herman, for their advice and help in bringing this project to completion. I was helped by so many wonderful people during my time at Texas A&M University and I would like to take some time to thank some of them. My friends and former lab-mates Dr. Zane Taylor, Dr. Tessily Hogancamp, and Dr. Drake Mellott were so generous with their time and ideas. They knew how to pick me up when I was low and push me to be a better scientist. Dr. Seth Cory was instrumental in the data collection for the X-ray crystal structure of Cj1441 and I am very grateful for his advice and assistance. Post-doctoral researchers are an integral part of scientific research and I had some truly special people helping me during my time in the Raushel lab and I'd like to thank them. Dr. Jamison Huddleston, Dr. DaoFeng Xiang, and Dr. Tamari Narindoshvili were always available to lend a hand or offer advice when I was stuck on a problem. There are often people behind the scenes that help our labs run smoothly and I would like to give a special thank-you to them as well. Andrea Scott was responsible for ordering supplies and coordinating conference travel, but most importantly she was always there to offer a smile and a piece of candy or two when you needed it. Ed Janusek is the ILSB building

manager and was always there when I needed a laugh or needed help fixing something. I also need to thank Dr. Hazel Holden and Dr. Jim Thoden for accepting me into their lab as an undergraduate at the University of Wisconsin-Madison. The time they spent to help train me was key to me having a successful transition to my graduate work. Finally, I would like to thank my family for their support over the last five years. Chuck, Val, Dan, Kelley, Oliver, Analicia, and Zelda; none of this would have been possible without your love and encouragement.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Frank M. Raushel, Professor Jennifer K. Herman, Professor Tatyana I. Igumenova of the Department of Biochemistry and Biophysics and Professor David P. Barondeau of the Department of Chemistry.

Dr. Tamari Narindoshvili was responsible for the chemical synthesis of numerous compounds used in this research including (*S*)-serinol phosphate, (*R*/*S*)-serinol phosphate, 1-*O*-methyl-β-D-glucuronate, *N*-acetyl ethanolamine phosphate, *N*-acetyl serinol phosphate, glucuronamide of *N*-acetyl serinol phosphate, *N*-acetyl ethanolamine phosphate, glucuronate (1→2)-1-*O*-methyl-α-D-ribose. Dr. Narindoshvili was assisted by Nicole Platzer.

Dr. Nigel Richards of Cardiff University and Dr. Adriana Coricello of the Dipartimento di Scienze della Salute, Università "Magna Græcia" di Catanzaro were able to calculate the kinetic constants for Cj1436 using membrane inlet mass spectrometry.

Mass spectrometry was performed by the Mass Spectrometry Laboratories in the Chemistry Department.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate student was supported by a teaching assistantship from Texas A&M University for 2017-2018 and a research assistantship from Texas A&M University and the Robert A. Welch Foundation under Grant Number A-840 for 2016-2022 and the National Institute of Health under grant number GM 122825 for 2018-2020 and GM 139428 for 2020-2022. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Robert A. Welch Foundation, National Institute of Health, and Texas A&M University.

NOMENCLATURE

αKG	Alpha-ketoglutarate
°C	Degrees Celsius
μm	Micrometer
μM	Micromolar
Å	angstrom
A. baumanii	*Acinetobacter baumannii*
ABC	ATP-binding-cassette
Ala	Alanine
Ara*f*	D-Arabinofuranose
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
ATCC	American Type Culture Collection
ATP	Adenosine 5’-triphosphate
BLAST	Basic Local Alignment Search Tool
C. jejuni	*Campylobacter jejuni*
C1	Carbon one
CDC	Centers for Disease Control and Prevention
CDP	Cytidine-5'-diphosphate
cm	Centimeter

CMP-Kdo	Cytidine monophosphate 3-deoxy-α-D-*manno*-octulosonate
CO_2	Carbon dioxide
CPS	Capsular polysaccharide
Cys	Cysteine
D_2O	Deuterium oxide
Da	Dalton
DHAP	Dihydroxyacetone phosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EFI-EST	Enzyme Function Initiative-Enzyme Similarity Tool
equiv	Equivalents
ESI	Electro spray ionization
FDA	Food and Drug Administration
GalA	D-Galacturonic acid
GalNAc	*N*-Acetyl-D-galactosamine
Gal*p*	D-Galactopyranose
GBS	Guillain-Barré syndrome
GDP	Guanosine 5'-diphosphate
GlcA	D-Glucuronic acid
GlcNAc	*N*-Acetyl-D-glucosamine
Glc*p*	D-Glucopyranose

Glu	Glutamate
Gly	Glycine
GNN	Genome Neighborhood Network
h	Hour
H. influenzae	*Haemophilus influenzae*
HAD	Haloalkanoic acid dehalogenase
Hep	Heptose
HEPES	*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid
HS:X	Heat stable
Hz	Hertz
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k_{cat}	Turnover number
KCl	Potassium chloride
Kdo	3-deoxy-D-*manno*-octulosonate
Kdo8P	3-deoxy-D-*manno*-octulosonate 8-phosphate
K_m	Michaelis constant
L	Liter
Leu	Leucine
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
Lys	Lysine
mM	Millimolar

M	Molar
M. jannaschii	*Methanocaldococcus jannaschii*
m/*z*	Mass to charge ratio
Man	D-Mannopyranose
MeOPN	*O*-methyl phosphoramidate
MFS	Miller Fisher syndrome
mg	Milligram
$MgCl_2$	Magnesium chloride
MIMS	Membrane Inlet Mass Spectrometry
min	Minute
mL	Milliliter
mM	Millimolar
MWCO	Molecular Weight Cut Off
N. meningitidis	*Neisseria meningitidis*
NaCl	Sodium chloride
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced)
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NCTC	National Collection of Type Cultures
nM	Nanomolar
nm	Nanometer
NMR	Nucleoar Magnetic Resonance

ORFs	Open reading frames
PCR	Polymerase chain reaction
PCV	Pneumococcal conjugate vaccine
PDB	Protein Data Bank
PEP	Phosphoenolpyruvate
Phe	Phenylalanine
P_i	Inorganic phosphate
PLP	Pyridoxal phosphate
poly-G	Poly-guanine
ppm	Parts per million
Pro	Proline
rcf	Relative centrifugal force
Rib*f*	D-Ribofuranose
RMSD	Root-mean-square deviation
s	Second
S. Boydii	*Shigella boydii*
S. multivorans	*Sabulilitoribacter multivorans*
S. pyogenes	*Streptococcus pyogenes*
Ser	Serine
serinol	2-amino-1,3-propanediol
SSN	Sequence Similarity Network
SSRL	Stanford Synchrotron Radiation Light Source

Thr	Threonine
Tyr	Tyrosine
UDP	Uridine-5'-diphosphate
UMP	Uridine-5'-monophosphate
UV	Ultraviolet
V. cholerae	Vibrio cholerae
V. vulnificus	*Vibrio vulnificus*
Val	Valine
vis	Visible
Xyl*p*	D-Xylopyranose

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

LIST OF SCHEMES

Page

1. INTRODUCTION

1.1. Capsular Polysaccharides

Capsular polysaccharides were first identified in *Escherichia coli* in 1945 by Kuaffmann and Valhne (1). They termed this sugary coat on the exterior of the bacterium the K antigen for the German word "kapsule". The capsular polysaccharide (CPS) is made up of a repeating unit of sugar monosaccharides and various modifications to those sugars. The CPS is the outermost layer of the bacterium and can mask the antigenic proteins on the surface as well as the O-antigen or lipopolysaccharide (LPS) (2). Capsular polysaccharides, and glycans in general, can be extremely diverse. Glycans can be found on proteins, cell surfaces, lipids, and even RNA (3). This diversity is driven by the wide range of stereoisomers that are possible with carbohydrates allowing for a myriad of different interactions. Indeed, a simple molecule of D-glucose has four chiral carbon atoms which means there are 16 possible stereoisomers. The driving force for the chemical diversity of the CPS is believed to be due to selective pressure from bacteriophage which recognize and bind to the capsular polysaccharide (4). However, some diversity is likely driven by horizontal gene transfer, gene duplication, and mutation (5). There are over 90 different capsule serotypes identified in *Streptococcus pneumoniae*, ~80 capsule serotypes in *Klebsiella pneumoniae*, and ~80 capsule serotypes in *Escherichia coli*, to name a few (6, 7). These capsules are firmly attached to the outer membrane of the bacteria and are often the first structure to come in contact with the host (7). Capsular polysaccharides function to

prevent desiccation, adhere to host tissues, and resist specific and nonspecific host immunity (8). Most capsular polysaccharides are found in mucosal pathogens and are necessary for virulence (9).

Campylobacter jejuni is a Gram-negative, pathogenic organism which is the leading cause of gastroenteritis worldwide (10). The bacterium is often found in chickens, cows, goats, and dogs (11). The infection which results from *Campylobacter jejuni* is known as campylobacteriosis which causes diarrhea, fever, cramps, and in rare cases death (12). Infections frequently result from improperly handled poultry, unclean water, pet feces, and unpasteurized milk (11). Campylobacteriosis is most commonly self-limiting and is cleared by the immune system in 7-10 days (12). Children, the elderly, and individuals with suppressed immune systems are most at risk from campylobacteriosis. The most recent estimates of *Campylobacter jejuni* infection found that it results in ~25000 deaths in children per year worldwide (13). Antibiotic resistance in *C. jejuni* has become more common in recent years. Indeed, resistance to ciprofloxacin has nearly doubled in the last 20 years (14-16). In the last 5 years in the United States, there have been two outbreaks of multi-drug resistant *Campylobacter jejuni* from puppies sold from Petland. These outbreaks resulted in 169 cases of campylobacteriosis and 32 hospitalizations across 24 different states (17, 18). In rare cases, infection of *C. jejuni* results in autoimmune diseases known as Guillain-Barré syndrome (GBS) and Miller Fisher syndrome (MFS). It is estimated that ~40% of all newly diagnosed cases of GBS are preceded by a *C. jejuni* infection (19, 20).

Campylobacter infection can be treated with antibiotics such as azithromycin and ciprofloxacin, but due to the growing resistance, new therapeutics need to be developed (14-16). Currently, there are no FDA approved vaccines for *Campylobacter jejuni* infection (13, 21). Due to the risk of GBS and MFS, whole cell attenuated vaccines cannot be used (22). The molecular mimicry of the sugars found on the surface of gangliosides and the sugars of the lipooligosaccharide (LOS) of *C. jejuni* make whole-cell attenuated vaccines unusable (22). Conjugate vaccines are an attractive alternative to whole-cell attenuated vaccines because they use only the capsular polysaccharide of the bacterium conjugated to a protein (13, 21). Conjugate vaccines which present the CPS of pathogenic bacteria such as *Neisseria meningitidis, Salmonella typhi, Haemophilus influenzae,* and *Streptococcus pneumoniae* have been used as preventative medicine for over 30 years (23, 24). The first pneumococcal pneumonia vaccine was introduced in 1977 as a 14 valent conjugate vaccine (25). Today, the standard of care is pneumococcal conjugate vaccine-13 (PCV13) for infants and children under two and pneumococcal conjugate vaccine-23 (PCV23) for adults (26). PCV13 is a 13 valent conjugate vaccine while PCV23 is a 23 valent conjugate vaccine. Pneumococcal conjugate vaccines represent a 7 billion dollar a year industry in the United States (27). The CPS of *Campylobacter jejuni* is an enticing target to study in order to help further the development of conjugate vaccines.

1.2. Capsule Types

The capsular polysaccharide is a thick sugar coat found on the surface of *Campylobacter jejuni*. The capsule extends outward from the surface of the bacterium and can be upwards of 200 units long (28). The CPS is the major determinant in the Penner serotyping system for *Campylobacter jejuni*. However, new methods of multiplex polymerase chain reaction (PCR) have largely replaced Penner serotyping in *C. jejuni* in the last decade (29). As of 2015, there are a total of 47 serotypes and 35 different capsule types found in *C. jejuni* (30). Unfortunately, of those 35 different capsule types, only 12 capsules have been chemically validated. A summary of the capsule types is presented in Table 1.

Table 1 *Campylobacter jejuni* capsule types and accession numbers. Adapted from Poly et al. (30).

Strain	Accession numbers	Penner type	CPS structure validated (Yes/No)	References
ATCC 43429	BX545859	HS:1/HS:44	**Y**	31, 38
NCTC 11168	AL111168.1	HS:2	**Y**	32, 39
ATCC 43431	HQ343268	HS:3	**Y**	31, 40
ATCC 43432	HQ343269	HS:4	**Y**	31, 41
GC8486	HQ343269	HS:4/13/64	N	33
ATCC 43433	KT868847	HS:5	**Y**	31, 42
NCTC 11828	NC_009839	HS:6/7	**Y**	34, 43
ATCC 43436	HQ343270	HS:8	N	31
ATCC 43437	KT868844	HS:9	N	31
ATCC 43438	HQ343271	HS:10	**Y**	31, 44
RM3415	KT868845	HS:11	N	35
RM3204	KT868848	HS:12	N	35
ATCC 43441	HQ343269	HS:13	N	31
ATCC 43442	HQ343272	HS:15	**Y**	31, 45
RM3417	AASY01000000	HS:16	N	35

ATCC 43444	HQ343270	HS:17	N	31
RM3419	KT932997	HS:18	N	35
ATCC 43446	BX545860	HS:19	**Y**	31, 46
ATCC 43447	KT868849	HS:21	N	31
ATCC 43448	KT893439	HS:22	N	31
81–176	BX545858	HS:23/36	**Y**	36, 47
RM3423	KT893437	HS:27	N	35
RM3424	KT868846	HS:29	N	35
ATCC 43452	KT868847	HS:31	N	31
RM3425	KT893427	HS:32	N	35
ATCC 43454	KT893436	HS:33	N	31
RM3426	KT893436	HS:35	N	35
RM3428	KT893431	HS:37	N	35
RM3429	KT893430	HS:38	N	35
ATCC 43459	KT893434	HS:40	N	31
ATCC 43460	BX545857	HS:41	**Y**	31, 48
ATCC 43461	HQ343274	HS:42	N	31
RM3432	KT893432	HS:45	N	35
ATCC 43465	HQ343269	HS:50	N	31
RM3434	KT893429	HS:52	N	35
RM1221	CP000025.1	HS:53	**Y**	37, 49
ATCC 43469	KT893433	HS:55	N	31
RM3436	KT893428	HS:57	N	35
ATCC 43470	KT893427	HS:58	N	31
RM3438	KT893426	HS:60	N	35
RM3439	HQ343269	HS:62	N	35
RM3440	KT893438	HS:63	N	35
ATCC 49302	AASY01000000	HS:64	N	31
RM3442	HQ343269	HS:65	N	35

1.3. CPS Gene Cluster Organization and Export

There are four capsular polysaccharide export groups in bacteria which are responsible for the synthesis and transport of the CPS to the surface of the bacterium (50). Group 1 and group 4 are known as the LPS model and they usually use an

undecaprenyl phosphate anchor which is then flipped to the inner membrane where it can undergo further modification. Group 2 and group 3 capsular polysaccharides are synthesized in the cytoplasm and are then exported to the outer membrane through an ATP-binding-cassette (ABC) transporter. The group 2 and group 3 export systems are organized into three distinct regions as shown in Figure 1 (adapted from Guerry et. al) (51). Regions one and three are home to the *kps* genes which are responsible for the formation of the phospholipid-linked 3-deoxy-D-*manno*-octulosonate (Kdo) as well as the transport of the CPS to the surface of the bacterium (52-54). The capsular polysaccharide of *Campylobacter jejuni* uses a hybrid system which contains parts of group 2 and group 3 to assemble and export its CPS (55).

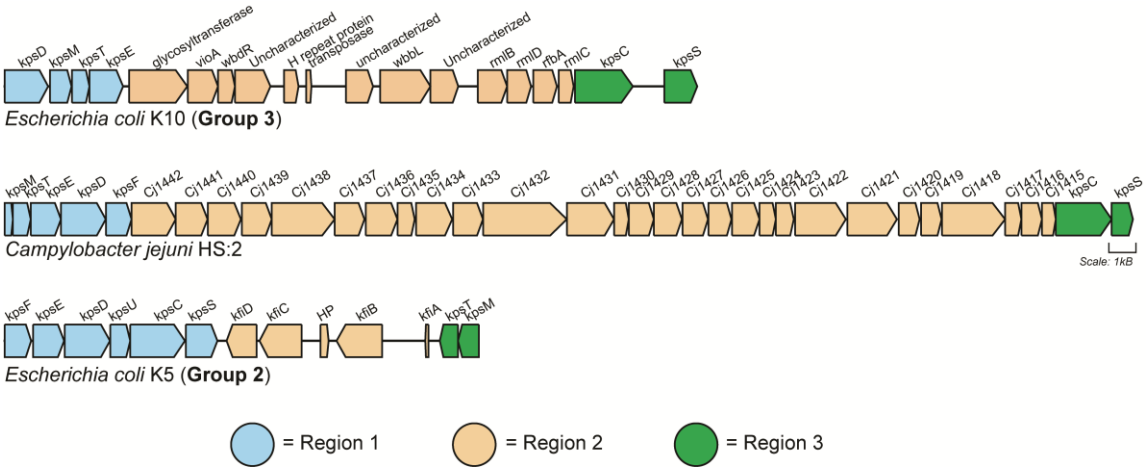
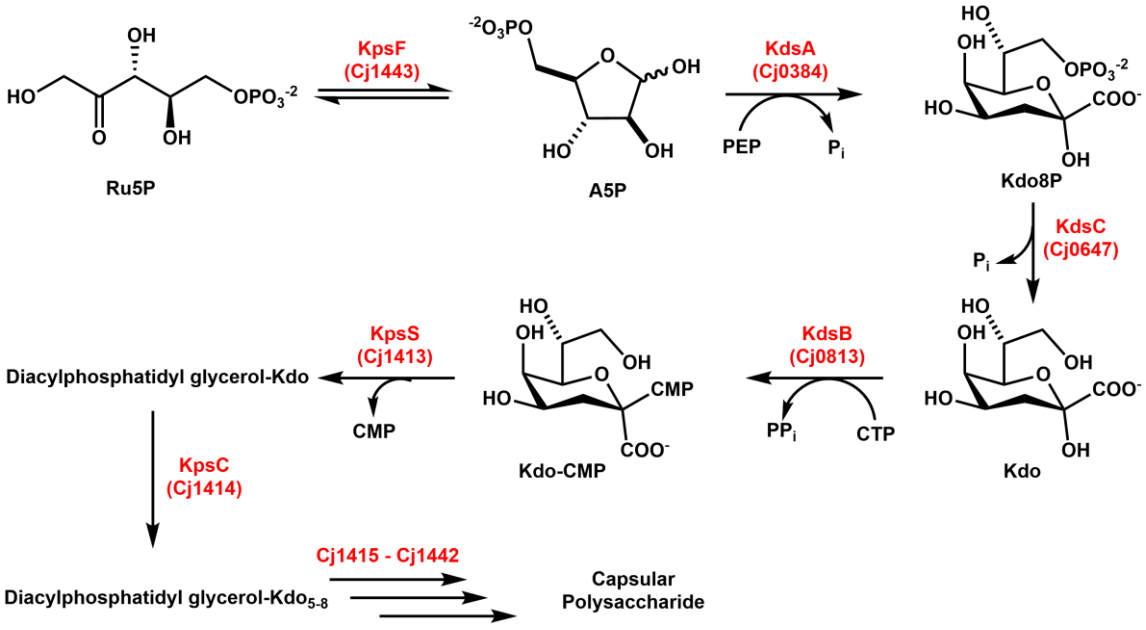


Figure 1 Genetic organization of group 2 and group 3 capsular polysaccharide gene clusters. CPS gene clusters for *Escherichia coli* K10, *Campylobacter jejuni* HS:2, and *Escherichia coli* K5. Shown in blue is region 1 of the CPS gene cluster. Shown in tan is region 2, which is also known as the variable region. Finally, shown in green is region 3 of the CPS gene cluster. Adapted from Guerry et. al (51, 56)

This system of CPS assembly and export can also be found in other pathogenic bacteria, such as *E. coli*, *H. influenzae*, and *N. meningitidis*, to name a few (57, 58). The organizational structure of the CPS gene cluster does not affect the activity of the *kps* genes (53). None of the Kps proteins in *C. jejuni* have been functionally characterized, but they share a high degree of homology to those in *E. coli* which have been characterized. KpsF is an arabinose 5-phosphate isomerase which catalyzes the isomerization reaction between ribulose 5-phosphate and arabinose 5-phosphate (59). KdsA, located elsewhere in the *C. jejuni* genome, will catalyze the formation of 3-deoxy-D-*manno*-octulosonate 8-phosphate (Kdo8P) from the condensation of arabinose 5-phosphate and phosphoenolpyruvate (60). Kdo8P will then be dephosphorylated by KdsC to form α-Kdo (61). KpsU will use the α-Kdo and CTP to form cytidine monophosphate 3-deoxy-α-D-*manno*-octulosonic acid (CMP-Kdo) (62). The *Campylobacter jejuni* CPS gene cluster does not contain a gene for KpsU, but it is hypothesized to use the homolog, KdsB (8). Next, KpsS will attach the first Kdo molecule to a phosphatidylglycerol which is embedded in the membrane. Unlike α-Kdo which acts as a bridge between lipid A and the core oligosaccharide of lipopolysaccharide (LPS), the Kdo residue attached to the lipid anchor of the CPS is in the β- configuration (52). Next, KpsC will extend the Kdo-lipid anchor by adding 5-9 residues of β-Kdo (52). These reactions are summarized in Scheme 1.

Scheme 1 Biosynthetic pathway of diacylphosphatidyl glycerol-Kdo_{5-8} formation.



In order to modify and extend the CPS, the organism will use enzymes found in region 2 of the gene cluster. This variable region will contain between 11 and 34 open reading frames (ORFs) in *Campylobacter jejuni* depending on the serotype (51). It is unknown what causes the termination of the polysaccharide or what initiates the export of the CPS, however, mutants of KpsT in *E. coli* accumulate CPS which is larger than average in the cytosol implying that synthesis and export may be coupled in some fashion (63). Export of the CPS is accomplished through use of the *kpsMTED* genes (64). KpsT couples the hydrolysis of adenosine triphosphate (ATP) to export the CPS through KpsM, a transmembrane transporter. KpsE and KpsD are responsible for

transporting the CPS through the periplasm and inserting it into the outer membrane. A summary of *C. jejuni* CPS export is presented in Figure 2.

There is some debate over the composition of the lipid anchor which attaches the CPS to the outer membrane. The lipid anchor to which the CPS is attached was initially believed to be a lyso-phosohpatidylglycerol molecule due to its copurification with the capsular polysaccharide (63). However, it was recently demonstrated that purified glycosyltransferases can use either monoacyl or diacyl acceptors like 16:0-18:1 phosphatidylglycerol (65). Another piece of evidence in support of the lipid molecule being a diacyl phosphatidylglycerol is the fact that lyso-phosphatidylglycerol is in low abundance in the cell and is implicated in membrane disruption (66). In opposition to this evidence, recent studies with super-resolution fluorescence microscopy of the K1 antigen of *E. coli* revealed the CPS clustered near the poles of the bacterium in lyso-phosphatidyglycerol rafts that diffuse away from the polysaccharide transport tunnels (67). Further study must be done to more fully characterize the phospholipid anchor. Additionally, structures of the ABC transporter system are not available, but are needed in order to further elucidate the mechanism of export for capsular polysaccharides. In summary, there are three main issues surrounding CPS export that remain to be solved: which enzyme catalyzes the glycosidic bond of the first sugar to the poly-Kdo-phosphatidylglycerol acceptor; which phosphatidylglycerol lipid anchor is used *in vivo*; and what do the structures of the ABC transporter complex look like.

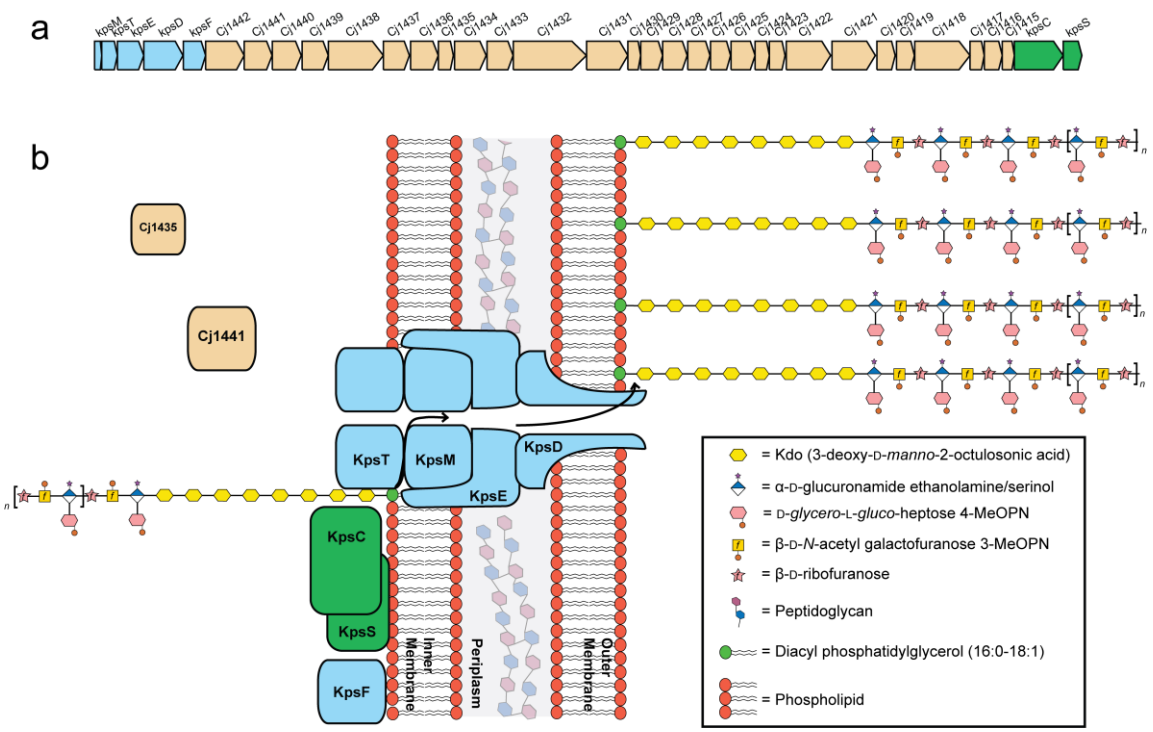


Figure 2 Capsular polysaccharide ABC-transporter export system. (a) *Campylobacter jejuni* HS:2 capsular polysaccharide gene cluster. Region 1, shown in blue, is made of genes involved in the ABC-transporter complex. Region 2, shown in tan, are responsible for modification and elongation of the capsular polysaccharide. Region 3, shown in green, is responsible for the synthesis of the poly-Kdo attached to the diacyl phosphatidylglycerol. (b) Representation of the CPS export system on the surface of *Campylobacter jejuni* NCTC 11168 HS:2. (56)

1.4. *Campylobacter jejuni* CPS structures

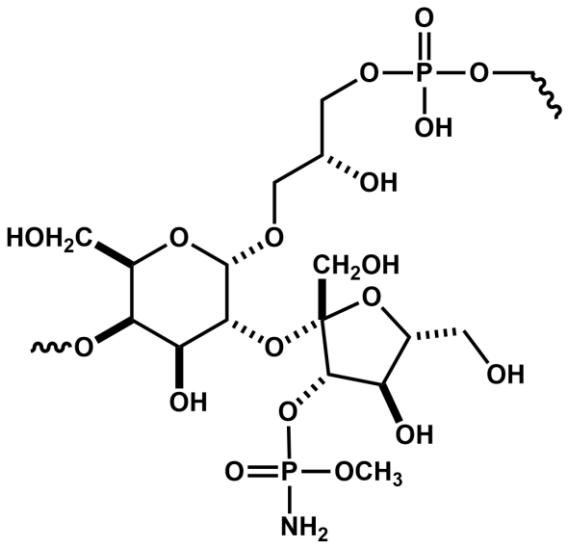
There are 12 structurally characterized capsules from *Campylobacter jejuni*. Briefly, the capsule structures are known for the following serotypes: HS:1/HS:44, HS:2, HS:3, HS:4, HS:5, HS:6/HS:7, HS:10, HS:15, HS:19, HS:23/HS:36, HS:41, and HS:53. These serotypes range from the simplest being a repeating unit of two monosaccharides to the most complex being a repeating unit of four monosaccharides. In addition to the

backbone sugars, some capsular polysaccharides are modified with features such as heptose sugars, *O*-methylation, phosphoramidylation, and amidation. Each structurally characterized capsular polysaccharide will be briefly described below.

1.4.1. HS:1/HS:44 Complex

The HS:1/HS:44 complex is composed of [→4)-[β-D-Fru*f*-(2→2)]-α-D-Gal*p*-(1→2)-D-Gro-(1-P→]$_n$ (38). This polysaccharide can be further modified with a non-stoichiometric addition of a β-D-Fru*f* at the 3-position of the D-Gal*p*. This complex also contains the non-stoichiometric addition of *O*-methyl phosphoramidate (MeOPN) on the 3-position of the β-D-Fru*f*. Interestingly, the linkage between the D-glycerol phosphate and galactose residue is a phosphodiester linkage. The structure of the CPS is shown in Scheme 2.

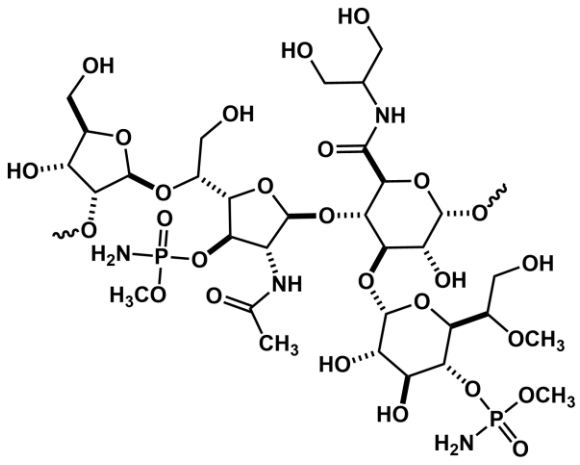
Scheme 2 HS:1/HS:44 CPS structure



1.4.2. HS:2

The HS:2 serotype is from strain NCTC 11168 which was the first *C. jejuni* genome sequenced (32). This is the most well studied serotype and also one of the largest and most complex. It is composed of a repeating unit of [→2)-β-D-Rib*f*- (1→5)-β-D-Gal*f*NAc-(1→4)-α-D-Glc*p*A6Etn-(1→]$_n$ (39). The HS:2 serotype contains a non-stoichiometric modification of MeOPN on the hydroxyl group of C3 of the β-D-Gal*f*NAc. This capsule is further modified non-stoichiometrically with a D-*glycero*-α-L-*gluco*-heptose. This heptose sugar can be further modified non-stoichiometrically with a MeOPN at C4, *O*-methylation at C3, and 6-dehydration. Additionally, the D-Glc*p*A6Etn has a phase variable modification which adds a CH_2OH group forming D-Glc*p*A6NGro. The structure of the CPS with serinol attached is shown in Scheme 3.

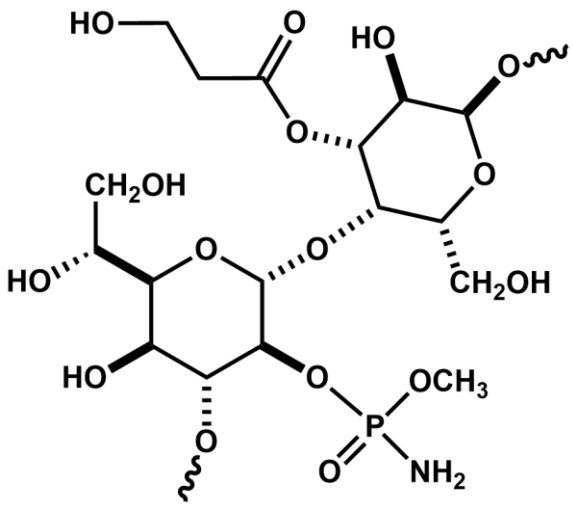
Scheme 3 HS:2 CPS structure with serinol



1.4.3. HS:3

The HS:3 capsular polysaccharide is composed of a repeating unit of [→3)-L-*glycero*-α-D-*ido*-Hep*p*-(1→4)-α-D-Gal*p*-(1→]$_n$ (40). The heptose moiety in this serotype contains non-stoichiometric ratios of the 6-deoxy and 6-oxy forms. It is hypothesized that a poly G tract found in the *dmhA* gene is responsible for the presence or absence of the 6-deoxy modification (manuscript in preparation). Additionally, the heptose can be further decorated with MeOPN at C2. The galactose residue can also be modified with a 3-hydroxypropanoyl at C3. It is unknown how this 3-hydroxypropanyol is attached to the galactose. In Scheme 4, the structure of the CPS is shown.

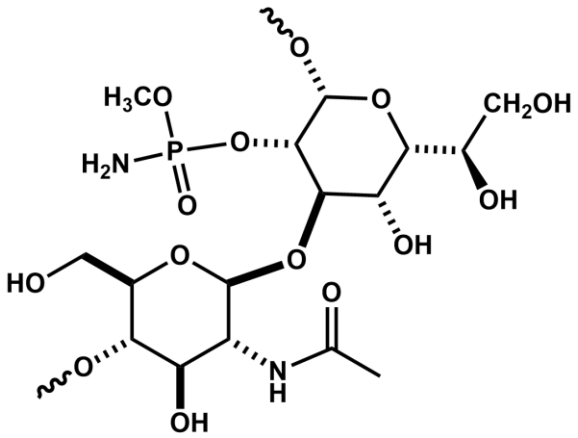
Scheme 4 HS:3 CPS structure



1.4.4. HS:4 Complex

Copies of the genes necessary for synthesis of the HS:4 capsule are found in serotypes HS:13, HS:16, HS:43, HS:50, HS:62, HS:64, and HS:65. This is likely due to a horizonal gene transfer event. The capsule structure is composed of a repeating unit of [→4)-β-D-Glc*p*NAc-(1→3)-L-*glycero*-β-D-*ido*-Hep*p*-(1→]$_n$ or [→4)-β-D-Glc*p*NAc-(1→3)-6-deoxy-β-D-*ido*-Hep*p*-(1→]$_n$. It was also observed that some of the *ido*-heptopyranoses have an MeOPN modification on either O2 or O7. In addition to these two capsule structures, a HS:13 serotype has been found to contain a repeating unit of [→4)-β-D-Glc*p*-(1→3)-6-deoxy-β-D-*ido*-Hep*p*-(1→]$_n$. It is unclear how these variations are biosynthesized from the gene clusters. The CPS is shown in Scheme 5 (41).

Scheme 5 HS:4 CPS structure

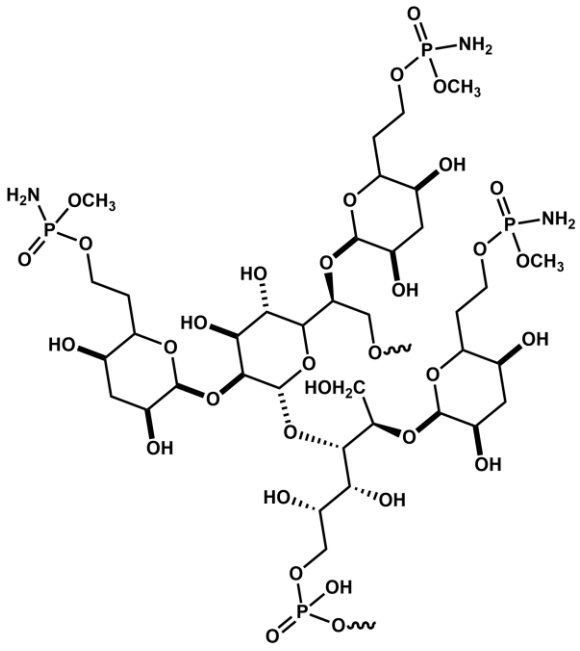


1.4.5. HS:5.

The dominant capsule type for the HS:5 serotype is [→6)-[α-3,6-dideoxy-*ribo*-Hep*p*-(1→2,6)]-α-D-*glycero*-D-*manno*-Hep*p*-(1→3)-[α-3,6-dideoxy-*ribo*-Hep*p*-(1→2)]-glucitol-(6→P]$_n$. There are also three less common variants of this this serotype which are sometimes observed and consist of [→6)-α-D-*glycero*-D-*manno*-Hep*p*-(1→2)-glucitol-(6→P]$_n$, [→6)-[α-3,6-dideoxy-*ribo*-Hep*p*-(1→2)]-α-D-*glycero*-D-*manno*-Hep*p*-(1→2)-glucitol-(6→P]$_n$, and [→6)-[α-3,6-dideoxy-*ribo*-Hep*p*-(1→2)]-α-D-*glycero*-D-*manno*-Hep*p*-(1→3)-[α-3,6-dideoxy-*ribo*-Hep*p*-(1→2)]-glucitol-(6→P]$_n$. Present on the

hydroxyl group of C7 of the α-3,6-dideoxy-*ribo*-Hep*p* is an MeOPN modification. The CPS for HS:5 is shown in Scheme 6 (42).

Scheme 6 HS:5 CPS structure

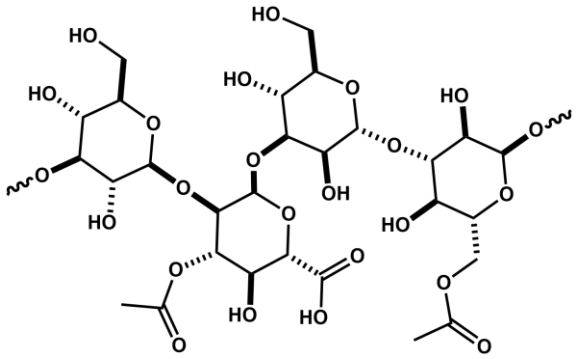


1.4.6. HS:6/HS:7

The complex of serotypes HS:6 and HS:7 has two known structures, termed A and B. The molar ratio of structure A to structure B is 3:1. The chemical structure of structure A is [→3)-β-D-Glc*p*-(1→2)-α-D-Glc*p*A-(1→3)-α-D-Man*p*-(1→3)-α-D-Glc*p*(1→]$_n$. Structure B is composed of [→3)-β-D-Glc*p*NAc-(1→6)-α-D-Glc*p*-(1→4)-(β-D-Glc*p*NAc-(1→3))-α-D-Gal*p*-(1→]$_n$. Structure A can be further modified to contain a non-stoichiometric amount of *O*-acetylation of the hydroxyl group of C3 of the α-D-Glc*p*A as well as the hydroxyl group of C6 of the α-D-Glc*p* molecule. Intriguingly, structure B is possibly associated with lipid A suggesting a different transport and

assembly system than structure A. It is unclear how the difference in the two structures arise. Scheme 7 shows the CPS structure of the HS:6/HS:7 serotype (43).

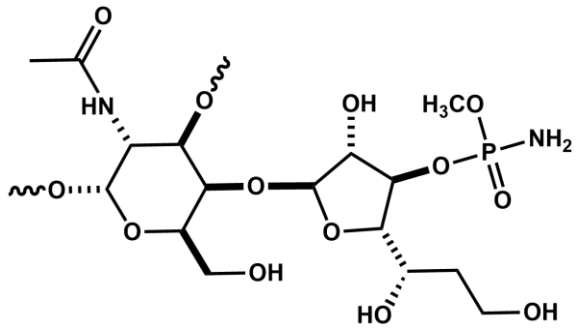
Scheme 7 HS:6/HS:7 CPS structure A



1.4.7. HS:10

The HS:10 serotype capsular polysaccharide is composed of [→3)-[6d-α-L-*gal*-Hep*f*-(1→4)-α-D-GalNAc-(1→]$_n$ with an MeOPN modification at the hydroxyl of C3 of the 6-deoxy-L-*galacto*-heptofuranose. The structure of the HS:10 CPS is shown in Scheme 8 (44).

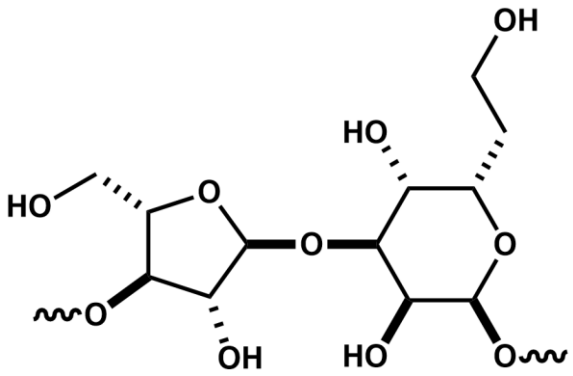
Scheme 8 HS:10 CPS structure



1.4.8. HS:15

The HS:15 capsule is composed of a repeating unit of [→3)-6d-α-L-*gulo*-Hep*p*-(1→3)-α-L-Ara*f*-(1→]$_n$. The heptose is present in the 6-deoxy form in all of the structural studies, however the *dmhA* gene that encodes a dehydratase contains a poly(G) tract which can be turned ON or OFF. It is possible that some of the capsules will have a 6-oxy heptose if the gene is turned OFF. The CPS structure of HS:15 serotype is shown in Scheme 9 (45).

Scheme 9 HS:15 CPS structure

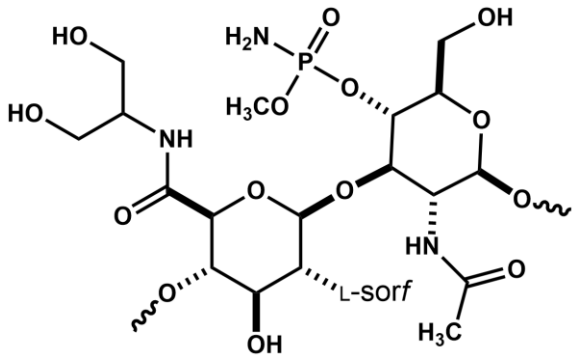


1.4.9. HS:19

The HS:19 capsule is the only structurally characterized capsule other than HS:2 which contains a glucuronamide modification. The repeating unit for the HS:19 capsular polysaccharide is [→4)-β-D-Glc*p*A6NGro-(1→3)-β-D-Glc*p*NAc-(1→]$_n$. The glucuronamide modification is stoichiometrically present on the capsule. Interestingly, there is a non-stoichiometric modification of α-L-sorbofuranose that is found on the C2

hydroxyl of the β-D-Glc*p*A6NGro. It is unknown how the sorbofurnanose is attached to the polysaccharide. There is also a phosphoramidate modification present on the C4 hydroxyl group of the β-D-Glc*p*NAc. The structure of the HS:19 serotype is shown in Scheme 10 (46).

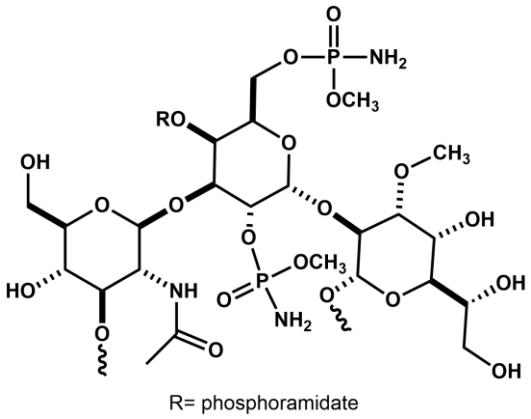
Scheme 10 HS:19 CPS structure



1.4.10. HS:23/HS:36

The HS:23/HS:36 capsule is composed of repeating trisaccharide units of [→3)-α-D-Gal*p*-(1→2)-6d-3-*O*-Me-α-D-*altro*-Hep*p*-(1→3)-β-D-Glc*p*NAc-(1]$_n$ or [→3)-β-D-Glc*p*NAc-(1→3)-α-D-Gal*p*-(1→2)-D-*glycero*-α-D-*altro*-Hep*p*-(1→]$_n$. The HS:23/36 seroclass is composed of serotypes HS:23, HS:36, and HS:23/36. Sequencing has revealed a 97% identity between the genomes for the three strains. The heptose molecule can be methylated at O3. It can also be dehydrated at C6 leading to either a 6-deoxy or 6-oxy form at C6. The galactose residue can be modified with a phosphoramidate at the hydroxyl of C2, C4, or C6. It is unknown if a phosphoramidate modification can occur at all three positions at the same time. The HS:23/HS:36 serotype is shown in Scheme 11 (47).

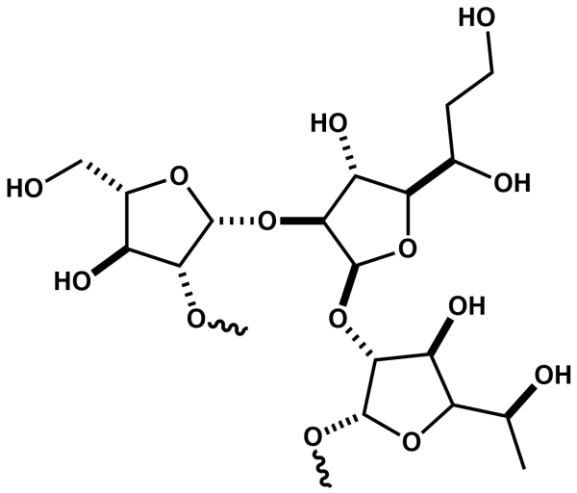
Scheme 11 HS:23/HS:36 CPS structure



1.4.11. HS:41

The HS:41 capsule is composed of a repeating unit of [→2)-β-L-Ara*f*-(1→2)-6d-β-D-*altro*-Hep*f*-(1→2)-α-D-Fuc*f*-(1→]$_n$. The α-D-Fuc*f* can be substituted for β-6d-L-Altf which is a C5 epimer of α-D-Fuc*f*. Other than HS:15, this is the only structurally characterized strain which contains the arabinose modification. Interestingly, the HS:41 serotype is one of only two capsule structures which contain a heptose in the furanose form. HS:10 is the only other CPS structure with a heptofuranose. The HS:41 CPS structure is depicted in Scheme 12 (48).

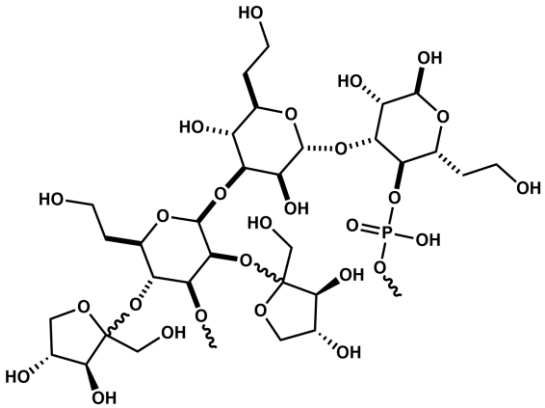
Scheme 12 HS:41 CPS structure



1.4.12. HS:53

The capsule structure of serotype HS:53 is composed of a repeating unit of [→3)-β-6d-D-*manno*-Hep*p*-(1→3)-α-6d-D-*manno*-Hep*p*-(1→3)-α-6d-D-*manno*-Hep*p*-(1-P→]$_n$. The β-6d-D-*manno*-Hep*p* can be modified non-stoichiometrically with D-xylulofuranose at the hydroxyl of C2 or C4. The 6-deoxy forms of the heptoses are likely caused by the phase variable *dmhA* gene. The CPS structure for HS:53 is presented in Scheme 13 (49).

Scheme 13 HS:53 CPS structure



1.5. CPS Monosaccharides

The capsular polysaccharide of *Campylobacter jejuni* is composed of numerous monosaccharides. Table 2 contains a list of the monosaccharides found in the various capsular polysaccharides of *C. jejuni*.

Table 2 *C. jejuni* monosaccharides

Monosaccharide	*C. jejuni* serotype
N-acetylgalactosamine	HS:2, HS:10
Galactose	HS:1/44, HS:3, HS:6/7, HS:23/36
Fructose	HS:1/44
Heptose	HS:1/44, HS:2, HS:3, HS:4, HS:5, HS:10, HS:15, HS:23/36, HS:41, HS:53
Ribose	HS:2
Glucose	HS:2, HS:4, HS:6/7, HS:19
N-acetylglucosamine	HS:6/7, HS:19, HS:23/36
Glucitol-1-phosphate	HS:5
Mannose	HS:6/7
Arabinose	HS:15, HS:41
Sorbose	HS:19
Fucose	HS:41
Altrose	HS:41
Xylulose	HS:53
Glycerol phosphate	HS:1

1.6. Capsule modifications

The capsular polysaccharide of *C. jejuni* can be highly modified. These modifications come in the form of the addition of sugars like heptose, arabinose, sorbose, or glycerol phosphate; as well as non-sugar modifications like, O-methylation, phosphoramidylation, and amidation. The function of these modifications is not well understood. Some speculate that these modifications aid the bacterium in the prevention of desiccation. Others believe the modifications are used to disguise the organism from the immune system of the host or phage attack. Many of the modifications result from genes which contain a poly-G tract which allows the bacterium to turn the modification on or off in its progeny. The ability to quickly modify the surface exposed sugars of the CPS may confer an advantage to the bacterium's offspring. Described below are some of the more prevalent modifications that are present on the capsules of different serotypes of *C. jejuni*.

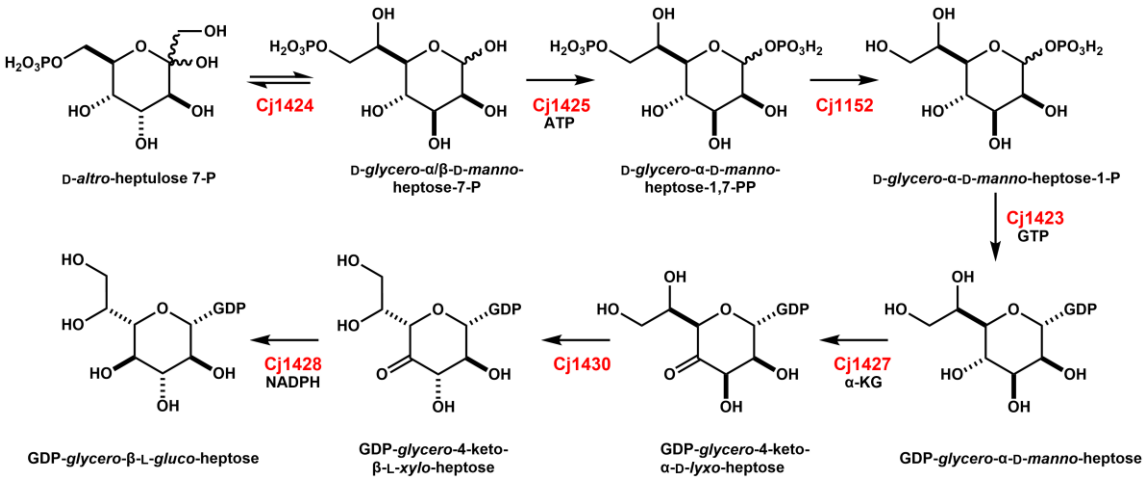
1.6.1. Heptose modification

The heptose modification is one of the most ubiquitous modifications found in *Campylobacter jejuni*. There are 13 different heptose modifications that are present in the various serotypes of *C. jejuni* (68). It is believed the sugars all start with a common precursor which is guanosine-5'-diphosphate (GDP)-D-*glycero*-α-D-*manno*-heptose (69). The enzymes necessary for the biosynthesis of the GDP-α-L-*gluco*-heptose have been characterized in the HS:2 serotype (69-73). The GDP-D-*glycero*-α-D-*manno*-heptose is synthesized using D-*altro*-heptulose-7-phosphate and 4 enzymes: Cj1424, Cj1425,

Cj1152, and Cj1423 (Scheme 14) (70). Starting with D-*altro*-heptulose 7-phosphate, Cj1424 catalyzes the formation of D-glycerol-α/β-D-*manno*-heptose-7-phosphate. Next, Cj1425 will phosphorylate the C1 position using ATP. Cj1152 was demonstrated to be a phosphatase capable of catalyzing the hydrolysis of the phosphate group at C7 of D-*glycero*-α-D-*manno*-heptose-1,7-bisphosphate. Finally, Cj1423 will add GDP to the C1 position to create the GDP-*glycero*-α-D-*manno*-heptose. The GDP-D-*glycero*-α-D-*manno*-heptose will serve as the starting material for subsequent transformations. In *C. jejuni* NCTC 11168 (HS:2), the enzymes required to transform the GDP-D-*glycero*-α-D-*manno*-heptose to GDP-D-*glycero*-β-L-*gluco*-heptose are Cj1427, Cj1430, and Cj1428 (Scheme 14) (69). Cj1427 acts as a dehydrogenase to form GDP-4-keto-α-D-*lyxo*-heptose. Cj1430 will then act as a 3/5 epimerase to generate the GDP-glycero-4-keto-β-L-*xylo*-heptose. Finally, Cj1428 will catalyze the NADPH-dependent reduction to the final product, GDP-*glycero*-β-L-*gluco*-heptose.

It has been demonstrated that the C6 position of the heptose modification can come in two distinct forms, either 6-deoxy or 6-oxy. In the HS:23/36 serotype, the CPS contains a 6-deoxy heptose. It was determined that the Cj1427 homolog found in the HS:23/36 gene cluster is a 4, 6-dehydratase that will dehydrate the GDP-D-*glycero*-α-D-*manno*-heptose to GDP-6-deoxy-4-keto-D-*lyxo*-heptose (74). It is believed that poly-G tracts play a key role in protein expression of these 4, 6-dehydratases which control whether the heptose has the 6-deoxy or 6-oxy heptose modifications.

Scheme 14 Biosynthetic pathway for formation of D-*glycero*-L-*gluco*-heptose in *C. jejuni* NCTC 11168 (HS:2)



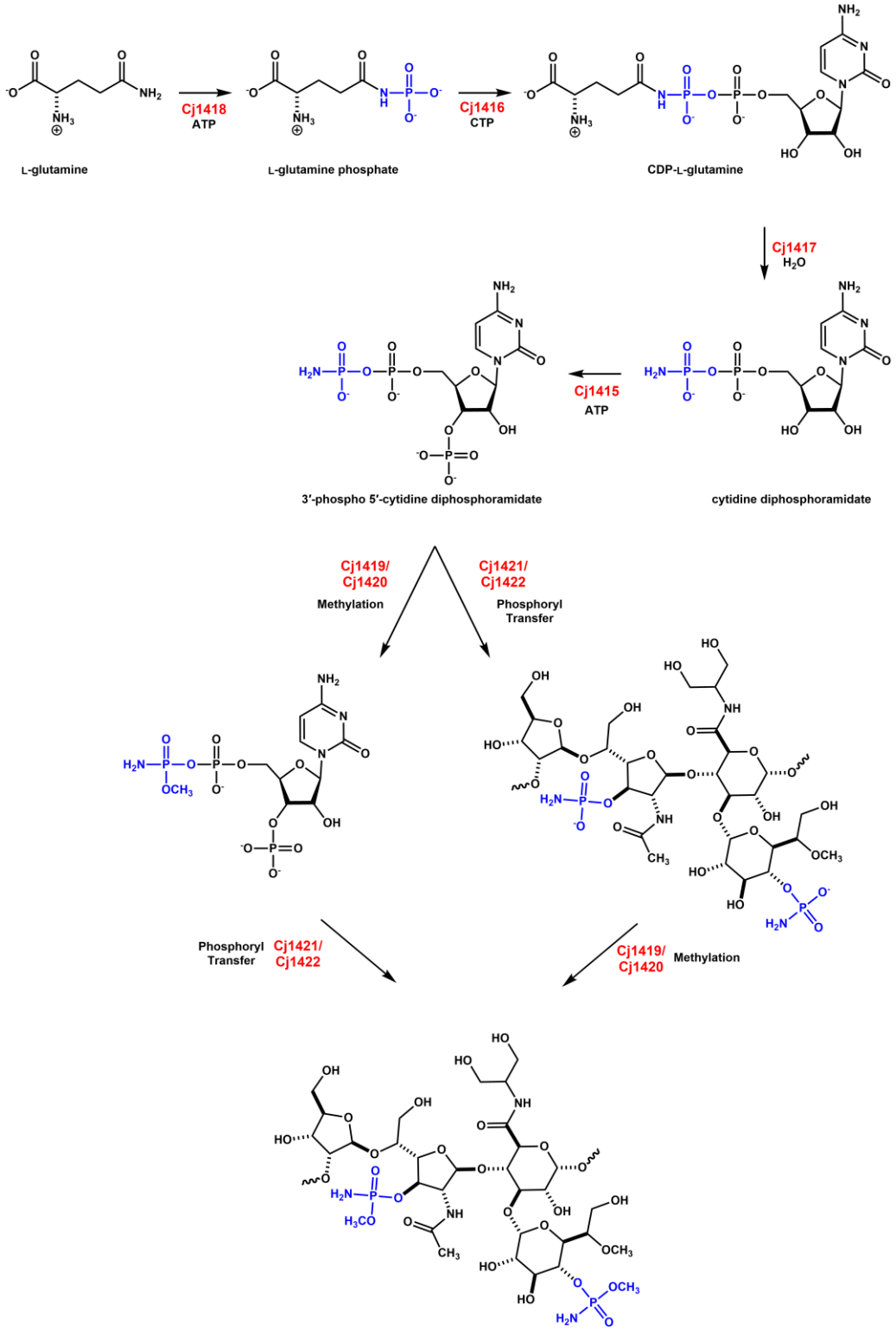
1.6.2. *O*-methyl phosphoramidate modification

The *O*-methyl phosphoramidate modification is found in 9 of the 12 known capsule structures. *C. jejuni* NCTC 11168 (HS:2) requires 8 genes to synthesize and transfer the MeOPN modification to the capsular polysaccharide: Cj1422, Cj1421, Cj1420, Cj1419, Cj1418, Cj1417, Cj1416, and Cj1415. The genes necessary to form the 3′-phospho-5′-cytidine diphosphoramidate cofactor have been functionally characterized (75-79). The reactions for Cj1418, Cj1416, Cj1417, and Cj1415 have been summarized in the top half of Scheme 15. The other four enzymes are likely responsible for the methylation and transfer of the *O*-methyl phosphoramidate modification, but it's currently unknown whether the 3′-phospho-5′-cytidine diphosphoramidate cofactor is methylated first or transferred to the CPS first. The hypothetical activities of the enzymes are represented as a branch point in Scheme 15 with the phosphoramidates

colored blue in the capsular polysaccharide. Through knockout studies, it has been determined that Cj1421 is responsible for the MeOPN transfer to the *N*-acetylgalactofuranose. In the same study, Cj1422 was determined to catalyze the phosphoryltransfer of the MeOPN to the heptose molecule (80).

The *in vivo* function of the *O*-methyl phosphoramidate is unclear. However, it has been demonstrated that *O*-methyl phosphoramidate modification found on the *N*-acetylgalactofuranose of the CPS from *C. jejuni* NCTC 11168 (HS:2) is a recognition site for phage F336 (81). The phosphoryltransferase, Cj1421, has a poly-G tract which can turn ON or OFF the transfer of the MeOPN to the *N*-acetylgalactofuranose. When the transfer is turned OFF, the bacterium is no longer sensitive to phage, whereas when it is turned ON phage will recognize the MeOPN modification (81). It has not been determined whether the other phosphoryltransferase, Cj1422, has any effect on phage recognition.

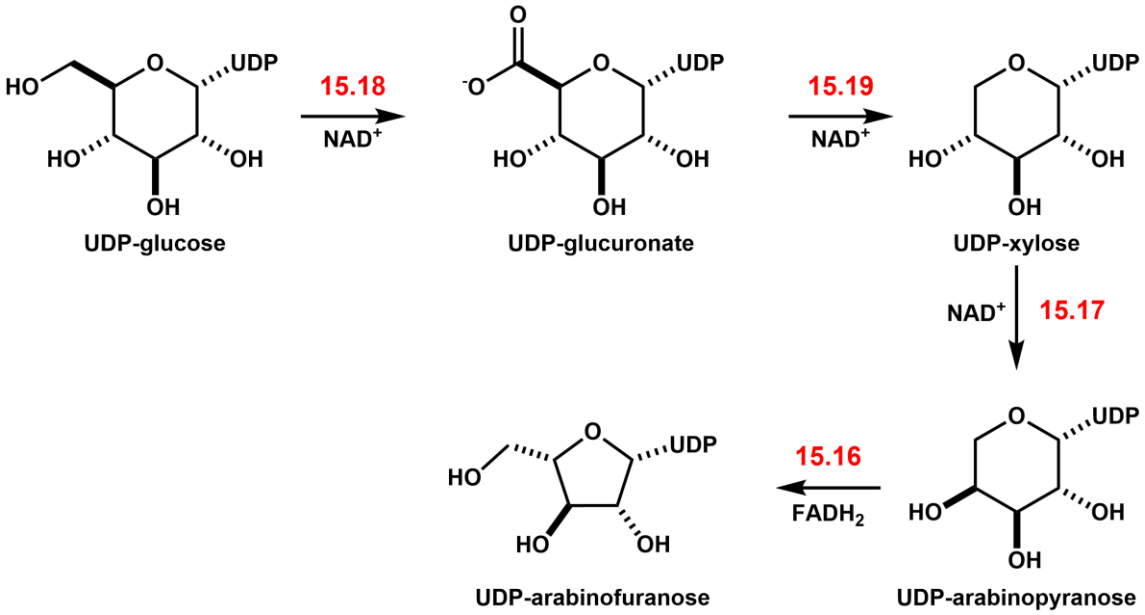
Scheme 15 Proposed biosynthetic pathway for *O*-methyl phosphoramidate modification



1.6.3. Arabinose modification

The arabinose modification is found on the CPS of HS:15 and HS:41 serotypes. The function and the biosynthesis of the arabinose modification are currently unknown. The enzymes hypothesized to be involved in biosynthesis of the arabinose modification in HS:15 are 15.18, 15.19, 15.17, and 15.16. The first enzyme in the proposed pathway, 15.18, catalyzes the nicotinamide adenine dinucleotide (NAD)-dependent oxidation of UDP-glucose to UDP-glucuronic acid. The next proposed enzyme, 15.19, is an NAD^+ dependent UDP-glucuronate decarboxylase that creates a UDP-4-keto-glucuronate which undergoes spontaneous decarboxylation at C6 and subsequent reduction to form UDP-xylose. The next enzyme in the pathway, 15.17, will epimerize the C4 position of UDP-xylose to form UDP-arabinopyranose. The final enzyme in the pathway is proposed to be an FAD-dependent mutase which catalyzes the transformation of UDP-arabinopyranose to UDP-arabinofuranose. This proposed pathway is summarized in Scheme 16. The genes hypothesized for arabinose modification are present in 14 of 35 gene clusters in *C. jejuni*, including serotypes HS:6/7 where the putative arabinose forming genes are present, but the modification is not observed on the chemically validated structure (34, 43). It is not known why the modification is not present on the surface of the CPS HS:6/7.

Scheme 16 Proposed biosynthetic pathway for arabinofuranose modification in *C. jejuni* HS:15



1.6.4. Glycerol phosphate modification

The glycerol phosphate modification is found only in the *C. jejuni* HS:1 serotype. The enzymes necessary for its synthesis have not been determined yet. It is hypothesized that it requires the genes HS1.9 and HS1.11. One of these proteins will use glycerol phosphate and CTP as substrates to catalyze the formation of CDP-glycerol. The other enzyme will facilitate the transfer of the CDP glycerol to a galactose on the polysaccharide chain. Presumably, a glycosyltransferase in the gene cluster will then catalyze the formation of a glycosidic bond between the C1 hydroxyl of glycerol phosphate and the anomeric carbon of galactose using UDP-galactose as the donor sugar. This proposed pathway is very similar to the wall teichoic acid biosynthesis in *Staphylococcus aureus* which uses TarD and TarB to synthesize a phosphodiester bond of glycerol phosphate and *N*-acetylmannosamine (82).

1.7. Capsule Assembly

Currently it is unknown which enzymes are responsible for the polymerization of the capsular polysaccharide in *Campylobacter jejuni*. The enzymes necessary for these transformations are undoubtedly members of the glycosyltransferase families. Currently, none of the glycosyltransferases from *C. jejuni* have been structurally or functionally characterized.

The number of glycosyltransferases found in the CPS of the HS:2 serotype exceeds the number of possible glycosidic bonds present on the CPS. There are 8 glycosyltransferases in the CPS gene cluster of serotype HS:2, however there are only 4 sugar monosaccharides. It is possible that one of the glycosyltransferases is used as the initiator enzyme which would attach the Kdo molecule to the first sugar of the CPS. It is unknown, however, what the first sugar of the CPS will be. It is possible that there are multiple sugars which can be used as the initiating sugar for CPS elongation.

Enzymes in the glycosyltransferase family can either invert or retain the stereochemistry of the anomeric carbon of the donor sugar (83). Most often sugars are linked to a nucleotide and the phosphate group activates the C1 carbon for glycosidic bond formation. Using amino acid sequence similarity, the glycosyltransferase families have been grouped into 115 families according to the CAZy database (84-86). The two most common and evolutionarily eldest glycosyltransferase families are the GT2 and GT4 families (87). The GT2 family is composed of a GT-A fold which contains two Rossmann-like folds tightly associated with one another, often leading to their characterization as a single domain fold (83). The GT4 family, which is the largest of all

the glycosyltransferase families, contains a GT-B fold which also consists of two Rossmann-like folds, however, unlike the GT2 family, these two Rossmann-like folds are loosely associated with one another (87). The GT2 family is a member of the inverting glycosyltransferase clan which will invert the stereochemistry of the anomeric carbon of the donor sugar (88). Meanwhile, GT4 enzymes are known as retaining glycosyltransferases as they will retain the stereochemistry of the anomeric carbon of the donor sugar (Figure 3) (87).

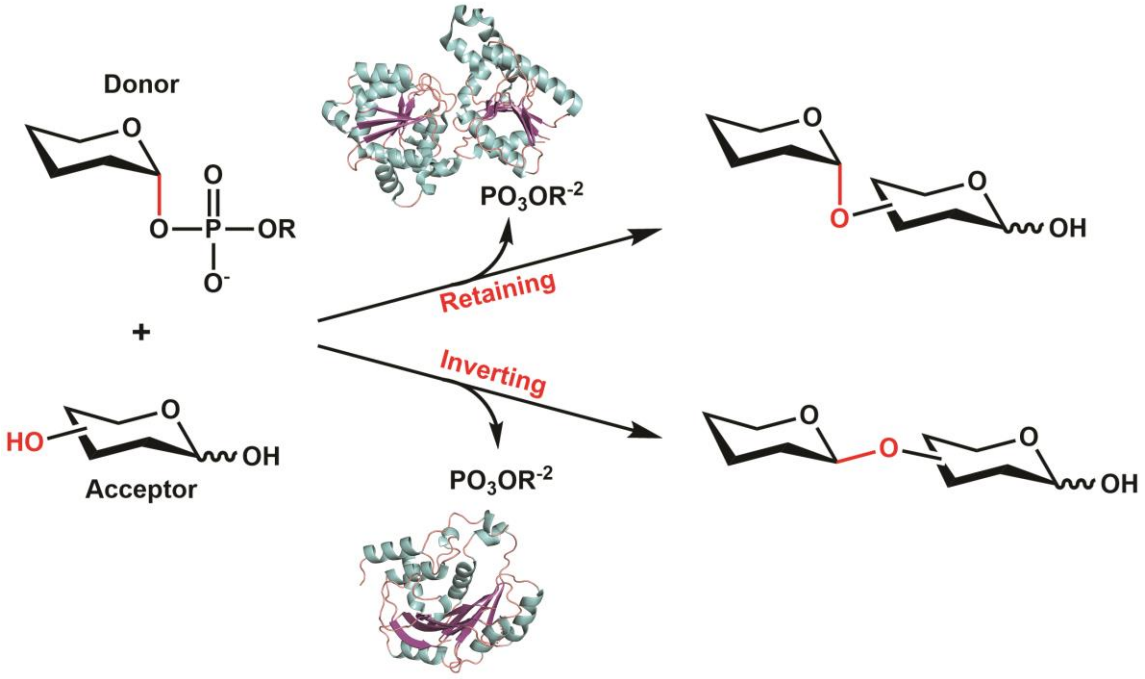


Figure 3 Inverting versus retaining glycosyltransferases (adapted from Withers et. al (83)). An activated donor substrate is attacked by the hydroxyl group of an acceptor molecule. The stereochemistry of the anomeric can either be inverted or retained. Examples of inverting and retaining enzymes are shown as ribbon diagrams colored by their secondary structure. The retaining enzyme is WaaG from *E. coli* (PDB id: 2IW1) (87). The inverting enzyme shown is SpsA from *Bacillus subtilis* (PDB id: 1QGQ) (89).

The assembly of the capsular polysaccharide by the glycosyltransferases occurs in the cytoplasm of the bacterium (90). It is currently unknown whether modifications to the monosaccharides of the CPS occur before or after glycosyltransfer. The non-stoichiometric modifications identified in certain serotypes suggest that at least some of the modifications are attached after glycosyltransfer of the monosaccharides to the polysaccharide chain. A knockout study of a selection of genes from the *C. jejuni* HS:2 serotype found that Cj1431 is required for heptosyltranfer to the capsular polysaccharide (91). High resolution magic angle spinning NMR spectroscopy data suggests the capsule will still be produced in the ΔCj1431 mutant, but will not contain the heptose modification. Additionally, Cj1432, Cj1434, Cj1438, Cj1440, and Cj1442 were individually knocked out and each knockout resulted in an acapsular bacterium (92). While this doesn't help to assign the function of the genes, it suggests these particular glycosyltransferases are necessary for CPS production.

Recently, an enzyme similar to Cj1432 was characterized by the Whitfield group at the University of Guelph (93). They were able to characterize *orf5* from *K. pneumoniae* serotype O4 that catalyzes the formation of a glycosidic bond between monosaccharides of galactopyranose and ribose using 5′-phospho-D-ribosyl-α-1-diphosphate (PRPP) as the donor substrate. The hydroxyl group of the galactopyranose will attack the activated anomeric carbon of PRPP and displace pyrophosphate to form a disaccharide of galactopyranose and ribose 5'-phosphate. The enzyme contains a second domain which functions as a HAD phosphatase to hydrolyze the ribose 5'-phosphate to form the final disaccharide of galactopyranose and ribose.

Similar to *orf5* from *K. pneumoniae* serotype O4, Cj1432 also contains a PRPP transferase domain and a HAD phosphatase domain. Unlike *orf5*, Cj1432 contains a GT4 glycosyltransferase domain on its N-terminus. If Cj1432 functions similar to *orf5*, then one would expect it to catalyze the formation of a glycosidic bond between *N*-acetylgalactosamine in the furanose form and PRPP forming a disaccharide of *N*-acetylgalactosamine in the furanose form and ribose 5'-phosphate. Next, the HAD phosphatase domain of Cj1432 would hydrolyze the ribose 5'-phosphate of the disaccharide to ribose. It is hypothesized that the newly formed disaccharide would be used as the acceptor group for the GT4 glycosyltransferase reaction using UDP-glucuronate as the donor group. This would retain the stereochemistry of the anomeric carbon of the glucuronate in the α- orientation and would form the trisaccharide backbone of the CPS of *C. jejuni* HS:2. This is the only GT4 retaining glycosyltransferase in the CPS gene cluster of *C. jejuni* HS:2.

1.8. Conclusions

Campylobacter jejuni is a pathogenic organism usually found in the intestinal tracts of chickens and other farm animals. The infection which results from *C. jejuni* is known as campylobacteriosis and it causes nausea, cramps, fever, and bloody diarrhea. Antibiotic resistance is a growing concern in *C. jejuni* which underscores the need for an effective vaccine to prevent infection. There is currently no FDA approved vaccine to treat or prevent campylobacteriosis, but the most successful candidates to date have been conjugate vaccines. The bacterium is coated in a thick outer layer of sugars known as

the capsular polysaccharide. It is believed these capsules help to prevent desiccation, evade the host immune system, and adhere to epithelial tissue. The chemical composition of the capsule can be different depending on the strain of *C. jejuni*. Some capsules are a simple repeating disaccharide, while others highly decorated tetrasaccharides. The CPS is exported to the surface of the bacterium using an ABC-transporter system that links the CPS to a poly Kdo-lipid anchor. The CPS can be highly modified with MeOPN, heptose, arabinose, *O*-methylation, and amidation. The genes responsible for the biosynthesis of these modifications have largely been characterized in *C. jejuni* NCTC 11168 (HS:2). It is not well understood how the synthesis of the CPS is initiated, elongated, or terminated. The elongation of the CPS is almost assuredly linked to the glycosyltransferase activities of enzymes in the CPS gene cluster, but this has yet to be experimentally established in *C. jejuni*.

1.9. References

1. Kauffmann, F.; and Vahlne, G. Uber die bedeutung des serologischen formenwechsels fur die bakteriophagen-wirkung in der Coli-Gruppe. *Acta Path. Microbiol. Scand*. 1945, 22: 119-121.
2. Fanning, S.; Hall, L. J.; Cronin, M.; Zomer, A.; MacSharry, J.; Goulding, D.; Motherway, M. O.; Shanahan, F.; Nally, K.; Dougan, G.; and van Sinderen, D. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(6), 2108–2113.
3. Flynn, R. A., Pedram, K., Malaker, S. A., Batista, P. J., Smith, B., Johnson, A. G., George, B. M., Majzoub, K., Villalta, P. W., Carette, J. E., & Bertozzi, C. R. (2021). Small RNAs are modified with N-glycans and displayed on the surface of living cells. Cell, 184(12), 3109–3124.e22.
4. Mostowy, R. J.; and Holt, K. E. Diversity-Generating Machines: Genetics of Bacterial Sugar-Coating. *Trends in microbiology*, 2018, 26(12), 1008–1021.
5. Bian, S., Zeng, W., Li, Q., Li, Y., Wong, N. K., Jiang, M., Zuo, L., Hu, Q., & Li, L. Genetic Structure, Function, and Evolution of Capsule Biosynthesis Loci in Vibrio parahaemolyticus. *Frontiers in microbiology*, 2021, 11, 546150.
6. Wyres, K. L., Wick, R. R., Gorrie, C., Jenney, A., Follador, R., Thomson, N. R., & Holt, K. E. Identification of Klebsiella capsule synthesis loci from whole genome data. *Microbial genomics*, 2016, 2(12), e000102.

7. Whitfield C. Biosynthesis and assembly of capsular polysaccharides in Escherichia coli. *Annual review of biochemistry*, 2006, 75, 39–68.
8. Roberts I. S. The biochemistry and genetics of capsular polysaccharide production in bacteria. *Annual review of microbiology*, 1996, 50, 285–315.
9. Whitfield, C., Wear, S. S., & Sande, C. Assembly of Bacterial Capsular Polysaccharides and Exopolysaccharides. *Annual review of microbiology*, 2020, 74, 521–543.
10 Epps, S. V., Harvey, R. B., Hume, M. E., Phillips, T. D., Anderson, R. C., & Nisbet, D. J. (2013). Foodborne Campylobacter: infections, metabolism, pathogenesis and reservoirs. International journal of environmental research and public health, 10(12), 6292–6304.
11. Burnham, P. M., & Hendrixson, D. R. (2018). Campylobacter jejuni: collective components promoting a successful enteric lifestyle. Nature reviews. Microbiology, 16(9), 551–565.
12. Allos B. M. (2001). Campylobacter jejuni Infections: update on emerging issues and trends. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 32(8), 1201–1206.
13. Poly, F., Noll, A. J., Riddle, M. S., & Porter, C. K. (2019). Update on Campylobacter vaccine development. Human vaccines & immunotherapeutics, 15(6), 1389–1400.

14. Luangtongkum, T., Jeon, B., Han, J., Plummer, P., Logue, C. M., and Zhang, Q. (2009). Antibiotic resistance in Campylobacter: emergence, transmission and persistence. *Future microbiology*, *4*(2), 189–200.
15. Gupta, A., Nelson, J. M., Barrett, T. J., Tauxe, R. V., Rossiter, S. P., Friedman, C. R., Joyce, K. W., Smith, K. E., Jones, T. F., Hawkins, M. A., Shiferaw, B., Beebe, J. L., Vugia, D. J., Rabatsky-Ehr, T., Benson, J. A., Root, T. P., Angulo, F. J., and NARMS Working Group (2004). Antimicrobial resistance among Campylobacter strains, United States, 1997-2001. *Emerging infectious diseases*, *10*(6), 1102–1109.
16. Engberg, J., Aarestrup, F. M., Taylor, D. E., Gerner-Smidt, P., and Nachamkin, I. (2001). Quinolone and macrolide resistance in Campylobacter jejuni and C. coli: resistance mechanisms and trends in human isolates. *Emerging infectious diseases*, *7*(1), 24–34.
17. Centers for Disease Control and prevention (2022, August 15). Outbreak of Multidrug-resistant *Campylobacter* Infections Linked to Contact with Pet Store Puppies. https://www.cdc.gov/campylobacter/outbreaks/puppies-12-19/index.html#:~:text=Investigation%20Details-,April%2015%2C%202021,be%20part%20of%20this%20outbreak.
18. Centers for Disease Control and prevention (2022, August 15). Multistate Outbreak of Multidrug-Resistant *Campylobacter* Infections Linked to Contact with Pet Store Puppies. https://www.cdc.gov/campylobacter/outbreaks/puppies-9-17/index.html

19. Pope, J. E., Krizova, A., Garg, A. X., Thiessen-Philbrook, H., and Ouimet, J. M. (2007). Campylobacter reactive arthritis: a systematic review. *Seminars in arthritis and rheumatism*, *37*(1), 48–55.
20. Nachamkin, I., Allos, B. M., and Ho, T. (1998). Campylobacter species and Guillain-Barré syndrome. *Clinical microbiology reviews*, *11*(3), 555–567.
21. Riddle, M. S., and Guerry, P. (2016). Status of vaccine research and development for Campylobacter jejuni. *Vaccine*, *34*(26), 2903–2906.
22. Scott D. A. (1997). Vaccines against Campylobacter jejuni. The Journal of infectious diseases, 176 Suppl 2, S183–S188.
23. Lindberg A. A. (1999). Glycoprotein conjugate vaccines. Vaccine, 17 Suppl 2, S28–S36.
24. Finn A. (2004). Bacterial polysaccharide-protein conjugate vaccines. British medical bulletin, 70, 1–14.
25. Mirsaeidi, M., & Schraufnagel, D. E. (2014). Pneumococcal vaccines: understanding centers for disease control and prevention recommendations. Annals of the American Thoracic Society, 11(6), 980–985.
26. Matanock, A., Lee, G., Gierke, R., Kobayashi, M., Leidner, A., & Pilishvili, T. (2019). Use of 13-Valent Pneumococcal Conjugate Vaccine and 23-Valent Pneumococcal Polysaccharide Vaccine Among Adults Aged ≥65 Years: Updated Recommendations of the Advisory Committee on Immunization Practices. MMWR. Morbidity and mortality weekly report, 68(46), 1069–1075.

27. Vaxcyte. Vaxcyte Corporate presentation, slide 9. (2022, Augsut 15). https://investors.vaxcyte.com/static-files/e767716a-6a08-4baa-8bea-de10e1809253.
28. Pelkonen, S., Häyrinen, J., & Finne, J. (1988). Polyacrylamide gel electrophoresis of the capsular polysaccharides of Escherichia coli K1 and other bacteria. Journal of bacteriology, 170(6), 2646–2653.
29. Poly, F., Serichatalergs, O., Schulman, M., Ju, J., Cates, C. N., Kanipes, M., Mason, C., & Guerry, P. (2011). Discrimination of major capsular types of Campylobacter jejuni by multiplex PCR. Journal of clinical microbiology, 49(5), 1750–1757.
30. Poly, F., Serichantalergs, O., Kuroiwa, J., Pootong, P., Mason, C., Guerry, P., & Parker, C. T. (2015). Updated Campylobacter jejuni Capsule PCR Multiplex Typing System and Its Application to Clinical Isolates from South and Southeast Asia. PloS one, 10(12), e0144349.
31. Penner, J. L., Hennessy, J. N., & Congi, R. V. (1983). Serotyping of Campylobacter jejuni and Campylobacter coli on the basis of thermostable antigens. European journal of clinical microbiology, 2(4), 378–383.
32. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, 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., Whitehead, S., … Barrell, B. G. (2000). The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. Nature, 403(6770), 665–668.

33. Poly, F., Read, T., Tribble, D. R., Baqar, S., Lorenzo, M., & Guerry, P. (2007). Genome sequence of a clinical isolate of Campylobacter jejuni from Thailand. Infection and immunity, 75(7), 3425–3433.
34. Palmer, S. R., Gully, P. R., White, J. M., Pearson, A. D., Suckling, W. G., Jones, D. M., Rawes, J. C., & Penner, J. L. (1983). Water-borne outbreak of campylobacter gastroenteritis. Lancet (London, England), 1(8319), 287–290.
35. Parker, C. T., Horn, S. T., Gilbert, M., Miller, W. G., Woodward, D. L., & Mandrell, R. E. (2005). Comparison of Campylobacter jejuni lipooligosaccharide biosynthesis loci from a variety of sources. Journal of clinical microbiology, 43(6), 2771–2781.
36. Poly, F., Threadgill, D., & Stintzi, A. (2005). Genomic diversity in Campylobacter jejuni: identification of C. jejuni 81-176-specific genes. Journal of clinical microbiology, 43(5), 2330–2338.
37. Fouts, D. E., Mongodin, E. F., Mandrell, R. E., Miller, W. G., Rasko, D. A., Ravel, J., Brinkac, L. M., DeBoy, R. T., Parker, C. T., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Sullivan, S. A., Shetty, J. U., Ayodeji, M. A., Shvartsbeyn, A., Schatz, M. C., Badger, J. H., Fraser, C. M., … Nelson, K. E. (2005). Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species. PLoS biology, 3(1), e15..
38. McNally, D. J., Jarrell, H. C., Li, J., Khieu, N. H., Vinogradov, E., Szymanski, C. M., & Brisson, J. R. (2005). The HS:1 serostrain of Campylobacter jejuni has a complex teichoic acid-like capsular polysaccharide with nonstoichiometric

fructofuranose branches and O-methyl phosphoramidate groups. The FEBS journal, 272(17), 4407–4422.

39. St Michael, F., Szymanski, C. M., Li, J., Chan, K. H., Khieu, N. H., Larocque, S., Wakarchuk, W. W., Brisson, J. R., & Monteiro, M. A. (2002). The structures of the lipooligosaccharide and capsule polysaccharide of Campylobacter jejuni genome sequenced strain NCTC 11168. European journal of biochemistry, 269(21), 5119–5136.
40. Aspinall, G. O., Lynch, C. M., Pang, H., Shaver, R. T., & Moran, A. P. (1995). Chemical structures of the core region of Campylobacter jejuni O:3 lipopolysaccharide and an associated polysaccharide. European journal of biochemistry, 231(3), 570–578.
41. Chen, Y. H., Poly, F., Pakulski, Z., Guerry, P., & Monteiro, M. A. (2008). The chemical structure and genetic locus of Campylobacter jejuni CG8486 (serotype HS:4) capsular polysaccharide: the identification of 6-deoxy-D-ido-heptopyranose. Carbohydrate research, 343(6), 1034–1040.
42. Pequegnat, B. M. Ph.D. thesis, University of Guelph, Guelph, ON, Canada, 2016. http://hdl.handle.net/10214/9737 (accessed 2022/08/10).
43. Muldoon, J., Shashkov, A. S., Moran, A. P., Ferris, J. A., Senchenkova, S. N., & Savage, A. V. (2002). Structures of two polysaccharides of Campylobacter jejuni 81116. Carbohydrate research, 337(21-23), 2223–2229.
44. Shin, J. E., Ackloo, S., Mainkar, A. S., Monteiro, M. A., Pang, H., Penner, J. L., & Aspinall, G. O. (1997). Lipo-oligosaccharides of Campylobacter jejuni serotype

O:10. Structures of core oligosaccharide regions from a bacterial isolate from a patient with the Miller-Fisher syndrome and from the serotype reference strain. Carbohydrate research, 305(2), 223–232.

45. Bertolo, L., Ewing, C. P., Maue, A., Poly, F., Guerry, P., & Monteiro, M. A. (2013). The design of a capsule polysaccharide conjugate vaccine against Campylobacter jejuni serotype HS15. Carbohydrate research, 366, 45–49.
46. Aspinall, G. O., McDonald, A. G., & Pang, H. (1994). Lipopolysaccharides of Campylobacter jejuni serotype O:19: structures of O antigen chains from the serostrain and two bacterial isolates from patients with the Guillain-Barré syndrome. Biochemistry, 33(1), 250–255.
47. Kanipes, M. I., Papp-Szabo, E., Guerry, P., & Monteiro, M. A. (2006). Mutation of waaC, encoding heptosyltransferase I in Campylobacter jejuni 81-176, affects the structure of both lipooligosaccharide and capsular carbohydrate. Journal of bacteriology, 188(9), 3273–3279.
48. Hanniffy, O. M., Shashkov, A. S., Moran, A. P., Prendergast, M. M., Senchenkova, S. N., Knirel, Y. A., & Savage, A. V. (1999). Chemical structure of a polysaccharide from Campylobacter jejuni 176.83 (serotype O:41) containing only furanose sugars. Carbohydrate research, 319(1-4), 124–132.
49. Gilbert, M., Mandrell, R. E., Parker, C. T., Li, J., & Vinogradov, E. (2007). Structural analysis of the capsular polysaccharide from Campylobacter jejuni RM1221. Chembiochem : a European journal of chemical biology, 8(6), 625–631.

50. Whitfield, C., & Roberts, I. S. (1999). Structure, assembly and regulation of expression of capsules in Escherichia coli. Molecular microbiology, 31(5), 1307–1319.

51. Guerry, P., Poly, F., Riddle, M., Maue, A. C., Chen, Y. H., & Monteiro, M. A. (2012). Campylobacter polysaccharide capsules: virulence and vaccines. Frontiers in cellular and infection microbiology, 2, 7.

52. Willis, L. M., & Whitfield, C. (2013). KpsC and KpsS are retaining 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) transferases involved in synthesis of bacterial capsules. Proceedings of the National Academy of Sciences of the United States of America, 110(51), 20753–20758.

53. Whitfield C. (2009). Structure and Assembly of Escherichia coli Capsules. EcoSal Plus, 3(2), 10.1128/ecosalplus.4.7.3.

54. Vimr, E. R., & Steenbergen, S. M. (2009). Early molecular-recognition events in the synthesis and export of group 2 capsular polysaccharides. Microbiology (Reading, England), 155(Pt 1), 9–15.

55. Stintzi A. (2003). Gene expression profile of Campylobacter jejuni in response to growth temperature variation. Journal of bacteriology, 185(6), 2009–2016.

56. Harrison, K. J., Crécy-Lagard, V., and Zallot, R. (2018). Gene Graphics: a genomic neighborhood data visualization web application. Bioinformatics (Oxford, England), 34(8), 1406–1408.

57. Ward, C. K., & Inzana, T. J. (1997). Identification and characterization of a DNA region involved in the export of capsular polysaccharide by Actinobacillus pleuropneumoniae serotype 5a. Infection and immunity, 65(6), 2491–2496.

58. Azurmendi, H. F., Veeramachineni, V., Freese, S., Lichaa, F., Freedberg, D. I., & Vann, W. F. (2020). Chemical structure and genetic organization of the E. coli O6:K15 capsular polysaccharide. Scientific reports, 10(1), 12608.

59. Meredith, T. C., & Woodard, R. W. (2006). Characterization of Escherichia coli D-arabinose 5-phosphate isomerase encoded by kpsF: implications for group 2 capsule biosynthesis. The Biochemical journal, 395(2), 427–432.

60. Strohmaier, H., Remler, P., Renner, W., & Högenauer, G. (1995). Expression of genes kdsA and kdsB involved in 3-deoxy-D-manno-octulosonic acid metabolism and biosynthesis of enterobacterial lipopolysaccharide is growth phase regulated primarily at the transcriptional level in Escherichia coli K-12. Journal of bacteriology, 177(15), 4488–4500.

61. Biswas, T., Yi, L., Aggarwal, P., Wu, J., Rubin, J. R., Stuckey, J. A., Woodard, R. W., & Tsodikov, O. V. (2009). The tail of KdsC: conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase. The Journal of biological chemistry, 284(44), 30594–30603.

62. Rosenow, C., Roberts, I. S., & Jann, K. (1995). Isolation from recombinant Escherichia coli and characterization of CMP-Kdo synthetase, involved in the expression of the capsular K5 polysaccharide (K-CKS). FEMS microbiology letters, 125(2-3), 159–164.

63. Willis, L. M., Stupak, J., Richards, M. R., Lowary, T. L., Li, J., & Whitfield, C. (2013). Conserved glycolipid termini in capsular polysaccharides synthesized by ATP-binding cassette transporter-dependent pathways in Gram-negative pathogens. Proceedings of the National Academy of Sciences of the United States of America, 110(19), 7868–7873.

64. Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J., & Koffas, M. A. (2014). Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules. FEMS microbiology reviews, 38(4), 660–697.

65. Lanz, N. D., Ming, S. A., Thon, V., Veeramachineni, V. M., Azurmendi, H. F., & Vann, W. F. (2021). Characterization of the β-KDO Transferase KpsS, the Initiating Enzyme in the Biosynthesis of the Lipid Acceptor for Escherichia coli Polysialic Acid. Biochemistry, 60, 25, 2044–2054.

66. Doyle, L., Ovchinnikova, O. G., Myler, K., Mallette, E., Huang, B. S., Lowary, T. L., Kimber, M. S., & Whitfield, C. (2019). Biosynthesis of a conserved glycolipid anchor for Gram-negative bacterial capsules. Nature chemical biology, 15(6), 632–640.

67. Phanphak, S., Georgiades, P., Li, R., King, J., Roberts, I. S., & Waigh, T. A. (2019). Super-Resolution Fluorescence Microscopy Study of the Production of K1 Capsules by Escherichia coli: Evidence for the Differential Distribution of the Capsule at the Poles and the Equator of the Cell. Langmuir : the ACS journal of surfaces and colloids, 35(16), 5635–5646.

68. Monteiro, M. A., Baqar, S., Hall, E. R., Chen, Y. H., Porter, C. K., Bentzel, D. E., Applebee, L., & Guerry, P. (2009). Capsule polysaccharide conjugate vaccine against diarrheal disease caused by Campylobacter jejuni. Infection and immunity, 77(3), 1128–1136.

69. Huddleston, J. P., Anderson, T. K., Girardi, N. M., Thoden, J. B., Taylor, Z., Holden, H. M., & Raushel, F. M. (2021). Biosynthesis of D-glycero-L-gluco-Heptose in the Capsular Polysaccharides of Campylobacter jejuni. Biochemistry, 60(19), 1552–1563.

70. Huddleston, J. P., & Raushel, F. M. (2019). Biosynthesis of GDP-D-glycero-α-D-manno-heptose for the Capsular Polysaccharide of Campylobacter jejuni. Biochemistry, 58(37), 3893–3902.

71. Huddleston, J. P., & Raushel, F. M. (2020). Functional Characterization of Cj1427, a Unique Ping-Pong Dehydrogenase Responsible for the Oxidation of GDP-D-glycero-α-D-manno-heptose in Campylobacter jejuni. Biochemistry, 59(13), 1328–1337.

72. Huddleston, J. P., Anderson, T. K., Spencer, K. D., Thoden, J. B., Raushel, F. M., & Holden, H. M. (2020). Structural Analysis of Cj1427, an Essential NAD-Dependent Dehydrogenase for the Biosynthesis of the Heptose Residues in the Capsular Polysaccharides of Campylobacter jejuni. Biochemistry, 59(13), 1314–1327.

73. McCallum, M., Shaw, G. S., & Creuzenet, C. (2013). Comparison of predicted epimerases and reductases of the Campylobacter jejuni D-altro- and L-gluco-heptose synthesis pathways. The Journal of biological chemistry, 288(27), 19569–19580.

74. Xiang, D. F., Thoden, J. B., Ghosh, M. K., Holden, H. M., & Raushel, F. M. (2022). Reaction Mechanism and Three-Dimensional Structure of GDP-D-glycero-α-D-manno-heptose 4,6-Dehydratase from Campylobacter jejuni. Biochemistry, 61(13), 1313–1322.

75. Taylor, Z. W., Brown, H. A., Holden, H. M., & Raushel, F. M. (2017). Biosynthesis of Nucleoside Diphosphoramidates in Campylobacter jejuni. Biochemistry, 56(46), 6079–6082.

76. Taylor, Z. W., Chamberlain, A. R., & Raushel, F. M. (2018). Substrate Specificity and Chemical Mechanism for the Reaction Catalyzed by Glutamine Kinase. Biochemistry, 57(37), 5447–5455.

77. Taylor, Z. W., & Raushel, F. M. (2019). Manganese-Induced Substrate Promiscuity in the Reaction Catalyzed by Phosphoglutamine Cytidylyltransferase from Campylobacter jejuni. Biochemistry, 58(16), 2144–2151.

78. Taylor, Z. W., & 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(15), 2238–2244.

79. Taylor, Z. W., Brown, H. A., Narindoshvili, T., Wenzel, C. Q., Szymanski, C. M., Holden, H. M., & 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. Journal of the American Chemical Society, 139(28), 9463–9466.

80. 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., & Szymanski, C. M. (2007). Commonality and biosynthesis of the O-methyl phosphoramidate capsule modification in Campylobacter jejuni. The Journal of biological chemistry, 282(39), 28566–28576.

81. Sørensen, M. C., van Alphen, L. B., Harboe, A., Li, J., Christensen, B. B., Szymanski, C. M., & Brøndsted, L. (2011). Bacteriophage F336 recognizes the capsular phosphoramidate modification of Campylobacter jejuni NCTC11168. Journal of bacteriology, 193(23), 6742–6749.

82. Brown, S., Zhang, Y. H., & Walker, S. (2008). A revised pathway proposed for Staphylococcus aureus wall teichoic acid biosynthesis based on in vitro reconstitution of the intracellular steps. Chemistry & biology, 15(1), 12–21.

83. Lairson, L. L., Henrissat, B., Davies, G. J., & Withers, S. G. (2008). Glycosyltransferases: structures, functions, and mechanisms. Annual review of biochemistry, 77, 521–555.

84. Campbell, J. A., Davies, G. J., Bulone, V., & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. The Biochemical journal, 326 (Pt 3)(Pt 3), 929–939.

85. Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. Nucleic acids research, 50(D1), D571–D577.

86. Coutinho, P. M., Deleury, E., Davies, G. J., & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. Journal of molecular biology, 328(2), 307–317.

87. Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2006). Insights into the synthesis of lipopolysaccharide and antibiotics through the structures of two retaining glycosyltransferases from family GT4. Chemistry & biology, 13(11), 1143–1152.

88. Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., & Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. Glycobiology, 16(2), 29R–37R.

89. Charnock, S. J., & Davies, G. J. (1999). Structure of the nucleotide-diphospho-sugar transferase, SpsA from Bacillus subtilis, in native and nucleotide-complexed forms. Biochemistry, 38(20), 6380–6385.

90. Arrecubieta, C., Hammarton, T. C., Barrett, B., Chareonsudjai, S., Hodson, N., Rainey, D., & Roberts, I. S. (2001). The transport of group 2 capsular polysaccharides across the periplasmic space in Escherichia coli. Roles for the KpsE and KpsD proteins. The Journal of biological chemistry, 276(6), 4245–4250.

91. Karlyshev, A. V., Linton, D., Gregson, N. A., Lastovica, A. J., & Wren, B. W. (2000). Genetic and biochemical evidence of a Campylobacter jejuni capsular polysaccharide that accounts for Penner serotype specificity. Molecular microbiology, 35(3), 529–541.

92. Sternberg, M. J., Tamaddoni-Nezhad, A., Lesk, V. I., Kay, E., Hitchen, P. G., Cootes, A., van Alphen, L. B., Lamoureux, M. P., Jarrell, H. C., Rawlings, C. J.,

Soo, E. C., Szymanski, C. M., Dell, A., Wren, B. W., & Muggleton, S. H. (2013). Gene function hypotheses for the Campylobacter jejuni glycome generated by a logic-based approach. Journal of molecular biology, 425(1), 186–197.

93. Kelly, S. D., Williams, D. M., Nothof, J. T., Kim, T., Lowary, T. L., Kimber, M. S., & Whitfield, C. (2022). The biosynthetic origin of ribofuranose in bacterial polysaccharides. Nature chemical biology, 18(5), 530–537.

2. FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF THE UDP-GLUCOSE DEHYDROGENASE INVOLVED IN CAPSULAR POLYSACCHARIDE BIOSYNTHESIS FROM *CAMPYLOBACTER JEJUNI*[*]

2.1. Introduction

Campylobacter jejuni is *Campylobacter jejuni* is a zoonotic organism commonly found in the intestinal tracts of poultry, cattle, and dogs. *C. jejuni* infection in humans leads to campylobacteriosis, which is the leading cause of gastroenteritis in the United States, accounting for 1.3 million new cases per year (1). Most commonly, *C. jejuni* infection is brought on by consumption of raw or undercooked poultry, unsanitary drinking water, unpasteurized milk, or pet feces (2). *C. jejuni* is also implicated in Guillain-Barré syndrome (GBS); ~40% of new GBS cases are preceded by a *C. jejuni* infection (3,4). Attached to the exterior of *C. jejuni* is a polymeric coat of sugar molecules known as the capsular polysaccharide (CPS). The sequence of monosaccharides that comprise the CPS in *C. jejuni* can be quite diverse among the different strains of this bacterium, and this diversity contributes to the evasion of the host immune system and epithelial attachment. More than 60 different strains of *C. jejuni* are known to exist (5,6).

[*] Reprinted with permission from "Functional and Structural Characterization of the UDP-Glucose Dehydrogenase Involved in Capsular Polysaccharide Biosynthesis from *Campylobacter jejuni*" by Alexander S. Riegert and Frank M. Raushel, *Biochemistry* (2021) *60* (9), pp 725-734, Copyright 2021 American Chemical Society

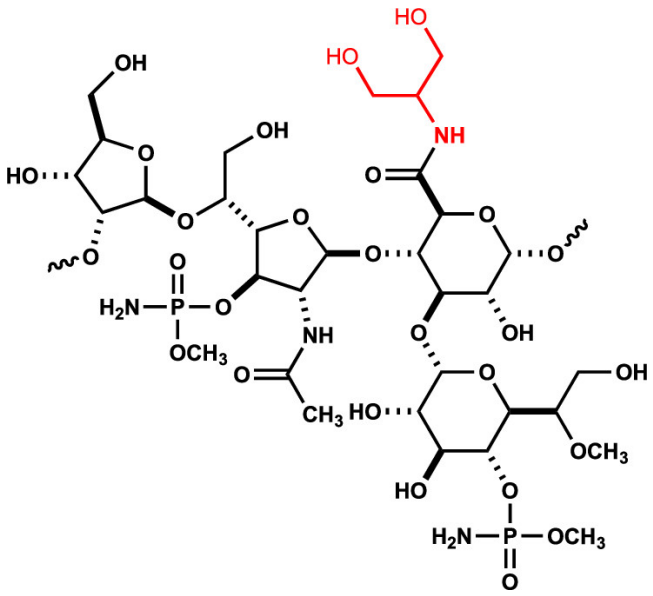


Figure 4 Structure of the repeating unit in the capsular polysaccharide of *C. jejuni* NCTC 11168. The amide attached to the glucuronamide moiety is colored red.

The CPS from *C. jejuni* NCTC 11168 is comprised of D-*glycero*-L-*gluco*-heptose, D-glucuronic acid, D-*N*-acetyl-galactose, and D-ribose (Figure 4) (7,8). These sugars are further decorated by methylation, phosphoramidylation, and amidylation. The catalytic activities of four enzymes (Cj1415–Cj1418) that are partially responsible for the phosphoramidate modification in *C. jejuni* NCTC 11168 have been elucidated (9−13). We, and others, have functionally characterized seven other enzymes that are required for the biosynthesis of the D-*glycero*-L-*gluco*-heptose moiety (Cj1152, Cj1423–Cj1425, Cj1427, Cj1428, and Cj1430) in *C. jejuni* NCTC 11168 (14−17). Knockout studies have shown that when the gene for Cj1423 from *C. jejuni* NCTC 11168 is deleted, the heptose sugar is no longer present in the capsular polysaccharide (7,18,19). In the same study, a knockout of the gene for Cj1441 resulted in a complete loss of the capsule. The gene cluster that encodes most, but not all, of the

enzymes required for the biosynthesis of the CPS of *C. jejuni* NCTC 11168 is presented in Figure 5.

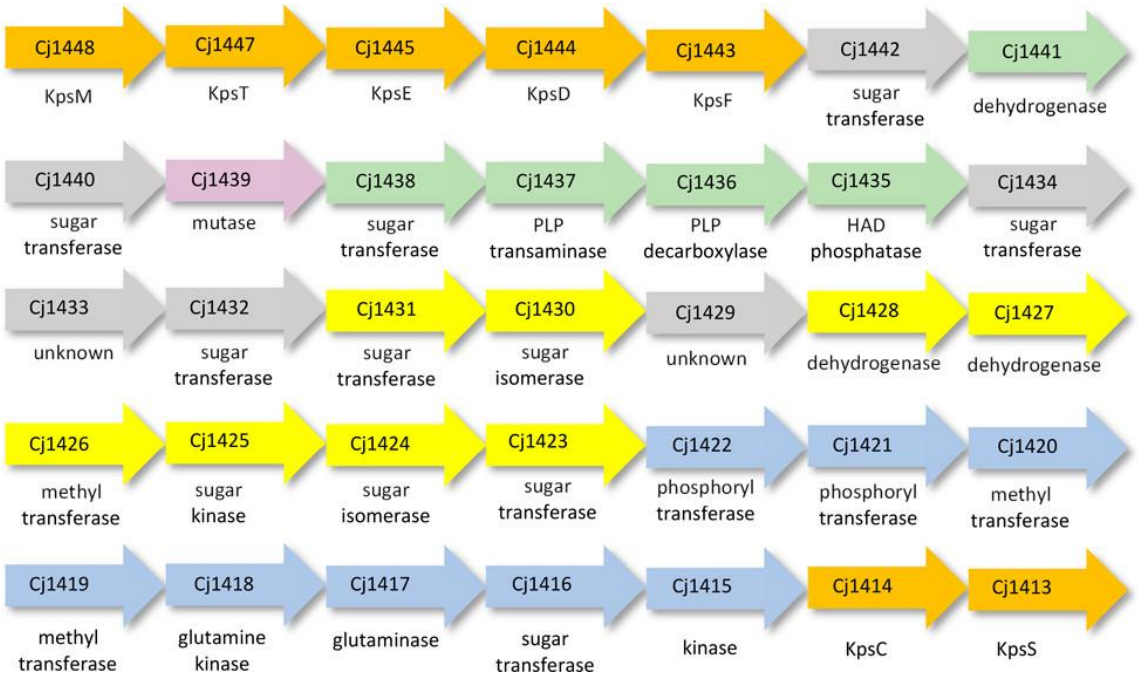


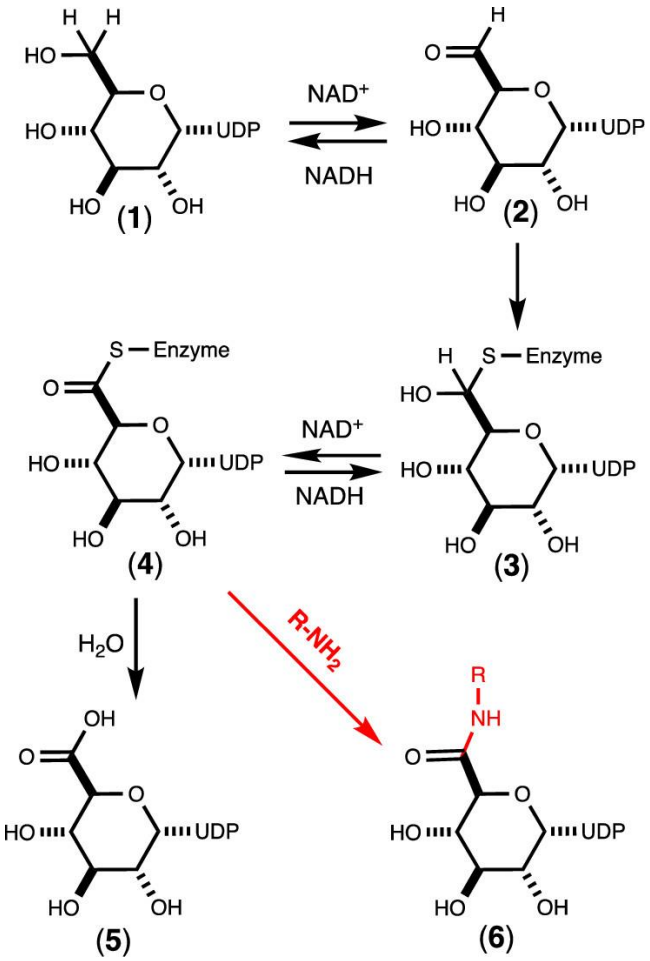
Figure 5 Gene cluster for the biosynthesis of the CPS in *C. jejuni* NCTC 11168. Enzymes of unknown function are colored gray. Enzymes responsible for the biosynthesis of the glucuronamide, heptose, and galactosyl moieties are colored green, yellow, and purple, respectively. Enzymes required for the methyl phosphoramidate modification are colored blue. Enzymes necessary for the export of the capsule are colored orange.

Unfortunately, relatively little is known about the biosynthesis of the glucuronamide moiety in the CPS of the NCTC 11168 (HS:2) and NCTC 12517 (HS:19) strains of *C. jejuni*. In strain NCTC 11168, the glucuronamide moiety is formed from either serinol (2-amino-1,3-propanediol) or ethanolamine, (7,20) whereas in NCTC 12517, the amide bond is formed with only serinol (21). The enzymes most likely involved in the biosynthesis and formation of the glucuronamide moiety to the growing

polysaccharide chain include Cj1435–Cj1438 and Cj1441 for HS:2. The genes for these five enzymes are found in the CPS gene cluster of strain NCTC 11168, and close homologues (except for Cj1436) are found in strain NCTC 12517 (HS:19) (22). Cj1435 is predicted to be a haloacid dehalogenase (HAD) phosphatase. Cj1436 and Cj1437 are predicted to be pyridoxal phosphate (PLP)-containing enzymes. Cj1438 is likely a glycosyltransferase. Cj1441 is a putative UDP-glucose 6-dehydrogenase.

UDP-glucose 6-dehydrogenase from *Streptococcus pyogenes* catalyzes the double NAD^+-dependent oxidation of UDP-D-glucose to UDP-D-glucuronic acid (23−26). The reaction mechanism has been shown to occur by the initial oxidation of UDP-D-glucose (**1**) to the corresponding aldehyde intermediate (**2**) at C6 (24). The aldehyde intermediate is tightly bound and not released to solution. The second oxidation step proceeds with the addition of a cysteine thiolate to form a thiohemiacetal intermediate (**3**) that is followed by a second hydride transfer to form a thioester intermediate (**4**) that is subsequently hydrolyzed by an activated water molecule to form the ultimate product, glucuronic acid (**5**). This transformation is summarized in Scheme 17.

Scheme 17 Mechanism for the reaction catalyzed by UDP-Glucose 6-Dehydrogenase[a]



[a] Shown with the red arrow is a potential mechanism for the formation of the UDP-glucuronamide (**6**) via nucleophilic attack of an amine with the thioester intermediate.

The source of the primary amine necessary for amide bond formation in the CPS of *C. jejuni* is likely controlled by the activity of the two PLP-containing enzymes (Cj1437 and Cj1436). PLP can be used in a number of reaction types, including transamination, decarboxylation, and racemization (27). It is hypothesized here that Cj1436 uses PLP to decarboxylate L-serine or L-serine phosphate to form ethanolamine

or ethanolamine phosphate, while Cj1437 functions to transaminate dihydroxyacetone phosphate (DHAP) to form serinol phosphate (28,29). In the absence of an obvious amide bond-forming enzyme in the putative gene cluster for the biosynthesis of the CPS of *C. jejuni*, one must look at alternative chemical strategies. One potential scenario is the nucleophilic attack of an amine substrate with the thioester intermediate that is postulated to occur during the reaction catalyzed by UDP-glucose 6-dehydrogenase to form the corresponding UDP-glucuronamide (**6**) as illustrated in Scheme 17 (25,30). This transformation could occur instead of the "normal" attack by an activated water molecule. This reaction is similar to that catalyzed by glyceraldehyde 3-phosphate dehydrogenases, where a thioester intermediate is attacked by phosphate to form 1,3-diphosphoglycerate. However, the enzyme directly responsible for amide bond formation may occur outside of the gene cluster for formation of the CPS in *C. jejuni*.

To test our proposal for formation of the glucuronamide moiety of the CPS from *C. jejuni*, we have cloned and expressed the gene for Cj1441. The enzyme has been purified and crystallized, and the three-dimensional structure determined. The catalytic properties of the enzyme have been determined in the presence and absence of amine substrates with NAD^+ and UDP-D-glucose.

2.2. Methods

2.2.1. Cloning, Expression, and Purification of Cj1441

The gene for Cj1441 (UniProt entry Q0P8H3) was cloned from the genomic DNA of *C. jejuni* NCTC 11168 (ATCC-700819D-5). Expression tests of the cloned gene for Cj1441 exhibited low levels of expression, so the gene for Cj1441 was codon optimized and synthesized by GenScript (Piscataway, NJ). The synthesized DNA served as the starting template for polymerase chain reaction using Phusion DNA polymerase (New England Biolabs). For all genes, primers were designed that incorporated NdeI and XhoI restriction sites and the resulting fragments were digested at 37 °C and ligated into a pET31b expression vector, which carried a C-terminal hexahistidine tag.

The pET31b vector was used to transform BL21 *Escherichia coli* cells (Novagen). The cells harboring the pET31b-Cj1441 plasmid were cultured in lysogeny broth with 100 mg/L ampicillin. The cells were grown at 37 °C while being shaken and induced with 1.0 mM isopropyl β-D-1-thiogalactopyranoside when the optical density reached 0.9 at 600 nm. The cells were allowed to express protein at 21 °C for 18 h after induction and then harvested by centrifugation at 15000 rcf and 4 °C. The cell pellet was then resuspended in loading buffer [50 mM HEPES/K^+, 300 mM KCl, and 20 mM imidazole (pH 8.0)] and lysed with sonication. The sonicated cells were centrifuged at 25000 rcf and 4 °C before the cell lysate was passed through a 0.45 μm filter. The sample was loaded onto a prepacked 5 mL HisTrap (GE Healthcare) nickel affinity column. The protein was eluted from the column using 50 mM HEPES/K^+ (pH 8.0), 300 mM KCl, and 250 mM imidazole over a gradient of 25 column volumes. The resulting protein was pooled and dialyzed against 10 mM HEPES/K^+ (pH 8.0) and 200 mM KCl. The protein was concentrated to 22 mg/mL and flash-frozen using liquid

nitrogen before being stored at −80 °C. Approximately 10 mg of protein was obtained per liter of cell culture.

2.2.2. Mass Spectral Analysis

Samples of the reaction catalyzed by Cj1441 were tested in 50 mM ammonium bicarbonate and allowed to incubate for 2 h at 25 °C in the presence of dithiothreitol, UDP-glucose, and NAD^+. The samples were also tested in the presence and absence of ethanolamine phosphate, ethanolamine, serinol phosphate, or serinol. The resulting solution was filtered through a GE Healthcare Vivaspin 500 10 kDa filter, and the flow-through was analyzed using a Thermo Scientific Q Exactive Focus mass spectrometer in ESI (negative) mode.

2.2.3. Reaction Stoichiometry

To determine the stoichiometry of the Cj1441-catalyzed reaction, solutions consisting of 1.0 μM Cj1441, 2.0 mM NAD^+, 1.0 mM DTT, 50 mM triethanolamine (pH 8.7), and variable concentrations of UDP-glucose from 10 to 100 μM were incubated for 60 min to enable the reaction to go to completion. The concentration of NADH was determined from the absorbance at 340 nm.

2.2.4. Determination of Kinetic Constants

The kinetic constants for Cj1441 were determined by following the reduction of NAD^+ to NADH at 340 nm and 25 °C with a Spectramax340 ultraviolet–visible

spectrophotometer. Assays were performed in 50 mM triethanolamine/K^+ (pH 8.7), 200 mM KCl, and 1.0 mM DTT in 96-well NucC plates. The reaction was initiated with the addition of 150 nM Cj1441 to the reaction well. Kinetic constants were determined with a fixed level of NAD^+ (2.0 mM) and variable levels of UDP-glucose (0.01–2.0 mM) as well as a fixed level of UDP-glucose (5.0 mM) and varying levels of NAD^+ (0.05–4.0 mM). The kinetic parameters were determined by fitting the initial rates to eq 1 using GraFit 5(1) where v is the initial velocity of the reaction, E_t is the enzyme concentration, k_{cat} is the turnover number, [A] is the substrate concentration, and K_m is the Michaelis constant.

$$v / E_t = k_{cat} [A] / (A + K_m) \quad (1)$$

2.2.5. Sequence Similarity Network for COG1004

The identifiers were used to generate a sequence similarity network using the Enzyme Function Initiative-Enzyme similarity tool and Option D (31,32). The full network was downloaded and used to create a sequence similarity network using Cytoscape 3.7.2 (33). The edges of the network were set to a 45% sequence identity using the Cytoscape filter function.

2.2.6. Crystallization of Cj1441

Crystallization conditions were initially surveyed by the hanging drop method of vapor diffusion using a sparse matrix screen from Wizard 3+4 purchased from Molecular Dimensions. The wild-type enzyme was tested in complex with 5.0 mM

UDP-glucose and 5.0 mM NAD^+ at room temperature. X-ray diffraction quality crystals appeared after 2–3 days and were subsequently grown by mixing, in a ratio of 1:1, the protein sample at 23 mg/mL and the crystallization buffer. The initial crystallization conditions were 30% pentaerythritol ethoxylate, 50 mM ammonium sulfate, and 50 mM Bis-Tris (pH 6.5) at 21 °C. These preliminary crystals were tested for diffraction, before being optimized. The crystallization conditions used to grow the samples for data collection were 5.0 mM UDP-glucose, 5.0 mM NAD^+, 1.0 mM DTT, 32% pentaerythritol ethoxylate, 50 mM ammonium sulfate, 50 mM Bis-Tris (pH 6.5), and 200 mM NaCl. The crystals used for data collection were roughly 50 μm × 50 μm in size.

Diffraction data were collected at the Stanford Synchrotron Radiation Light Source (SSRL), beamline 14-1, on an Eiger 16M detector at 100 K. The data were integrated, scaled, and merged (34). Merged observed intensities were then imported into the Phenix reflection file editor, where an R_{free} set was assigned using the default settings (35). Relevant X-ray data collection statistics are listed in Table 3. The unit cell was determined to be $P2_1$ with the following cell dimensions: a = 43.79 Å, b = 148.78 Å, c = 62.34 Å, α = 90.0°, β = 107.5°, and γ = 90.0°. The structure was determined using PHASER from the PHENIX suite and using Protein Data Bank (PDB) entry 1DLJ as the search model with 49% sequence identity (36,37). The structure was built using COOT, and iterative rounds of refinement using PHENIX refine in the PHENIX suite reduced R_{work} and R_{free} to 19.0% and 25.6%, respectively, from 46 to 2.09 Å resolution (35,38). Refinement statistics are listed in Table 3.

Table 3 X-ray Data Collection Statistics and Model Refinement Statistics of Cj1441 with NAD^+ and UDP-Glucose

	Cj1441 with NAD and UDP-glucose
resolution limits	46.4 - 2.09 (2.17 – 2.09)[b]
Space Group	$P2_1$
Unit Cell *a* (Å) *b* (Å) *c* (Å) α (°) β (°) γ (°)	 43.790 148.780 62.340 90.00 90.00 107.52
number of independent reflections	43332 (4336)
completeness (%)	96.6 (97.9)
redundancy	12.8 (4.1)
avg I/avg σ(I)	11.3 (1.43)
R_{sym} (%)[a]	4.1 (18.4)
resolution limits (Å)	46.4 - 2.09
[c]*R*-factor (overall)%/no. reflections	19.1/54971
R-factor (working)%/no. reflections	18.9/ 43332
R-factor (free)%/no. reflections	24.9/ 4336
number of protein atoms	6350
number of ligands (NAD or UDP-glucose)	4
number of water molecules	249
average B values ($Å^2$)	
protein atoms	48.7
ligand NAD^+, UDP-glc	37.8, 47.7
solvent	46.4
weighted RMS deviations from ideality	
bond lengths (Å)	0.008
bond angles (°)	0.99
general planes (°)	0.005
Ramachandran regions (%)	
most favored	97.7
additionally allowed	2.1
generously allowed	0.3

[a]$R_{sym} = (\sum|I - |/ \sum I) \times 100$.

[b]Statistics for the highest resolution bin.
[c] R-factor = ($\Sigma|Fo - Fc|/\Sigma|Fo|$) × 100 where F_0 is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.

2.3. Results

2.3.1. Sequence Similarity Network for UDP-Glucose Dehydrogenase

Cj1441 is a member of COG1004, and the sequence similarity network (SSN) is presented in Figure 6 at a percent identity cutoff of 45% (32). Cj1441 clusters with other known UDP-glucose 6-dehydrogenases, including those from *S. pyogenes*, and *E. coli*. The enzymes denoted as UDP-mannose dehydrogenases segregate with one another in a group that is separated from those enzymes that have been shown to catalyze the oxidation of UDP-glucose. Of the 2883 enzymes in COG1004, ~600 of them cluster in the same group as Cj1441. In this group are examples from other pathogenic bacteria, including *E. coli* O157:H7 and *E. coli* O6:H1. Cj1441 and the UDP-glucose 6-dehydrogenases from *S. pyogenes* and *E. coli* O157:H7 exhibit 47% and 51% sequence identity, respectively.

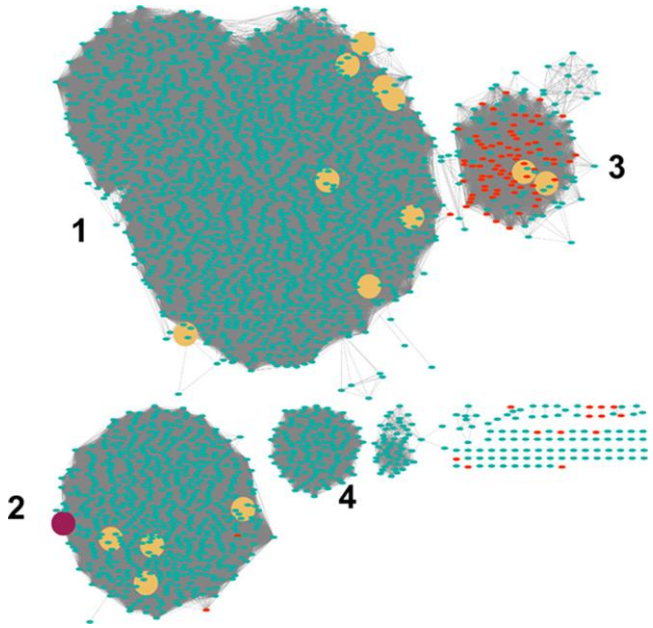


Figure 6 Sequence similarity network for COG1004. Each node in the network represents a single sequence, and each edge (depicted as lines) represents the pairwise connection between two sequences at a sequence identify better than 45%. The maroon-colored circle represents Cj1441 in group 2. Group 2 contains the UDP-glucose 6-dehydrogenases from *E. coli* 0157:H7, *E. coli* K12, *E. coli* O6:H1, and *S. pyogenes*. The red circles are examples annotated as mannose 6-dehydrogenases. The yellow circles are functionally characterized enzymes from uniprot.org.

2.3.2. Kinetic Analysis of Cj1441

Cj1441 was shown to catalyze the reduction of NAD^+ to NADH in the presence of excess UDP-glucose by following the change in absorbance at 340 nm. To measure the reaction stoichiometry, Cj1441 was mixed with excess NAD^+ (2.0 mM) and variable amounts of UDP-glucose (10–100 μM). The reaction was followed to completion, and the amount of NADH formed was determined from the change in the absorbance at 340 nm. On average, the ratio of NADH formed at equilibrium with the initial UDP-glucose concentration was determined to be 2:1. This value matches the predicted value from the reaction mechanism presented in Scheme 17.

The apparent kinetic constants for the oxidation of UDP-glucose by Cj1441 were determined by variation of the UDP-glucose concentration at a fixed concentration of NAD^+ (5.0 mM) and then variation of the NAD^+ concentration at a fixed concentration of UDP-glucose (5.0 mM). The plot for the variation of UDP-glucose is presented in Figure 7. The apparent k_{cat} values are 0.80–1.1 s^{-1}, with k_{cat}/K_m values of 4.2×10^3 and 4.8×10^3 M^{-1} s^{-1} for the variation of UDP-glucose and NAD^+, respectively. To determine whether the addition of potential amine substrates to the reaction mixtures could alter the rate of NAD^+ reduction, UDP-glucose was varied at a fixed concentration of NAD^+ and the apparent kinetic constants determined for the oxidation of UDP-glucose. Within experimental error, the addition of either ethanolamine, ethanolamine phosphate, serinol, or serinol phosphate had no measurable effect on the catalytic properties of Cj1441. The apparent kinetic constants are listed in Table 4.

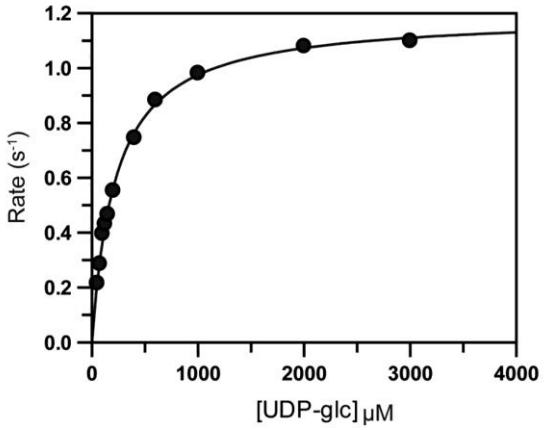


Figure 7 Michaelis–Menten plot for the oxidation of UDP-glucose catalyzed by Cj1441 at 25 °C. The reaction mixture contained saturating conditions of NAD^+ (2.0 mM) and 1.0 mM DTT. The solid line represents the fit of the data to eq 1.

Table 4 Steady State Kinetic Constants for Cj1441[a]

fixed substrate	variable substrate (mM)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}\ s^{-1}$)
NAD^+ (5.0 mM)	UDP-glucose (0.05 – 3.0)	1.10 ± 0.02	210 ± 8.0	4800 ± 180
UDP-glucose (5.0 mM)	NAD^+ (0.05 – 3.0)	0.80 ±0.02	190 ±12	4200 ± 290
ethanolamine (2.0 mM) NAD^+ (5.0 mM)	UDP-glucose (0.05 – 3.0)	0.90 ± 0.04	304 ± 14	3000 ± 190
ethanolamine-P (2.0 mM) NAD^+ (5.0 mM)	UDP-glucose (0.05 – 3.0)	0.80 ±0.03	210 ±11	3800 ±210
serinol (2.0 mM) NAD^+ (5.0 mM)	UDP-glucose (0.05 – 3.0)	0.80 ±0.03	230 ±15	3500 ± 240
serinol phosphate (2.0 mM) NAD^+ (5.0 mM)	UDP-glucose (0.05 – 3.0)	0.80 ± 0.03	200 ± 11	4000 ± 220

[a] At pH 8.7 and 25 °C.

Table 4 Steady State Kinetic Constants for Cj1441[aa] At pH 8.7 and 25 °C.

2.3.3. Verification of Reaction Products

The products of the reaction catalyzed by Cj1441 were analyzed by negative ion ESI mass spectrometry. As a control, commercial UDP-glucuronic acid was shown to have an *m*/*z* value of 579.02, identical to the calculated *m*/*z* value of 579.02. Cj1441 was subsequently incubated with NAD^+, DTT, and UDP-glucose, and the products were analyzed similarly with *m*/*z* values of 565.04 (UDP-glucose), 579.02 (UDP-glucuronate), and 664.11 (NADH). Amidated products were not observed in any of the mass spectra

when ammonia, ethanolamine, serinol, ethanolamine phosphate, or serinol phosphate was added to the reaction mixtures (Appendix Figures S1–S6).

2.3.4. Crystallization and Determination of the Structure of Cj1441

Cj1441 crystallized as a dimer in the asymmetric unit, and its structure was determined by molecular replacement using UDP-glucose 6-dehydrogenase from *S. pyogenes* (PDB entry 1DLJ) as the search model. Cj1441 co-crystallized with UDP-glucose and NAD^+ in the active site, and the structure was determined to a resolution of 2.09 Å. The Cj1441 monomer contains 393 amino acids, and the dimeric form of the enzyme is presented in Figure 8. The enzyme consists of two major domains, each containing an α/β fold. The N-terminal domain contains a Rossmann fold bearing the conserved GXGXXG motif between residues 7 and 12, in addition to six parallel β-strands arranged in the canonical order [3, 2, 1, 4, 5, 6]. The two subunits were aligned using Pymol, and they superimpose with a root-mean-square deviation of 0.3 Å (39). The overall buried surface area between the dimeric interface was calculated using PISA to be 2504 $Å^2$ (http://pdbe.org/pisa). The two α/β domains of the protein are linked to one another by a 48 Å long central α-helix from residue Asp191 to Asn223. The N-terminal domain (residues 1–191) is responsible for binding the NAD^+ cofactor. The C-terminal domain (residues 223–393) binds UDP-glucose using a similar α/β fold. The defining difference between the two domains of the protein is the lack of the Ω loop from residue 301 to 317. A structural comparison of Cj1441 with the UDP-glucose 6-dehydrogenase from *S. pyogenes* (PDB entry 1DLJ) is shown in Appendix Figure S7.

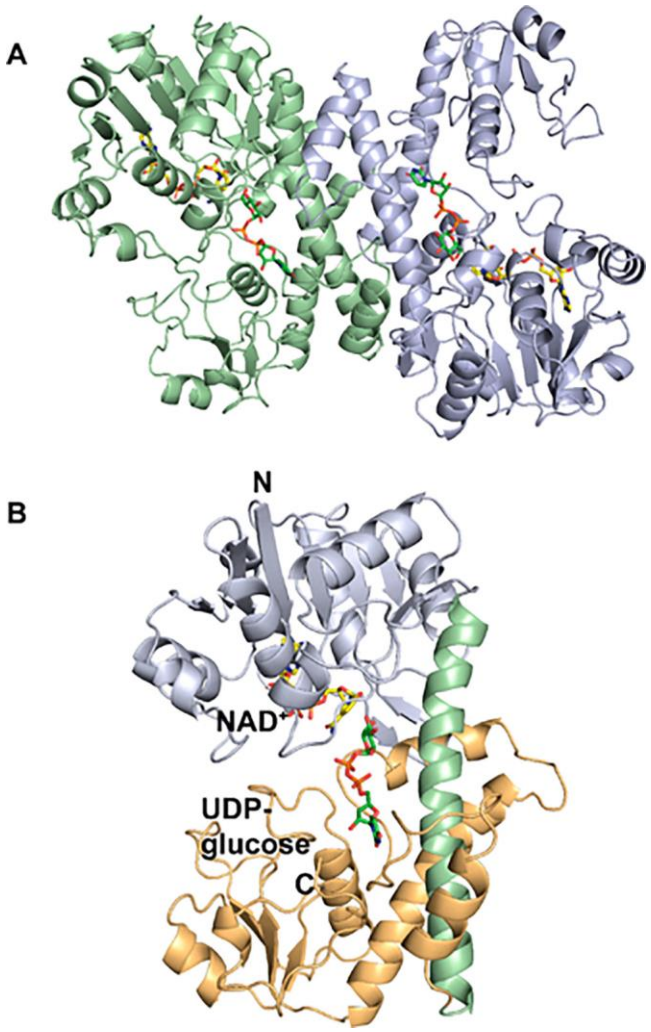


Figure 8 (A) Ribbon diagram of the Cj1441 homodimer. Chain A is colored pale green, and chain B is colored pale blue. UDP-glucose is colored green, and NAD^+ is colored yellow. (B) Ribbon diagram of the Cj1441 monomer in which the N-terminal domain is colored pale blue and the C-terminal domain is colored tan. The α helix at the dimer interface is colored light green.

2.3.5. Binding Site for NAD^+

The N-terminal domain of the protein is primarily responsible for binding the NAD^+ cofactor. The binding site for NAD^+ is marked by the β1−α1 turn of the Rossmann fold that directly interacts with the adenine ribose. With the exception of a

small portion of the adenine ring, which is buried in the interior of the protein, the vast majority of the cofactor is exposed to solvent. This may enable the cofactor to be readily exchangeable, which is consistent with the 2-fold oxidation mechanism (40). Additionally, the nicotinamide ring is in the *syn* conformation, as determined previously for the enzyme from *S. pyogenes* (PDB entry 1DLI) and *Klebsiella pneumoniae* (PDB entry 3PLR) (37,40). The residues directly responsible for binding the coenzyme are Val11, Gly12, Asp31, Lys36, Ala59, Thr85, Thr120, and Arg316. Additionally, four ordered waters interact with the cofactor to help orient and bind the cofactor to the protein surface. UDP-glucose points directly toward the *si* face of the nicotinamide ring. There are six residues that are apparently conserved across all UDP-glucose 6-dehydrogenases involved in NAD^+ binding (41). These include the three glycine residues from the Rossmann fold, a Thr83/Pro84 pair that interacts with the adenine ring, and Arg316, which forms a salt bridge with the pyrophosphate of the nucleotide (Figure 9A).

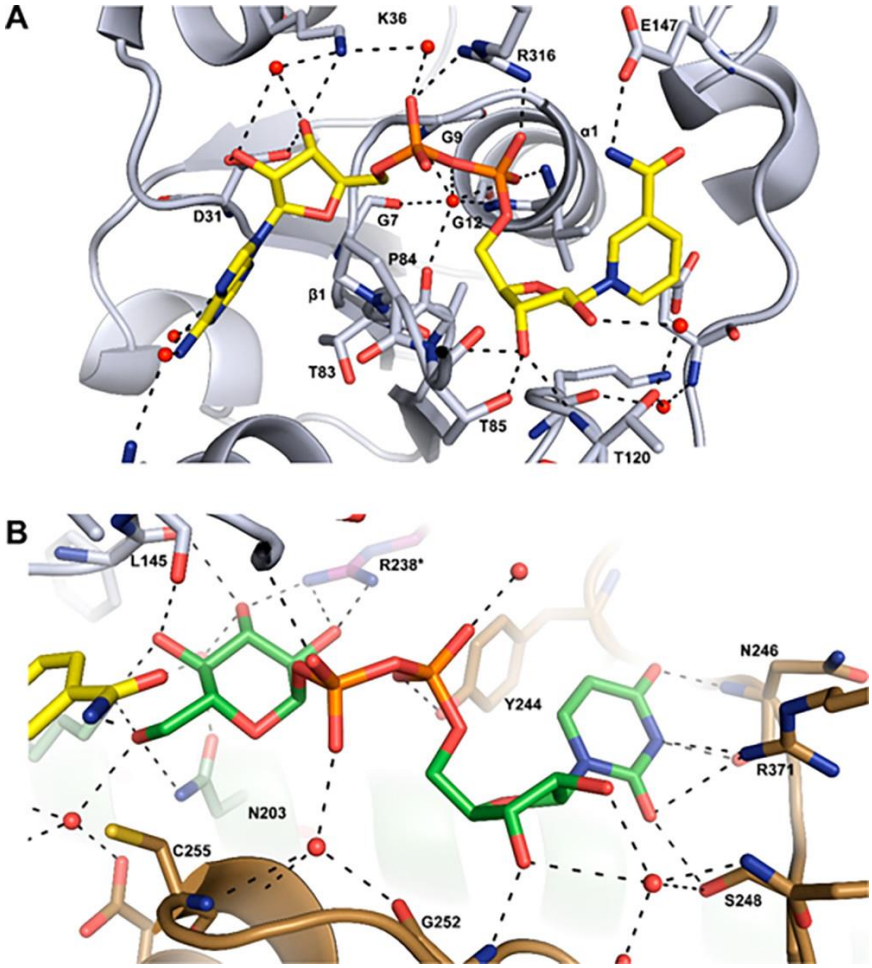


Figure 9 Cartoon representations of the binding sites for NAD^+ and UDP-glucose. Potential hydrogen bonding interactions (<3.2 Å) are colored black. (A) Binding site for NAD^+ (yellow) in the N-terminal domain of the protein. (B) Binding site for UDP-glucose (green) in the C-terminal domain of the protein. Additional details are provided in the text.

2.3.6. Binding of UDP-Glucose

UDP-glucose primarily interacts with the C-terminal domain of Cj1441. The protein crystallized at pH 6.5. This explains, in part, why UDP-glucose, rather than UDP-glucuronate, is present in the active site because the reaction equilibrium favors the substrates at low pH (Appendix Figure S8). Cys255 is located 3.6 Å from C6 of UDP-

glucose, which is sufficient to facilitate nucleophilic attack with the aldehyde intermediate and formation of a thioester intermediate during the catalytic cycle (Scheme 17) (25). Additionally, C4 of the nicotinamide ring is positioned 3.2 Å from C6 of UDP-glucose to facilitate hydride transfer (Figure 10). UDP-glucose is held in place by interactions with Asn203, Leu145, Lys199, Glu147, Tyr244, Gly252, Asn246, Ser248, and Arg371. Arg238 from the adjacent subunit is found on a flexible loop near the dimeric interface and is capable of interacting with the C2′ and C3′ hydroxyl groups of the UDP-glucose (Figure 9B).

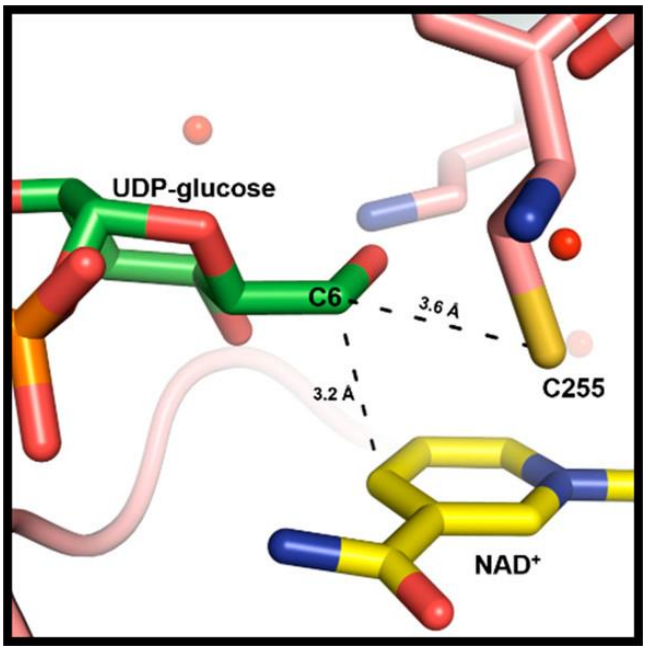


Figure 10 Ligand interactions between NAD^+ and UDP-glucose in the active site of Cj1441. The nicotinamide portion of NAD^+ is colored yellow, and the glucose portion of UDP-glucose is colored green. The catalytic cysteine is 3.6 Å from C6 of the glucose moiety. NAD^+ is 3.2 Å from C6 of the glucose moiety of the substrate.

Binding of the uridine nucleotide moiety of UDP-glucose is facilitated by the C-terminal domain through backbone interactions with residues 244–252. Of the six

hydrogen bonds formed with the UMP portion of the substrate, only two of them are from amino acid side chains (Ser248 and Tyr244). The rest of the interactions are made through hydrogen bond interactions with the carbonyl oxygens and amide nitrogens of the protein backbone. The glucose 1-phosphate binding region is rather compact compared to the UMP binding region, consisting of only four residues (Phe144–Glu147). There is an additional hydrogen bond provided by Asn203 on the central helix.

2.3.7. Other Amide Bond-Forming Enzymes

A bioinformatic search was conducted in an attempt to find additional candidates for the amide bond-forming enzyme within the CPS of *C. jejuni*. The initial probe focused on the LOS region of the *C. jejuni* genome. Specifically, we searched for regions of genes that might contain a ligase domain that would be capable of catalyzing amide bond formation. Generally, these bonds are formed through the use of an ATP-grasp domain, so our investigation focused on finding genes that contained such a region. It was subsequently determined that the LOS of *E. coli* L19 and *Shigella boydii* type 8 contained a glucuronamide moiety with serinol (42−44). The appropriate gene clusters within these organisms contained a phosphatase (WfdQ), a PLP-dependent transaminase (WfdP), and a putative transferase with an ATP-grasp domain (WfdR). This discovery led to the interrogation of the sugar transferases of unknown specificities in the capsular polysaccharide cluster of *C. jejuni*. We identified a putative glycosyltransferase (Cj1438) in the gene cluster of *C. jejuni* NCTC 11168 that also contains an ATP-grasp domain at the C-terminal half of the protein. This enzyme has all

of the necessary residues to bind ATP when compared to other ATP-grasp enzymes, including those homologues in *E. coli* L19 and *S. boydii* type 8.

2.4. Discussion

The putative UDP-glucose 6-dehydrogenase (Cj1441) from *C. jejuni* NCTC 11168 was successfully purified to homogeneity after recombinant expression in *E. coli*. We demonstrated that the enzyme requires 2 equiv of NAD^+ and that the ultimate product is UDP-D-glucuronate (**5**). The enzyme was crystallized in the presence of UDP-D-glucose and NAD^+, and the three-dimensional structure subsequently determined to a resolution of 2.09 Å. However, attempts to demonstrate amide bond formation through nucleophilic attack of added amines on the proposed thioester intermediate (**4**) were unsuccessful (Scheme 17). Therefore, it is highly unlikely that this enzyme is directly responsible for formation of the glucuronamide moiety found in the CPS of *C. jejuni* NCTC 11168 (Figure 4). However, it remains possible that amines other than the ones tested are required for amide bond formation.

How then is the glucuronamide moiety actually synthesized for the HS:2 capsule? A cursory look at the annotations of the enzymes derived from the *C. jejuni* NCTC 11168 gene cluster (Figure 5) for the biosynthesis of the CPS does not identify any obvious candidates. However, we have previously demonstrated that Cj1152 can be used for the biosynthesis of GDP-D-*glycero*-D-*manno*-heptose in *C. jejuni* NCTC 11168, and thus, genes outside of this specific gene cluster are required for the biosynthesis of the CPS (16). In this specific case, Cj1152 catalyzes the hydrolysis

of phosphate from D-*glycero*-α-D-*manno*-heptose-1,7-bisphosphate for the biosynthesis of lipooligosaccharides (LOS) in *C. jejuni* (45). However, no amide bonds are formed in the LOS of *C. jejuni*, and thus, the missing enzyme is not to be found there.

The aminoglycerol modification within the CPS of *C. jejuni* NCTC 11168 (HS:2) is, however, also found in other structurally elucidated polysaccharide capsules, including those from *C. jejuni* NCTC 12517 (HS:19), (46,47) *Vibrio cholerae* H1122, (48) *E. coli* O143, *S. boydii* type 8, (42,43) and *E. coli* L19 (44). The capsule from *C. jejuni* NCTC 12517 is presented in Figure 11. Comparison of the gene clusters for the biosynthesis of the capsular polysaccharides in *C. jejuni* NCTC 11168 (HS:2) and *C. jejuni* NCTC 12517 (HS:19) highlights the production of enzymes of nearly identical function (Table 5). These proteins include the eight enzymes known to be responsible for the *O*-methyl phosphoramidate decoration of the capsule [Cj1415–Cj1421 in *C. jejuni* NCTC 11168 (HS:2) and HS19.01–HS19.07 in *C. jejuni* NCTC 12517 (HS:19) (9−13,22,49). Except for HS19.07 and Cj1421/Cj1422, the sequence identities for these proteins are ≥97%. The other relevant homologous enzymes include Cj1435 (HAD phosphatase), Cj1437 (PLP-dependent aminotransferase), and Cj1441 (UDP-glucose 6-dehydrogenase) where the amino acid sequence identity is >55%. The two putative sugar transferases (Cj1438 and Cj1434) are less similar to their apparent counterparts for assembly of the HS19 capsule (34% and 21% identical to HS19.11 and

HS19.08, respectively). This is quite reasonable because the acceptor/donor pairs are not identical in the two capsules.

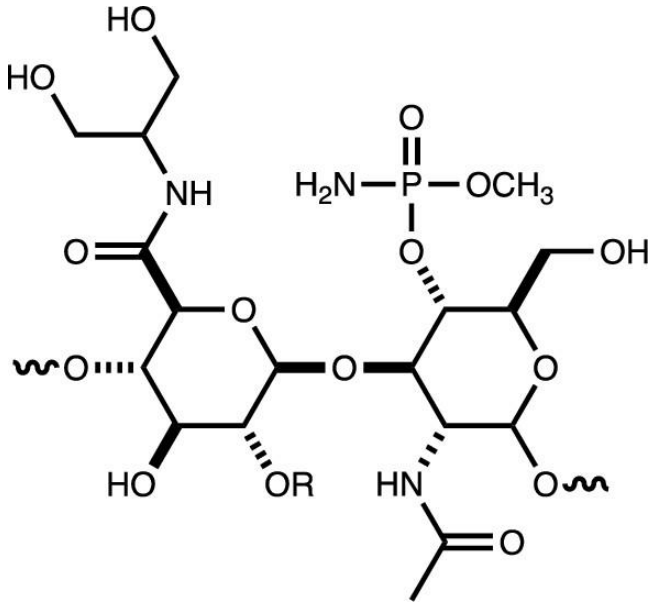


Figure 11 Structure of the repeating unit for the capsular polysaccharide from *C. jejuni* NCTC 12517 (HS:19). C2 of the glucuronamide moiety can be derivatized with L-sorbose.

Table 5 Sequence Identities in Similar Proteins for Assembly of the CPS in HS:2 and HS:19

C. jejuni NCTC 11168 (HS:2)			*C. jejuni* NCTC 12517 (HS:19)		
protein	UniProt	function	protein	UniProt	sequence identity (%)
Cj1415	Q0P8J9	kinase	HS19.01	Q6EF72	98
Cj1416	Q0P8J8	transferase	HS19.02	Q5M6N1	98
Cj1417	Q0P8J	hydrolase	HS19.03	Q5M6N0	97
Cj1418	Q0P8J6	glutamine kinase	HS19.04	Q5M6M9	98
Cj1419	Q0P8J5	methyl transferase	HS19.05	Q5M6M8	98
Cj1420	Q0P8J4	methyl transferase	HS19.06	Q5M6M7	98
Cj1421	Q0P8J3	phosphotransferase	HS19.07	Q5M6M6	47
Cj1422	Q0P8J2	phosphotransferase	HS19.07	Q5M6M6	49
Cj1434	Q0P8I0	sugar transferase	HS19.08	Q5M6M5	21
Cj1435	Q0P8H9	phosphatase	HS19.09	Q5M6M4	65
Cj1437	Q0P8H7	aminotransferase	HS19.10	Q5M6M3	55
Cj1438	Q0P8H6	sugar transferase	HS19.11	Q5M6M2	36

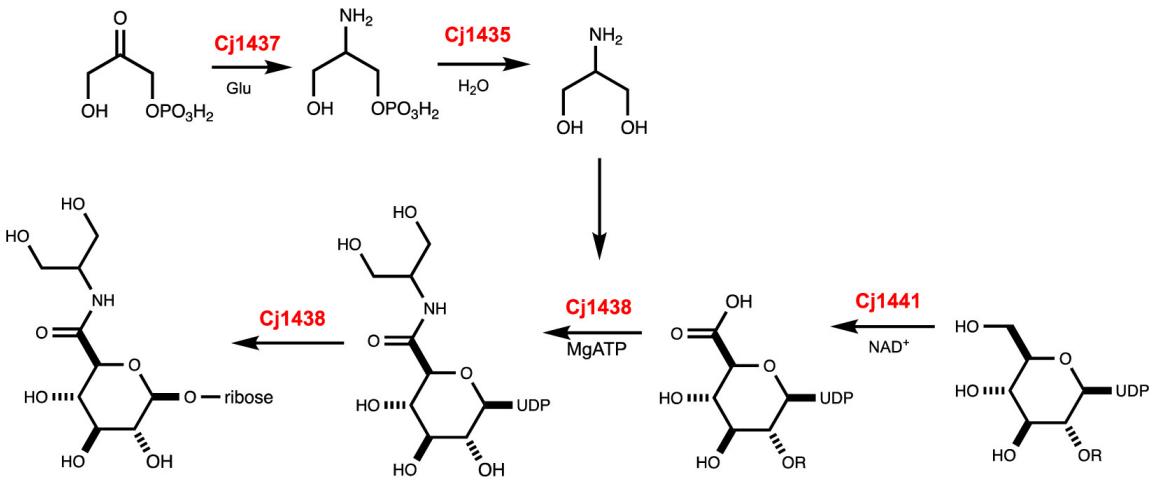
Cj1441	Q0P8H3	dehydrogenase	HS19.12	Q5M6M1	56

Upon further inspection, the putative sugar transferase (Cj1438) is particularly interesting because the N-terminal half of this protein is homologous to a large number of other functionally characterized members of the GT2 family of sugar transferases, (50−53) whereas the C-terminal half of the enzyme shows significant similarity to TupA (54). TupA is an enzyme from the ATP-grasp family that is important for amide bond formation during the synthesis of the teichuronopeptide where an amide bond is formed between glucuronate and polyglutamate (55). This observation suggests that the C-terminal half of Cj1438 harbors the catalytic machinery for amide bond formation with glucuronate. This conclusion is also supported by the high sequence identity between the C-terminal half of Cj1438 (residues 420–776) and the C-terminal half of HS19.11 (residues 471–832) and the TupA-like sequence found in the gene cluster for capsule formation in *E. coli* O143 (WfdR). The residues responsible for the binding of ATP are clearly conserved (Appendix Figure S9).

Our proposal for the assembly of the glucuronamide moiety of the HS2 capsule in *C. jejuni* NCTC 11168 utilizes the catalytic function of Cj1437 as a PLP-dependent aminotransferase to convert dihydroxyacetone phosphate to serinol phosphate. Cj1435 functions to hydrolyze the phosphorylated product to serinol, and then Cj1438 catalyzes the ATP-dependent amide bond formation between UDP-glucuronate and serinol. However, the exact order of these chemical transformations is clearly not known at this time. It is quite possible that amide bond formation does not occur until after the glucuronate has been incorporated into the growing polysaccharide and the phosphatase

does not have to function until after amide bond formation. One possibility is highlighted in Scheme 18.

Scheme 18 Proposed Reactions for the Catalytic Activities of Cj1435, Cj1437, Cj1438, and Cj1441



2.5. Conclusions

The putative UDP-glucose 6-dehydrogenase (Cj1441) from *C. jejuni* NCTC 11168 was purified to homogeneity and shown to catalyze the oxidation of UDP-D-glucose to UDP-D-glucuronate in the presence of NAD^+. The enzyme does not facilitate the nucleophilic attack of added amines with the proposed thioester intermediate, and thus, this enzyme is not responsible for formation of the glucuronamide moiety in the HS:2 capsular polysaccharide. The enzyme was crystallized in the presence of NAD^+ and UDP-glucose, and the three-dimensional structure determined to a resolution of 2.09 Å. The positioning of the hydroxyl group attached to C6 of UDP-glucose is ideal for hydride transfer to C4 of the bound NAD^+. A bioinformatic analysis of the

respective gene clusters for the biosynthesis of the HS:2 and HS:19 capsular polysaccharides in *C. jejuni* suggests that the catalytic activity for amide bond formation for assembly of the glucuronamide moiety resides in the C-terminal domain of Cj1438 from *C. jejuni* NCTC 11168.

2.6. References

1. Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., Jones, J. L., and Griffin, P. M. (2011) Foodborne illness acquired in the United States--major pathogens. *Emerging Infect. Dis. 17*, 7– 15
2. Burnham, P. M. and Hendrixson, D. R. (2018) *Campylobacter jejuni*: collective components promoting a successful enteric lifestyle. *Nat. Rev. Microbiol. 16*, 551– 565.
3. Nachamkin, I., Allos, B. M., and Ho, T. (1998) Campylobacter species and Guillain-Barré syndrome. *Clin. Microbiol. Rev. 11*, 555– 567.
4. Yuki, N. (1997) Molecular mimicry between gangliosides and lipopolysaccharides of *Campylobacter jejuni* isolated from patients with Guillain-Barré Syndrome and Miller Fisher Syndrome. *J. Infect. Dis. 176*, S150 – S153.
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. 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.

7. St. Michael, F.; Szymanski, C. M.; Li, J.;, Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A., (2002). The Structures of the Lipooligosaccharide and the Capsule Polysaccharide of *Campylobacter jejuni* Genome Sequenced Strain NCTC 11168. *Eur. J. Biochem. 269*, 5119-5136.
8. Young, K. T., Davis, L. M., and Dirita, V. J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol. 5*, 665– 679.
9. 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.
10. Taylor, Z. W., Chamberlain, A. R., and Raushel, F. M. (2018) Substrate specificity and chemical mechanism for the reaction catalyzed by glutamine kinase. *Biochemistry 57*, 5447– 5455.
11. Taylor, Z. W. and Raushel, F. M. (2019) Manganese-induced substrate promiscuity in the reaction catalyzed by phosphoglutamine cytidylyltransferase from *Campylobacter jejuni*. *Biochemistry 58*, 2144– 2151.
12. 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.
13. 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.

14. Huddleston, J. P., Anderson, T. K., Spencer, K. D., Thoden, J. B., Raushel, F. M., and Holden, H. M. (2020) Structural analysis of Cj1427, an essential NAD-dependent dehydrogenase for the biosynthesis of the heptose residues in the capsular polysaccharides of *Campylobacter jejuni*. *Biochemistry 59*, 1314– 1327.
15. Huddleston, J. P. and Raushel, F. M. (2020) Functional characterization of Cj1427, a unique ping-pong dehydrogenase responsible for the oxidation of GDP-D-*glycero*-α-D-*manno*-heptose in *Campylobacter jejuni*. *Biochemistry 59*, 1328– 1337.
16. Huddleston, J. P. and Raushel, F. M. (2019) Biosynthesis of GDP-D-*glycero*-α-D-*manno*-heptose for the capsular polysaccharide of *Campylobacter jejuni*. *Biochemistry 58*, 3893– 3902.
17. McCallum, M., Shaw, G. S., and Creuzenet, C. (2013) Comparison of predicted epimerases and reductases of the *Campylobacter jejuni* D-*altro*- and L-*gluco*-heptose synthesis pathways. *J. Biol. Chem. 288*, 19569– 19580.
18. 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.
19. Sternberg, M. J., Tamaddoni-Nezhad, A., Lesk, V. I., Kay, E., Hitchen, P. G., Cootes, A., van Alphen, L. B., Lamoureux, M. P., Jarrell, H. C., Rawlings, C. J., Soo, E. C., Szymanski, C. M., Dell, A., Wren, B. W., and Muggleton, S.

H. (2013) Gene function hypotheses for the *Campylobacter jejuni* glycome generated by a logic-based approach. *J. Mol. Biol. 425*, 186– 197.

20. Szymanski, C. M., Michael, F. S., 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.
21. McNally, D. J., Jarrell, H. C., Khieu, N. H., Li, J., Vinogradov, E., Whitfield, D. M., Szymanski, C. M., and Brisson, J. R. (2006) The HS:19 serostrain of *Campylobacter jejuni* has a hyaluronic acid-type capsular polysaccharide with a nonstoichiometric sorbose branch and O-methyl phosphoramidate group. *FEBS J. 273*, 3975– 3989.
22. Parker, C. T., Huynh, S., and Heikema, A. P. (2016) Complete genomic sequence of *Campylobacter jejuni* subsp. jejuni HS:19 Strain RM1285 isolated from packaged chicken. *Genome Announcements 4*, e01100-16.
23. Campbell, R. E., Sala, R. F., van de Rijn, I., and Tanner, M. E. (1997) Properties and kinetic analysis of UDP-glucose dehydrogenase from group A streptococci. Irreversible inhibition by UDP-chloroacetol. *J. Biol. Chem. 272*, 3416– 3422.
24. Campbell, R. E. and Tanner, M. E. (1997) Uridine diphospho-α-D-gluco-hexodialdose: Synthesis and kinetic competence in the reaction catalyzed by UDP-glucose dehydrogenase. *Angew. Chem., Int. Ed. Engl. 36*, 1520– 1522.

25. Ge, X., Campbell, R. E., van de Rijn, I., and Tanner, M. E. (1998) Covalent adduct formation with a mutated enzyme: Evidence for a thioester intermediate in the reaction catalyzed by UDP-glucose dehydrogenase. *J. Am. Chem. Soc. 120*, 6613 – 6614.

26. Campbell, R. E. and Tanner, M. E. (1999) UDP-glucose analogues as inhibitors and mechanistic probes of UDP-glucose dehydrogenase. *J. Org. Chem. 64*, 9487– 9492.

27. Eliot, A. C. and Kirsch, J. F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem. 73*, 383– 415.

28. Keller, S., Wetterhorn, K. M., Vecellio, A., Seeger, M., Rayment, I., and Schubert, T. (2019) Structural and functional analysis of an L-serine O-phosphate decarboxylase involved in norcobamide biosynthesis. *FEBS Lett. 593*, 3040– 3053

29. Andreeßen, B. and Steinbüchel, A. (2011) Serinol: small molecule – big impact. *AMB Express 1*, 12

30. Egger, S., Chaikuad, A., Klimacek, M., Kavanagh, K. L., Oppermann, U., and Nidetzky, B. (2012) Structural and kinetic evidence that catalytic reaction of human UDP-glucose 6-dehydrogenase involves covalent thiohemiacetal and thioester enzyme intermediates. *J. Biol. Chem. 287*, 2119 – 2129.

31. Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics 1854*, 1019– 1037.

32. Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One 4*, e4345.

33. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res. 13*, 2498– 2504

34. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 67*, 271– 281.

35. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 213– 221.

36. 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.

37. Campbell, R. E., Mosimann, S. C., van De Rijn, I., Tanner, M. E., and Strynadka, N. C. (2000) The first structure of UDP-glucose dehydrogenase reveals the catalytic residues necessary for the two-fold oxidation. *Biochemistry 39*, 7012– 7023.

38. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 486– 501.

39. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA.

40. Chen, Y. Y., Ko, T. P., Lin, C. H., Chen, W. H., and Wang, A. H. (2011) Conformational change upon product binding to *Klebsiella pneumoniae* UDP-glucose dehydrogenase: a possible inhibition mechanism for the key enzyme in polymyxin resistance. *J. Struct. Biol. 175*, 300– 310.

41. Egger, S., Chaikuad, A., Kavanagh, K. L., Oppermann, U., and Nidetzky, B. (2011) Structure and mechanism of human UDP-glucose 6-dehydrogenase. *J. Biol. Chem. 286*, 23877 – 23887.

42. L'vov, V. L., Tokhtamysheva, N. V., Shashkov, A. S., Dmitriev, B. A., and Kochetkov, N. K. (1983) Bacterial antigenic polysaccharides. 12. Structure and ^{13}C NMR spectrum of the polysaccharide chain of *Shigella boydii* type 8 lipopolysaccharide. *Bioorg. Khim. 9*, 60– 73.

43. Landersjö, C., Weintraub, A., and Widmalm, G. (1996) Structure determination of the O-antigen polysaccharide from the enteroinvasive *Escherichia coli* (EIEC) O143 by component analysis and NMR spectroscopy. *Carbohydr. Res. 291*, 209– 216.

44. Zdorovenko, E. L., Varbanets, L. D., Liu, B., Valueva, O. A., Wang, Q., Shashkov, A. S., Garkavaya, E. G., Brovarskaya, O. S., Wang, L., and Knirel, Y.

A. (2014) Structure and gene cluster of the O antigen of *Escherichia coli* L-19, a candidate for a new O-serogroup. *Microbiology (London, U. K.) 160*, 2102– 2107.

45. Gilbert, M., Brisson, J. R., Karwaski, M. F., Michniewicz, J., Cunningham, A. M., Wu, Y., Young, N. M., and Wakarchuk, W. W. (2000) Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-MHz ^{1}H and ^{13}C NMR analysis. *J. Biol. Chem. 275*, 3896– 3906.
46. McNally, D. J., Jarrell, H. C., Khieu, N. H., Li, J., Vinogradov, E., Whitfield, D. M., Szymanski, C. M., and Brisson, J. R. (2006) The HS:19 serostrain of Campylobacter jejuni has a hyaluronic acid-type capsular polysaccharide with a nonstoichiometric sorbose branch and O-methyl phosphoramidate group. *FEBS J. 273*, 3975– 3989.
47. Aspinall, G. O., McDonald, A. G., Pang, H., Kurjanczyk, L. A., and Penner, J. L. (1994) Lipopolysaccharides of Campylobacter jejuni serotype O:19: structures of core oligosaccharide regions from the serostrain and two bacterial isolates from patients with the Guillain–Barre′ syndrome. *Biochemistry 33*, 241– 249.
48. Vinogradov, E. V., Holst, O., Thomas-Oates, J. E., Broady, K. W., and Brade, H. (1992) The structure of the O-antigenic polysaccharide from lipopolysaccharide of *Vibrio cholerae* strain H11 (non-O1). *Eur. J. Biochem. 210*, 491– 498.
49. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, 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., Whitehead, S., and Barrell, B. G. (2000) The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature 403*, 665– 668.

50. Vrielink, A., Rüger, W., Driessen, H. P., and Freemont, P. S. (1994) Crystal structure of the DNA modifying enzyme beta-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J. 13*, 3413– 3422.
51. Tarbouriech, N., Charnock, S. J., and Davies, G. J. (2001) Three-dimensional structures of the Mn and MgdTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J. Mol. Biol. 314*, 655– 661.
52. Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology 16*, 29R– 37R.
53. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem. 77*, 521– 555.
54. Aono, R., Ito, M., and Machida, T. (1999) Contribution of the cell wall component teichuronopeptide to pH homeostasis and alkaliphily in the alkaliphile *Bacillus lentus* C-125. *J. Bacteriol. 181*, 6600– 6606.
55. Iyer, L. M., Abhiman, S., Maxwell Burroughs, A., and Aravind, L. (2009) Amidoligases with ATP-grasp, glutamine synthetase-like and

acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol. BioSyst. 5*, 1636– 1660.

3. FUNCTIONAL CHARACTERIZATION OF TWO PLP-DEPENDENT ENZYMES INVOLVED IN CAPSULAR POLYSACCHARIDE BIOSYNTHESIS FROM *CAMPYLOBACTER JEJUNI*[*]

3.1. Introduction

Campylobacter jejuni is a Gram-negative, zoonotic organism that causes campylobacteriosis, the leading food-borne illness in the world (1). In the United States, campylobacteriosis accounts for 1.3 million cases of gastroenteritis each year (2). *C. jejuni* is a commensal organism in chickens that can easily spread to other farm animals and pets through unsanitary drinking water, meat processing facilities, unpasteurized milk, and pet feces (3). Symptoms of campylobacteriosis include nausea, cramps, fever, and bloody diarrhea (4). Additionally, *Campylobacter* infections precede ∼40% of all Guillain–Barré syndrome diagnoses (5,6). It is estimated that *Campylobacter* causes over 25 000 deaths annually in children under five years of age, primarily affecting Asia and Africa (7). There is no known FDA-approved vaccine for *Campylobacter* infection, making it a significant public health concern. The best vaccine candidates to date have been glycoconjugate vaccines, which are based on the surface-exposed sugars of *C. jejuni* (7).

[*] Reprinted with permission from "Functional Characterization of Two PLP-Dependent Enzymes Involved in Capsular Polysaccharide Biosynthesis from *Campylobacter jejuni*" by Alexander S. Riegert, Tamari Narindoshvili, Adriana Coricello, Nigel G. J. Richards, and Frank M. Raushel, *Biochemistry*, (2021), *60* (37), pp 2836-2843, Copyright 2021 American Chemical Society

Coating the exterior surface of *C. jejuni* is a polymeric layer of carbohydrates known as the capsular polysaccharide (CPS). This polysaccharide is quite diverse among different strains of *C. jejuni* and is used to serotype ~60 different variants of *C. jejuni* (8). The CPS is known to aid in evasion of the host immune system and adherence to epithelial tissue (9). The CPS of *C. jejuni* NCTC 11168 (HS:2) is composed of a repeating unit of D-*glycero*-L-*gluco*-heptose, D-glucuronic acid, D-*N*-acetyl-galactosamine, and D-ribose (10,11). The capsule is further modified by methylation, phosphoramidylation, and amidation. The glucuronic acid moiety of the CPS from *C. jejuni* serotype HS:2 is found amidated with either ethanolamine or serinol. In the HS:19 serotype, the glucuronate is amidated only with serinol (12). The biosynthetic pathways for the formation of these two modifications to the glucuronate moiety are currently unknown. The structures of the capsules found in serotypes HS:2 and HS:19 are provided in Figure 12.

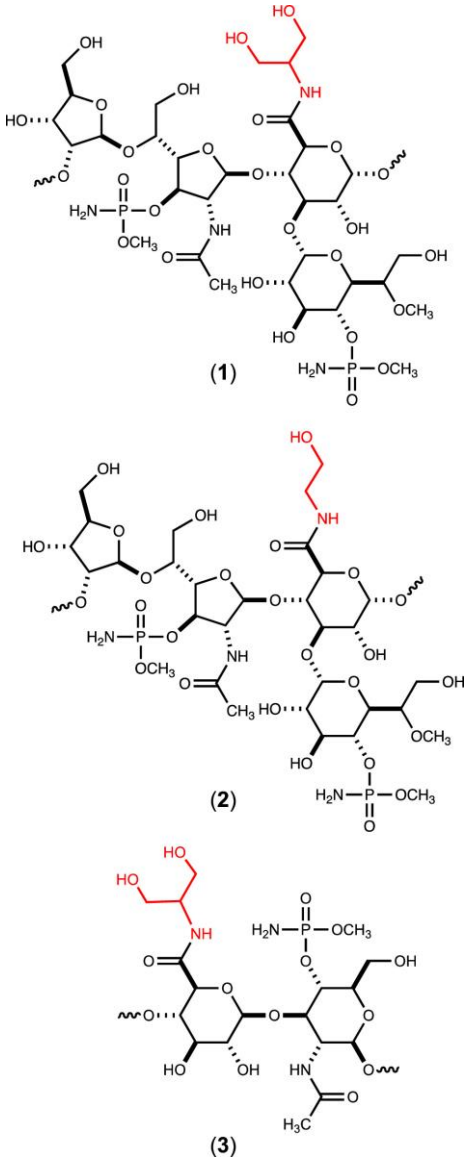


Figure 12 Structures of the capsular polysaccharides found in *C. jejuni* NCTC 11168 (HS:2) (**1** and **2**) and *C. jejuni* strain RM1285 (HS:19) (**3**). The amide moiety in each structure is highlighted in red.

The DNA cluster responsible for the formation of the CPS in *C. jejuni* NCTC 11168 is composed of 35 genes, as illustrated in Appendix Figure S10 (13). Previous investigations have functionally characterized the genes necessary for the biosynthesis of

the D-*glycero*-L-*gluco*-heptose moiety and modification (Cj1152, Cj1423, Cj1424, Cj1425, Cj1427, Cj1428, and Cj1430) (14−19). In addition, the enzymes Cj1415, Cj1416, Cj1417, and Cj1418 have been shown to be required for the biosynthesis of the phosphoramidate modification, and gene knockout experiments have identified the enzymes (Cj1422 and Cj1421) required for the transfer of the phosphoramidate modifications to specific sugar receptors (8,20−25). More recently, we have shown that Cj1441 is required for the conversion of uridine diphosphate (UDP)-glucose to UDP-glucuronate and that this enzyme is unable to catalyze the formation of an amide bond via the attack of an amine substrate with the putative thioester intermediate formed during the oxidation of UDP-glucose (26). The enzymes required for the biosynthesis of the two amines (serinol and ethanolamine) found in the HS:2 capsule are currently unknown.

The three most likely enzymes required for the biosynthesis of serinol and ethanolamine in the HS:2 capsule of *C. jejuni* are Cj1435, Cj1436, and Cj1437. Cj1435 is predicted to be a HAD phosphatase, while Cj1436 and Cj1437 are anticipated to be pyridoxal phosphate (PLP)-dependent enzymes. Close homologs to Cj1435 and Cj1437 are found in the gene cluster for the HS:19 serotype, but none to Cj1436. A sequence similarity network (SSN) for the closest 1000 protein sequences to Cj1436 is presented in Figure 13a, at a threshold of 39% sequence identity (27,28). The nearest functionally characterized enzyme is L-serine phosphate decarboxylase (SMUL_1544), with an overall sequence identity of 46%. This enzyme has been shown to catalyze the formation of ethanolamine phosphate during the biosynthesis of

norcobamides (29,30). The SSN with Cj1437 for the nearest 1000 sequences is presented in Figure 13b, at a sequence identity of 41%. The nearest functionally characterized enzyme to Cj1437 is histidinol phosphate aminotransferase, which catalyzes the transamination of imidazole acetol phosphate, in the presence of L-glutamate, during the biosynthesis of L-histidine. Here, we have purified Cj1436 and Cj1437 and have functionally characterized these two enzymes. Cj1436 is found to catalyze the decarboxylation of L-serine phosphate to generate ethanolamine phosphate and CO_2, while Cj1437 catalyzes the transamination of dihydroxyacetone phosphate to (*S*)-serinol phosphate in the presence of L-glutamate.

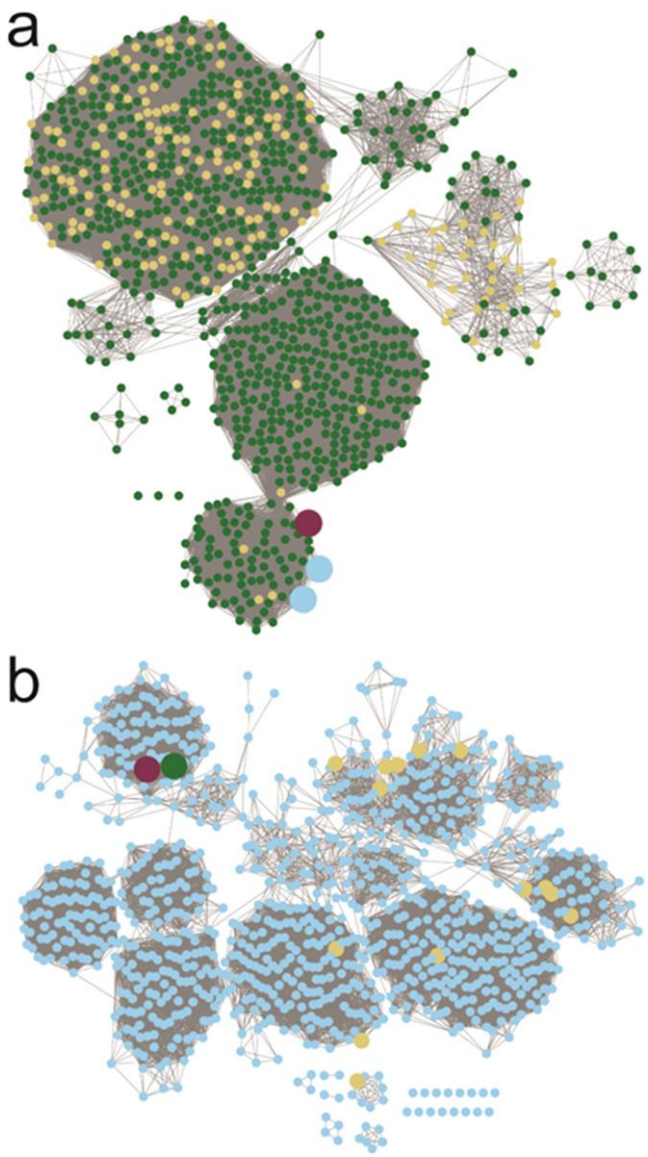


Figure 13 Sequence similarity networks for Cj1436 and Cj1437 using the 1000 nearest protein sequences. (a) The SSN for Cj1436 clustered at a sequence identity of 39%. Maroon node is Cj1436. Blue nodes are enzymes sharing 99% sequence identity to the functionally characterized L-serine phosphate decarboxylase (SMUL_1544 from *Sulfurospirillum multivorans*). Yellow nodes are annotated as L-threonine phosphate decarboxylases. Green nodes are described as aminotransferases. (b) SSN for Cj1437 clustered at a sequence identity of 41%. The maroon node is Cj1437, while the green node is HS19.10. Yellow nodes are proteins with a Swiss-Prot designation as histidinol phosphate aminotransferases. Blue nodes do not have a Swiss-Prot annotation but are generically described as aminotransferases.

3.2. Materials and Methods

3.2.1. Cloning, Expression, and Purification of Cj1436 and Cj1437

The genes for Cj1436 (UniProt id: Q0P8H8) and Cj1437 (UniProt id: Q0P8H7) from serotype HS:2 were cloned from the genomic DNA of *C. jejuni* NCTC 11168 obtained from ATCC. The genomic DNA served as the starting template for PCR using Phusion DNA polymerase (New England Biolabs). Primers were designed that incorporated *Nde*I and *Xho*I restriction sites, and the resulting fragments were digested at 37 °C and ligated into a pET31b expression vector, which carried a C-terminal hexahistidine tag. The gene encoding the homolog to Cj1437 (HS19.10) from the HS:19 serotype in *C. jejuni* strain RM1285 (Uniprot id: Q5M6M3) was codon-optimized for expression in *E. coli* and ordered from Twist Biosciences in a pET28a expression vector with an N-terminal hexahistidine tag.

The pET31b vector was used to transform BL21 *E. coli* cells (Novagen) via electroporation. Cells harboring the pET31b-Cj1436, pET31b-Cj1437, and pET28a-HS19.10 plasmids were cultured in lysogeny broth with 100 mg/L ampicillin for the pET31b constructs and 50 mg/L kanamycin for the pET28a construct. The cells were grown at 37 °C with shaking and induced with 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the optical density reached 0.8 at 600 nm. The cells were allowed to express protein at 21 °C for 18 h after induction and then harvested by centrifugation at 15 000 rcf at 4 °C. The cell pellet was resuspended in the loading buffer (50 mM *N*-(2-hydroxyethyl)piperazine-*N*′-ethanesulfonic acid (HEPES)/K^+, 300

mM KCl, 20 mM imidazole, pH 8.0) and lysed with sonication. The sonicated cells were centrifuged at 25 000 rcf at 4 °C before the lysate was passed through a 0.45 μm filter (Whatman). The sample was loaded onto a prepacked 5 mL HisTrap (GE Healthcare) nickel affinity column. The protein was eluted from the column using 50 mM HEPES/K^+, pH 8.0, 300 mM KCl, and 250 mM imidazole over a gradient of 25 column volumes. The protein was pooled and dialyzed against 10 mM HEPES/K^+, pH 8.0 and 200 mM KCl. Cj1436 was concentrated to 10 mg/mL and flash-frozen using liquid nitrogen before being stored at −80 °C. Approximately 5 mg of protein was obtained per liter of cell culture. Cj1437 (~3 mg of protein per liter of cell culture) was concentrated to 3.5 mg/mL and flash-frozen before being stored at −80 °C. HS19.10 (~10 mg of protein per liter of cell culture) was concentrated to 6 mg/mL and flash-frozen before being stored at −80 °C.

3.2.2. Bioinformatic Analysis of Cj1436 and Cj1437

The amino acid sequences for Cj1436 and Cj1437 were obtained from the Uniprot database. Individually, the protein sequences for the two genes were input as the search queries for the BLAST section of the EFI-EST webtool to generate the sequence similarity networks (SSNs) (27,28). The parameters were set such that the BLAST search would retrieve the 1000 closest sequences to the search query. Sequence identities for Cj1436 and Cj1437 were set at 39 and 41%, respectively. The networks were created using the yFiles organic layout, and all images were created using Cytoscape version 3.7.2.

3.2.3. Determination of Kinetic Constants for Cj1437 and HS19.10

The kinetic constants for Cj1437 and HS19.10 were determined using an L-glutamate dehydrogenase coupled assay to monitor the formation of α-ketoglutarate. The reaction mixtures contained 150 nM enzyme, 250 μM PLP, 250 μM reduced nicotinamide adenine dinucleotide phosphate (NADPH), 15 mM ammonium chloride, 2.0 mM L-glutamate, 5.0 units of L-glutamate dehydrogenase, varying levels of dihydroxyacetone phosphate (DHAP) (0.005–2.0 mM), and 50 mM HEPES/K^+ buffer at pH 8.0. The reaction was monitored spectrophotometrically at 340 nm and 25 °C with a SpectraMax340 UV–visible spectrophotometer. The kinetic parameters were determined by fitting the initial rates to eq 1 using GraFit 5, where v is the initial velocity of the reaction, E_t is the enzyme concentration, k_{cat} is the turnover number, A is the substrate concentration, and K_m is the Michaelis constant.

$$v / E_t = k_{cat} [A] / (A + K_m) \quad (1)$$

3.2.4. Determination of Reaction Products for Cj1436

The reaction catalyzed by Cj1436 was determined by following the decarboxylation of L-serine phosphate via ^1H and ^{31}P NMR spectroscopy. The reaction was conducted in 50 mM phosphate buffer, pD 8.0, in D_2O and initiated by the addition of 1.0 μM Cj1436 with a fixed level of PLP (250 μM) and a fixed concentration of L-serine phosphate (5.0 mM). The catalytic activity of Cj1436 was also tested with L-serine, L-threonine phosphate, L-aspartate, and L-glutamate at a fixed concentration of

5.0 mM. The reaction was also monitored using ^{31}P NMR spectroscopy by monitoring the phosphorus resonance for the product (ethanolamine phosphate) relative to the ^{31}P resonance for substrate (serine phosphate) as a function of time. The reaction was conducted in 50 mM HEPES/K^+ at pH 8.0 containing 250 μM PLP, 5.0 mM L-serine and initiated by the addition of 1.0 μM enzyme.

3.2.5. Membrane Inlet Mass Spectrometry (MIMS) for the Determination of the Kinetic Constants for Cj1436

Steady-state kinetic constants for Cj1436 were determined by following the decarboxylation of L-serine phosphate using MIMS (31−33). In these experiments, assay mixtures contained 0.1, 0.5, 1.0, 5.0, and 10 mM L-serine phosphate dissolved in 100 mM HEPES/K^+ buffer, pH 8.0. Reactions were initiated at 25 °C by the addition of an enzyme/PLP mixture in 100 mM HEPES/K^+ buffer, pH 8.0, to give final concentrations of 1.0 μM Cj1436 and 250 μM PLP (total volume 2.0 mL). The catalytic activity of Cj1436 was also tested with L-serine at a fixed concentration of 5.0 mM.

3.2.6. Reaction Product Confirmation Using Mass Spectrometry

Cj1436 (1.0 μM) was incubated with 5.0 mM L-serine phosphate, 250 μM PLP, and 50 mM ammonium bicarbonate at pH 8.0. The reaction was monitored using ^{31}P NMR spectroscopy until the reaction was complete. The solution was filtered through a GE Healthcare Vivaspin 500 10 kDa molecular weight cutoff (MWCO) filter, and 100 μL of the flow-through was submitted for mass spectrometry analysis using a Thermo

Scientific Q Exactive Focus mass spectrometer. Cj1437 (1.0 μM) was incubated with 5.0 mM DHAP, 15 mM L-glutamate, and 250 μM PLP in 50 mM ammonium bicarbonate buffer at pH 8.0 until the reaction reached equilibrium as determined by ^{31}P NMR spectroscopy. The solution was filtered through a GE Healthcare Vivaspin 500 10 kDa MWCO filter, and 100 μL of the flow-through was analyzed using a Thermo Scientific Q Exactive Focus mass spectrometer.

3.2.7. Chemical Synthesis of (*R*/*S*)-Serinol Phosphate and (*S*)-Serinol Phosphate

The synthesis of racemic and (*S*)-serinol phosphate is described in the Appendix Supporting Information.

3.2.8. Stereochemical Analysis of Serinol Phosphate Produced by Cj1437 and HS19.10

The stereochemistry of the chiral center at C2 of the serinol phosphate produced by Cj1437 and HS19.10 was determined by monitoring the exchange of the hydrogen at C2 of (*R*/*S*)-serinol phosphate and (*S*)-serinol phosphate with solvent deuterium. Cj1437 or HS19.10, at a concentration of 3.0 μM, was allowed to incubate with 3.0 mM (*R*/*S*)-serinol in 50 mM phosphate buffer (D_2O) at pD 8.0. The exchange of the hydrogen at C2 of serinol phosphate with deuterium was monitored using ^{1}H NMR spectroscopy (Avance III 400 MHz). A similar experiment was conducted with the (*S*)-serinol phosphate enantiomer. The exchange reactions were also conducted in the presence of 250 μM PLP and 250 μM α-ketoglutarate using the HS19.10 enzyme.

3.3. Results

3.3.1. Determination of the Reaction Catalyzed by Cj1436

The closest functionally characterized enzyme to Cj1436 is SMUL_1544 from *S. multivorans,* and this enzyme has been shown to catalyze the decarboxylation of L-serine phosphate to ethanolamine phosphate and CO_2 (29,30). ^{1}H NMR spectroscopy was used to monitor the reaction catalyzed by Cj1436 in D_2O with L-serine phosphate, and the results are presented in Figure 14, showing that the product is ethanolamine phosphate. The two hydrogens from C2 resonate at 3.96 ppm and appear as a doublet of doublets due to coupling with the adjacent phosphate and the single proton from C1 with coupling constants of 1.89 and 12.14 Hz, respectively (Figure 14a). The single proton from C1 appears as a broadened triplet at 3.19 ppm. The ^{1}H NMR spectrum of the substrate prior to the addition of the enzyme is provided in Figure 14b.

In the ^{31}P NMR spectrum, the substrate exhibits a single resonance at 3.89 ppm (Figure 15b), whereas the product resonates at 3.82 ppm (Figure 15a). The formation of ethanolamine phosphate was confirmed using electrospray ionization (ESI) negative mode mass spectrometry. The substrate of the reaction, L-serine phosphate, is present at an *m*/*z* of 184.0 for the M–H ion (Figure 16a). The product of the reaction, ethanolamine

phosphate, is found at an *m*/*z* of 140.01 when the reaction is conducted in H_2O (Figure 16b) and at 141.0 when the reaction is conducted in D_2O (Figure 16c).

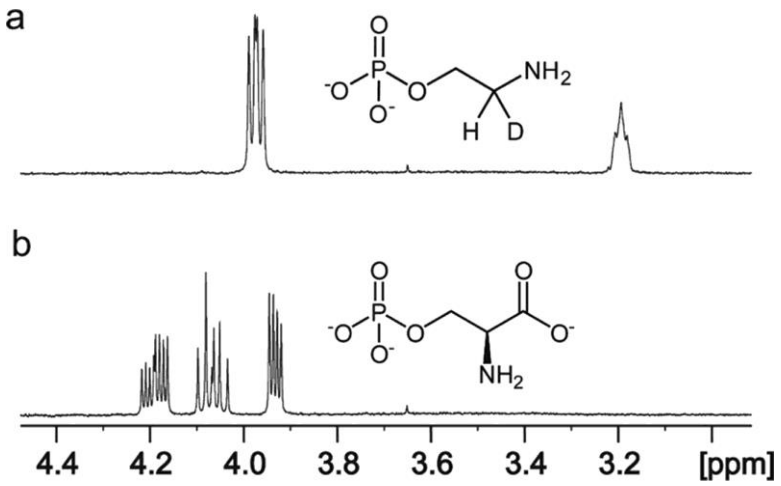


Figure 14 ^{1}H NMR spectrum of the reaction catalyzed by Cj1436 in D_2O. (a) The product of the reaction catalyzed by Cj1436 in the presence of L-serine phosphate is ethanolamine phosphate. (b) The substrate of the reaction, L-serine phosphate, prior to the addition of enzyme. Additional details are provided in the text.

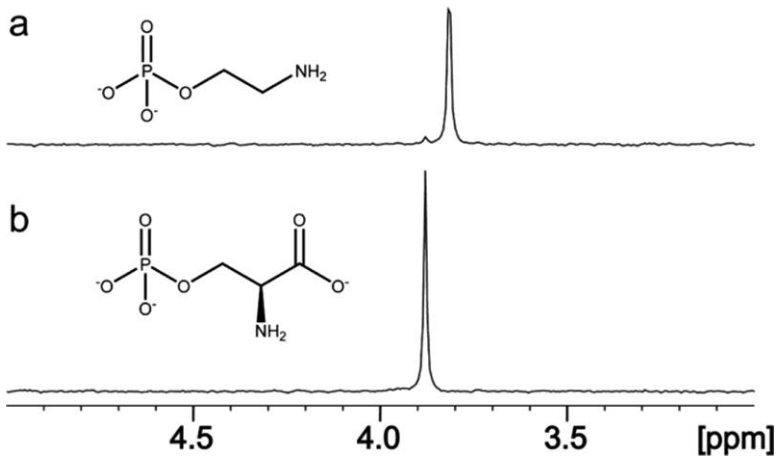


Figure 15 ^{31}P NMR spectrum of the reaction catalyzed by Cj1436. (a) Ethanolamine phosphate resonates at 3.82 ppm. A small amount of the remaining substrate appears at 3.89 ppm. (b) Control reaction in the absence of enzyme showing the substrate, L-serine phosphate, at 3.89 ppm. Additional details are provided in the text.

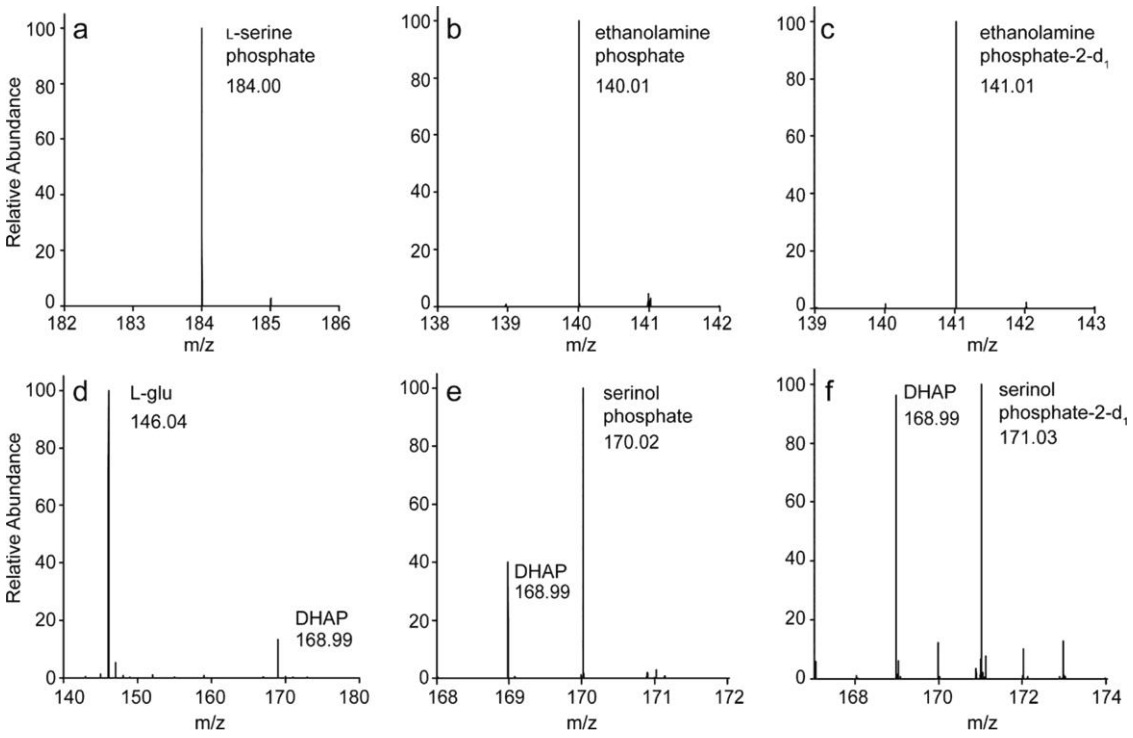


Figure 16 Mass spectrometry data for Cj1436 and Cj1437. (a) L-Serine phosphate, the substrate for Cj1436, appears in a no-enzyme control at an *m*/*z* of 184.00 for the M–H ion. (b) Product of the reaction catalyzed by Cj1436. Ethanolamine phosphate appears at an *m*/*z* of 140.01 for the M–H ion. (c) Product of the reaction catalyzed by Cj1436 in the presence of D_2O. The deuterated product, (2-d_1)-ethanolamine phosphate, appears at an *m*/*z* of 141.01 for the M–H ion. (d) L-glutamate and DHAP, substrates for the reaction catalyzed by Cj1437, appear in a no-enzyme control at *m*/*z* of 146.04 and 168.99, respectively, for the M–H ions. (e) Product of the reaction catalyzed by Cj1437. The product serinol phosphate appears at an *m*/*z* equal to 170.02, while the leftover DHAP substrate appears at 168.99 for the M–H ions. (f) The product of the reaction catalyzed by Cj1437 in the presence of D_2O. (2-d_1)-Serinol phosphate appears at an m/z of 171.03, while the leftover DHAP substrate appears at 168.99 for the M–H ions.

3.3.2. Determination of the Reaction Catalyzed by Cj1437

The sequence similarity network for Cj1437 indicates that this enzyme will most likely catalyze a transamination reaction since the closest functionally characterized enzyme is histidinol phosphate aminotransferase. Since the pathway for the biosynthesis of the CPS in serotype HS:2 must eventually make serinol, the most probable reaction

catalyzed by Cj1437 is the transamination of DHAP, with an amino donor such as L-glutamate to make serinol phosphate. When tested with dihydroxyacetone (DHA) as a potential substrate, the rate was less than 3% of the rate with DHAP. The reaction catalyzed by Cj1437 was initially characterized using ^{1}H NMR spectroscopy using 3.0 mM DHAP and 15 mM L-glutamate. A portion of the ^{1}H NMR spectrum of the mixture of DHAP and L-glutamate in D_2O is presented in Figure 17a. After the addition of Cj1437, deuterated serinol phosphate is formed as indicated by the pair of doublet of doublets for the two protons at C1 centered at 3.95 ppm and the doublet of doublets for the two protons at C3 centered at 3.80 ppm (Figure 17b). The NMR spectrum of the chemically synthesized racemic serinol phosphate is shown for comparison in Figure 17c, where the hydrogen at C2 resonates at 3.51 ppm, thus confirming the formation of serinol phosphate.

The product of the reaction catalyzed by Cj1437 has a chiral center at C2. To determine the stereochemistry at C2, the enzyme was incubated with racemic serinol phosphate (3.0 mM) and α-ketoglutarate (250 μM), and the exchange of the hydrogen at C2 with solvent deuterium was monitored by ^{1}H NMR spectroscopy. Prior to the addition of the enzyme (3.0 μM), the integrated ratio of the resonances for the hydrogens at C1 and C3 relative to the hydrogen at C2 was 4.0:1.04. After the addition of the enzyme, the ratio gradually changed to 4.0:0.55 (Figure 17d), where the mixture of the serinol phosphate with hydrogen and deuterium at C2 is now clearly visible. When the same reaction was conducted using chemically synthesized (*S*)-serinol phosphate (Figure 17e), the hydrogen at C2 completed exchange with solvent deuterium after 2 h

(Figure 17f), thus demonstrating that Cj1437 catalyzes the synthesis of (*S*)-serinol phosphate from DHAP and L-glutamate.

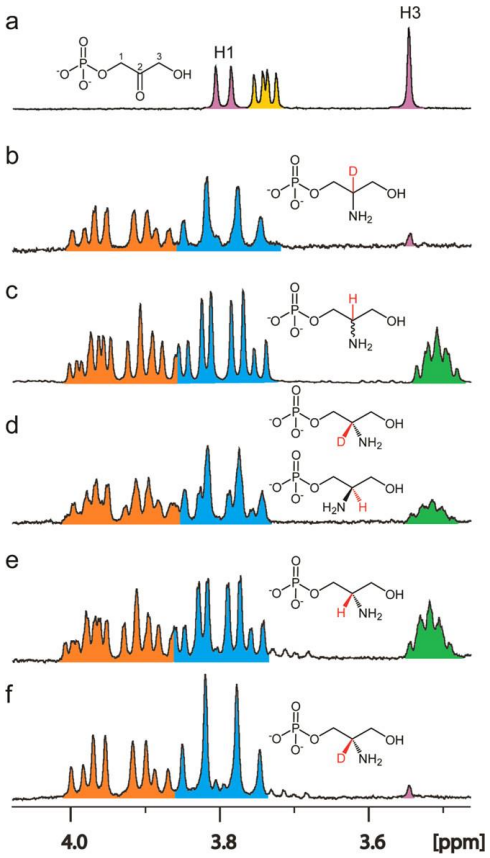


Figure 17 ^{1}H NMR spectra of reactions catalyzed by Cj1437. (a) Control spectrum of a mixture of DHAP (pink) and L-glutamate (yellow). The resonance for the H1 hydrogen of DHAP is centered at 3.79 ppm, while the H3 hydrogen is centered at 3.55 ppm. The Cα of L-glutamate appears as a doublet of doublets centered at 3.74. (b) Product of the reaction catalyzed by Cj1437. The H1 hydrogens (orange) are centered at 3.94 ppm. The C3 hydrogens (blue) are centered at 3.80 ppm. A small amount of DHAP is left over, as shown by the pink singlet at 3.55 ppm. (c) Chemically synthesized (*R*/*S*)-serinol phosphate. The H1 protons are centered at 3.93 ppm, and the H2 hydrogen (green) is centered at 3.51 ppm. (d) Incubation of Cj1437 with (*R*/*S*)-serinol phosphate with 50 μM αKG in D_2O for 2 h. (e) Chemically synthesized (*S*)-serinol phosphate. (f) Incubation of Cj1437 with (*S*)-serinol phosphate and 250 μM α-KG for 2 h. Additional details are provided in the text.

The reaction products formed after the incubation of Cj1437 with DHAP and L-glutamate were confirmed to be serinol phosphate and α-ketoglutarate by ESI mass spectrometry (negative mode). In the control experiment, in the absence of enzyme, the substrates DHAP and L-glutamate are identified at *m*/*z* of 146.05 and 168.99, respectively, for the M–H ions (Figure 16d). The product, serinol phosphate, is found at 170.02 (Figure 16e). Furthermore, when the reaction is conducted in D_2O, the serinol phosphate product incorporates a single deuterium, as demonstrated by the appearance of the signal at an *m*/z of 171.03 (Figure 16f).

3.3.3. Kinetic Analysis of the Reaction Catalyzed by Cj1436

Steady-state kinetic parameters for Cj1436 were obtained by monitoring the rate of decarboxylation at various concentrations of L-serine phosphate using membrane inlet mass spectrometry. These experiments provided values for k_{cat}, k_{cat}/K_m, and K_m of 0.21 ± 0.02 s^{-1}, 350 ± 120 M^{-1} s^{-1}, and 0.6 ± 0.2 mM, respectively. The value of k_{cat}/K_m for Cj1436 is 3–4 fold lower than that determined for Cj1437 (see below). No activity (<0.01 s^{-1}) was observed when L-serine replaced L-serine phosphate in the assay.

3.3.4. Kinetic Analysis of the Reaction Catalyzed by Cj1437

The reaction catalyzed by Cj1437 and HS19.10 was followed using a glutamate dehydrogenase coupled assay to measure the rate of formation of α-ketoglutarate. Using DHAP and L-glutamate as substrates, the values of k_{cat}, k_{cat}/K_m, and K_m with Cj1437 were determined to be 0.50 ± 0.02 s^{-1}, 1.2 (± 0.2) x 10^4 M^{-1} s^{-1}, and 42 ± 6 μM,

respectively. Using the homolog to Cj1437 from the HS:19 serotype (HS19.10), the values of k_{cat}, k_{cat}/K_m, and K_m were determined to be 1.10 ± 0.01 s^{-1}, of 2.7 (± 0.1) × 10^4 M^{-1} s^{-1}, and 41 ± 2 μM, respectively. The turnover using L-aspartate (15 mM) was less than 50% of the rate exhibited by L-glutamate (15 mM) at a fixed concentration of DHAP (3.0 mM) as monitored by ^{1}H NMR spectroscopy. Essentially, no activity could be detected using L-alanine.

3.4. Discussion

The HS:2 serotype of *C. jejuni* is surrounded by a capsular polysaccharide consisting of a repeating unit of D-*glycero*-L-*gluco*-heptose, D-glucuronic acid, D-*N*-acetyl-galactosamine, and D-ribose (10,11). The glucuronic acid is amidated with either serinol or ethanolamine. This same modification can be found within the CPS of the *C. jejuni* serotype HS:19, except that the amide bond is formed only with serinol. The exact role of the amide bond-containing molecule within the CPS is unknown. Other modifications to the CPS are critical for the evasion of the immune system and adherence to epithelial tissue (9,34). Additionally, the phase variable phosphoramidate modification is a target for recognition by bacteriophage F336 (35). Phase variation may play a role in the biosynthesis of glucuronamide. The gene, which encodes for Cj1437 appears to have a poly-G tract of 9 nucleotides [877–885]. Poly-G tracts are common to *C. jejuni* and are known to introduce phase variation, which creates phenotypic heterogeneity in populations of *C. jejuni*. This allows *C. jejuni* to avoid the immune system as well as phage attack (37).

To form the amide bond found in the CPS of *C. jejuni* HS:2, the bacterium must synthesize two primary amines, ethanolamine and serinol. Here, we have shown that the first step in the biosynthesis of the ethanolamine product is the decarboxylation of L-serine phosphate to ethanolamine phosphate catalyzed by Cj1436. This result is consistent with the prior identification of an L-serine phosphate decarboxylase (SMUL_1544), which was shown previously to catalyze the decarboxylation of L-serine phosphate (29). SMUL_1544 is the only other functionally characterized L-serine phosphate decarboxylase reported to date and is used in the biosynthesis of norpseudovitamin B_{12} in *S. multivorans* (29,30). L-Threonine phosphate is also a weaker substrate, but no activity could be detected with L-serine, L-threonine, L-aspartate, or L-glutamate. The phosphate group of the substrate is thus necessary for substrate recognition and is required for catalytic activity. The homologous three-dimensional structures of SMUL_1544 and CobD of *S. enterica* have identified the residues used to bind the phosphate group of the substrate (29,38). In SMUL_1544, these residues are Arg-356, Arg-368, and Ser-25. A sequence alignment of Cj1436 with SMUL_1544 reveals the conservation of those residues required to bind the phosphate group of the substrate (Appendix Figure S11).

Cj1437 participates in the first step in serinol formation, where it catalyzes the transamination of DHAP to serinol phosphate using L-glutamate as the amine donor. The stereochemistry of the reaction was shown to produce the (*S*)-enantiomer of serinol phosphate using solvent deuterium exchange with chemically synthesized (*S*)-serinol phosphate. This stereochemical outcome is consistent with the known product formation

of L-histidinol phosphate in the reaction catalyzed by histidinol phosphate transaminase (39). Serinol phosphate has been shown to be an important precursor for serinol production (40). Serinol is an intermediate in the formation of other chemicals, including the anti-cancer drugs *N*-palmitoyl-2-amino-1,3-propanediol, and the first orally administered treatment for multiple sclerosis, Gilenya (Novartis) (41−43).

The precise sequence of the enzymatic transformations required for the biosynthesis of the glucuronamide moiety in the capsular polysaccharide of *C. jejuni* HS:2 has not yet been elucidated. A previous bioinformatic investigation identified a probable candidate enzyme, Cj1438, which contains an N-terminal glycosyltransferase domain as well as a C-terminal ATP-grasp domain that could be used for amide bond formation. However, at this point, it is unclear whether the ATP-grasp domain would utilize serinol phosphate or serinol as the amine substrate or whether the acceptor carboxylate would be contributed by UDP-glucuronate or the growing polysaccharide chain. In either case, the putative phosphatase (Cj1435) will ultimately have to be used to remove the phosphoryl group from either serinol phosphate or the phosphorylated glucuronamide moiety. Efforts to unravel these possible scenarios are in progress.

Glucuronamides are found within the lipopolysaccharide O-antigen and CPS of many bacteria, including *Acinetobacter baumannii* G7, *E. coli* O143, *E. coli* L-19, *Vibrio cholerae* H11 (non-O1), *Shigella boydii* type 8, *Pseudomonas aeruginosa*, *Proteus mirabilis* O27, *Vibrio vulnificus* strain 6353, *Bacillus halodurans*, *Providencia stuartii* O25, and *Proteus penneri* strain 14 (44−36). In *A.*

baumannii G7 *E. coli* O143, *E. coli* L-19, *S. boydii* type 8, and *B. halodurans*, there is an ATP-grasp enzyme that is located in the O-antigen or CPS biosynthetic gene cluster. In addition to the organisms listed, *C. jejuni* HS:19 contains a glucuronamide as well as an ATP-grasp-containing enzyme. An analysis of other *C. jejuni* genomes revealed that *C. jejuni* HS:22 also contains the genes necessary for the biosynthesis of a CPS modified with a glucuronamide. A comparison of the capsular polysaccharide biosynthetic gene clusters for HS:2, HS:19, and HS:22 is presented in Appendix Figure S10.

3.5. Conclusions

Here, we report the functional characterization of two enzymes necessary for the formation of the *C. jejuni* HS:2 capsular polysaccharide, Cj1436 and Cj1437. Cj1436 is a PLP-dependent decarboxylase, which catalyzes the decarboxylation of L-serine phosphate to ethanolamine phosphate. The products were confirmed by mass spectrometry and ^{1}H NMR spectroscopy. The rate and catalytic constants were measured using NMR spectroscopy and membrane inlet mass spectrometry. Cj1437 is a PLP-dependent transaminase, which catalyzes the transamination of DHAP using L-glutamate to form (*S*)-serinol phosphate. The product of the reaction was characterized using mass spectrometry and ^{1}H NMR spectroscopy. The rate and catalytic constants were determined spectrophotometrically, using a glutamate dehydrogenase coupled assay.

3.6. References

1. Epps, S. V.; Harvey, R. B.; Hume, M. E.; Phillips, T. D.; Anderson, R. C.; Nisbet, D. J. Foodborne *Campylobacter*: infections, metabolism, pathogenesis and reservoirs. *Int. J. Environ. Res. Public Health* 2013, *10*, 6292– 6304.
2. Scallan, E.; Hoekstra, R. M.; Angulo, F. J.; Tauxe, R. V.; Widdowson, M. A.; Roy, S. L.; Jones, J. L.; Griffin, P. M. Foodborne illness acquired in the United States--major pathogens. *Emerging Infect. Dis.* 2011, *17*, 7 – 15.
3. Burnham, P. M.; Hendrixson, D. R. *Campylobacter jejuni*: collective components promoting a successful enteric lifestyle. *Nat. Rev. Microbiol.* 2018, *16*, 551– 565.
4. Acheson, D.; Allos, B. M. *Campylobacter jejuni* Infections: update on emerging issues and trends. *Clin. Infect. Dis.* 2001, *32*, 1201 – 1206.
5. Nachamkin, I.; Allos, B. M.; Ho, T. *Campylobacter* species and Guillain-Barré syndrome. *Clin. Microbiol. Rev.* 1998, *11*, 555– 567.
6. Yuki, N. Molecular mimicry between gangliosides and lipopolysaccharides of *Campylobacter jejuni* isolated from patients with Guillain-Barré Syndrome and Poly, F.; Noll, A. J.; Riddle, M. S.; Porter, C. K. Update on *Campylobacter* vaccine development. *Hum. Vaccines Immunother.* 2019, *15*, 1389– 1400.
8. 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.; Szymanski, C. M. Commonality and biosynthesis of the O-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *J. Biol. Chem.* 2007, *282*, 28566– 28576.

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.; Szymanski, C. M. Biological roles of the O-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One* 2014, *9*, e87051 .
10. St.Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* 2002, *269*, 5119– 5136.
11. Young, K. T.; Davis, L. M.; Dirita, V. J. *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* 2007, *5*, 665– 679.
12. McNally, D. J.; Jarrell, H. C.; Khieu, N. H.; Li, J.; Vinogradov, E.; Whitfield, D. M.; Szymanski, C. M.; Brisson, J. R. The HS:19 serostrain of *Campylobacter jejuni* has a hyaluronic acid-type capsular polysaccharide with a nonstoichiometric sorbose branch and O-methyl phosphoramidate group. *FEBS J.* 2006, *273*, 3975– 3989.
13. Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R. M.; Feltwell, T.; Holroyd, 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.; Whitehead, S.; Barrell, B. G. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000, *403*, 665– 668.

14. Huddleston, J. P.; Anderson, T. K.; Spencer, K. D.; Thoden, J. B.; Raushel, F. M.; Holden, H. M. Structural analysis of Cj1427, an essential NAD-dependent dehydrogenase for the biosynthesis of the heptose residues in the capsular polysaccharides of *Campylobacter jejuni*. *Biochemistry* 2020, *59*, 1314– 1327.
15. Huddleston, J. P.; Raushel, F. M. Functional characterization of Cj1427, a unique ping-pong dehydrogenase responsible for the oxidation of GDP-D-glycero-α-D-manno-heptose in *Campylobacter jejuni*. *Biochemistry* 2020, *59*, 1328– 1337.
16. Huddleston, J. P.; Raushel, F. M. Biosynthesis of GDP-D-glycero-α-D-manno-heptose for the capsular polysaccharide of *Campylobacter jejuni*. *Biochemistry* 2019, *58*, 3893– 3902.
17. McCallum, M.; Shaw, G. S.; Creuzenet, C. Comparison of predicted epimerases and reductases of the *Campylobacter jejuni* D-altro- and L-gluco-heptose synthesis pathways. *J. Biol. Chem.* 2013, *288*, 19569– 19580.
18. Barnawi, H.; Woodward, L.; Fava, N.; Roubakha, M.; Shaw, S. D.; Kubinec, C.; Naismith, J. H.; Creuzenet, C. Structure-function studies of the C3/C5 epimerases and C4 reductases of the *Campylobacter jejuni* capsular heptose modification pathways. *J. Biol. Chem.* 2021, *296*, 100352.
19. Huddleston, J. P.; Anderson, T. K.; Girardi, N. M.; Thoden, J. B.; Taylor, Z.; Holden, H. M.; Raushel, F. M. Biosynthesis of D-glycero-L-gluco-Heptose in the Capsular Polysaccharides of *Campylobacter jejuni*. *Biochemistry* 2021, *60*, 1552– 1563.

20. Taylor, Z. W.; Brown, H. A.; Holden, H. M.; Raushel, F. M. Biosynthesis of nucleoside diphosphoramidates in *Campylobacter jejuni*. *Biochemistry* 2017, *56*, 6079– 6082.

21. Taylor, Z. W.; Chamberlain, A. R.; Raushel, F. M. Substrate specificity and chemical mechanism for the reaction catalyzed by glutamine kinase. *Biochemistry* 2018, *57*, 5447– 5455.

22. Taylor, Z. W.; Raushel, F. M. Manganese-induced substrate promiscuity in the reaction catalyzed by phosphoglutamine cytidylyltransferase from *Campylobacter jejuni*. *Biochemistry* 2019, *58*, 2144– 2151.

23. Taylor, Z. W.; Raushel, F. M. Cytidine diphosphoramidate kinase: an enzyme required for the biosynthesis of the O-methyl phosphoramidate modification in the capsular polysaccharides of *Campylobacter jejuni*. *Biochemistry* 2018, *57*, 2238– 2244.

24. Taylor, Z. W.; Brown, H. A.; Narindoshvili, T.; Wenzel, C. Q.; Szymanski, C. M.; Holden, H. M.; Raushel, F. M. 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.* 2017, *139*, 9463– 9466.

25. Sternberg, M. J.; Tamaddoni-Nezhad, A.; Lesk, V. I.; Kay, E.; Hitchen, P. G.; Cootes, A.; van Alphen, L. B.; Lamoureux, M. P.; Jarrell, H. C.; Rawlings, C. J.; Soo, E. C.; Szymanski, C. M.; Dell, A.; Wren, B. W.; Muggleton, S. H. Gene

function hypotheses for the *Campylobacter jejuni* glycome generated by a logic-based approach. *J. Mol. Biol.* 2013, *425*, 186– 197.

26. Riegert, A. S.; Raushel, F. M. Functional and Structural Characterization of the UDP-Glucose Dehydrogenase Involved in Capsular Polysaccharide Biosynthesis from *Campylobacter jejuni*. *Biochemistry* 2021, *60*, 725– 734.
27. Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics* 2015, *1854*, 1019– 1037.
28. Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One* 2009, *4*, e4345.
29. Keller, S.; Wetterhorn, K. M.; Vecellio, A.; Seeger, M.; Rayment, I.; Schubert, T. Structural and functional analysis of an L-serine O-phosphate decarboxylase involved in norcobamide biosynthesis. *FEBS Lett.* 2019, *593*, 3040– 3053.
30. Keller, S.; Treder, A.; von Reuss, S. H.; Escalante-Semerena, J. C.; Schubert, T. The SMUL_1544 gene product governs norcobamide biosynthesis in the tetrachloroethene-respiring bacterium *Sulfurospirillum multivorans*. *J. Bacteriol.* 2016, *198*, 2236– 2243.
31. Sheng, X.; Zhu, W.; Huddleston, J.; Xiang, D. F.; Raushel, F. M.; Richards, N. G. J.; Himo, F. A combined experimental-theoretical study of the LigW-catalyzed

decarboxylation of 5-carboxyvanillate in the metabolic pathway for lignin degradation. *ACS Catal.* 2017, *7*, 4968– 4974.

32. Zhu, W.; Easthon, L. M.; Reinhard, L. A.; Tu, C.-K.; Cohen, S. E.; Silverman, D. N.; Allen, K. N.; Richards, N. G. J. Substrate binding mode and molecular basis of a specificity switch in oxalate decarboxylase. *Biochemistry* 2016, *55*, 2163– 2173.
33. Moral, M. E. G.; Tu, C.-K.; Richards, N. G. J.; Silverman, D. N. Membrane inlet for mass spectrometric measurement of catalysis by enzymatic decarboxylases. *Anal. Biochem.* 2011, *418*, 73– 77.
34. Guerry, P.; Poly, F.; Riddle, M.; Maue, A. C.; Chen, Y. H.; Monteiro, M. A. *Campylobacter* polysaccharide capsules: virulence and vaccines. *Front. Cell. Infect. Microbiol.* 2012, *2*, 7.
35. Sørensen, M. C.; van Alphen, L. B.; Harboe, A.; Li, J.; Christensen, B. B.; Szymanski, C. M.; Brøndsted, L. Bacteriophage F336 recognizes the capsular phosphoramidate modification of *Campylobacter jejuni* NCTC11168. *J. Bacteriol.* 2011, *193*, 6742– 6749.
36. Bayliss, C. D.; Bidmos, F. A.; Anjum, A.; Manchev, V. T.; Richards, R. L.; Grossier, J. P.; Wooldridge, K. G.; Ketley, J. M.; Barrow, P. A.; Jones, M. A.; Tretyakov, M. V. Phase variable genes of *Campylobacter jejuni* exhibit high mutation rates and specific mutational patterns but mutability is not the major determinant of population structure during host colonization. *Nucleic Acids Res.* 2012, *40*, 5876– 5889.

37. Cheong, C. G.; Escalante-Semerena, J. C.; Rayment, I. Structural studies of the L-threonine-O-3-phosphate decarboxylase (CobD) enzyme from *Salmonella enterica*: the apo, substrate, and product-aldimine complexes. *Biochemistry* 2002, *41*, 9079– 9089.

38. Haruyama, K.; Nakai, T.; Miyahara, I.; Hirotsu, K.; Mizuguchi, H.; Hayashi, H.; Kagamiyama, H. Structures of *Escherichia coli* histidinol-phosphate aminotransferase and its complexes with histidinol-phosphate and N-(5'-phosphopyridoxyl)-L-glutamate: double substrate recognition of the enzyme. *Biochemistry* 2001, *40*, 4633– 4644.

39. Andreeßen, B.; Steinbüchel, A. Biotechnological conversion of glycerol to 2-amino-1,3-propanediol (serinol) in recombinant Escherichia coli. *Appl. Microbiol. Biotechnol* 2012, *93*, 357– 365.

40. Bieberich, E.; Kawaguchi, T.; Yu, R. K. N-acylated serinol is a novel ceramide mimic inducing apoptosis in neuroblastoma cells. *J. Biol. Chem.* 2000, *275*, 177– 181.

41. Ueoka, R.; Fujita, T.; Iwashita, T.; van Soest, R. W.; Matsunaga, S. Inconspicamide, new N-acylated serinol from the marine sponge *Stelletta inconspicua*. *Biosci., Biotechnol., Biochem.* 2008, *72*, 3055– 3058.

42. Buranachokpaisan, T.; Dannenfesler, R. M.; Li, P. Compound formulations of 2-amino-1,3-propanediol compounds. WO Patent. WO2006010630A12006. World Intellectual Property Organization (PCT).

43. Kenyon, J. J.; Senchenkova, S.; Shashkov, A. S.; Shneider, M. M.; Popova, A. V.; Knirel, Y. A.; Hall, R. M. K17 capsular polysaccharide produced by *Acinetobacter baumannii* isolate G7 contains an amide of 2-acetamido-2-deoxy-D-galacturonic acid with D-alanine. *Int. J. Biol. Macromol.* 2020, *144*, 857– 862.

44. Landersjö, C.; Weintraub, A.; Widmalm, G. Structure determination of the O-antigen polysaccharide from the enteroinvasive *Escherichia coli* (EIEC) O143 by component analysis and NMR spectroscopy. *Carbohydr. Res.* 1996, *291*, 209– 216.

45. Vinogradov, E. V.; Holst, O.; Thomas-Oates, J. E.; Broady, K. W.; Brade, H. The structure of the O-antigenic polysaccharide from lipopolysaccharide of Vibrio cholerae strain H11 (non-O1). *Eur. J. Biochem.* 1992, *210*, 491– 498.

46. Vinogradov, E. V.; Knirel, Y. A.; Shashkov, A. S.; Kochetkov, N. K. Determination of the degree of amidation of 2-deoxy-2-formamido-D-galacturonic acid in O-specific polysaccharides of Pseudomonas aeruginosa O4 and related strains. *Carbohydr. Res.* 1987, *170*, c1– c4.

47. Vinogradov, E. V.; Krajewska-Pietrasik, D.; Kaca, W.; Shashkov, A. S.; Knirel, Y. A.; Kochetkov, N. K. Structure of *Proteus mirabilis* O27 O-specific polysaccharide containing amino acids and phosphoethanolamine. *Eur. J. Biochem.* 1989, *185*, 645– 650.

48. Reddy, G. P.; Hayat, U.; Xu, Q.; Reddy, K. V.; Wang, Y.; Chiu, K. W.; Morris, J. G., Jr; Bush, C. A. Structure determination of the capsular polysaccharide from *Vibrio vulnificus* strain 6353. *Eur. J. Biochem.* 1998, *255*, 279– 288.

49. Iyer, L. M.; Abhiman, S.; Maxwell Burroughs, A.; Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol. BioSyst.* 2009, *5*, 1636– 1660.

50. Aono, R. The poly-alpha- and -beta-1,4-glucuronic acid moiety of teichuronopeptide from the cell wall of the alkalophilic Bacillus strain C-125. *Biochem. J.* 1990, *270*, 363– 367.

51. Kocharova, N. A.; Ovchinnikova, O. G.; Bialczak-Kokot, M.; Shashkov, A. S.; Knirel, Y. A.; Rozalski, A. Structure of the O-polysaccharide of *Providencia alcalifaciens* O25 containing an amide of D-galacturonic acid with N(ε)-[(R)-1-carboxyethyl]-L-lysine. *Biochemistry (Moscow)* 2011, *76*, 707– 712.

52. Vinogradov, E.; Sidorczyk, Z.; Knirel, Y. A. Structure of the core part of the lipopolysaccharides from *Proteus penneri* strains 7, 8, 14, 15, and 21. *Carbohydr. Res.* 2002, *337*, 643– 649.

53. Zdorovenko, E. L.; Varbanets, L. D.; Liu, B.; Valueva, O. A.; Wang, Q.; Shashkov, A. S.; Garkavaya, E. G.; Brovarskaya, O. S.; Wang, L.; Knirel, Y. A. Structure and gene cluster of the O antigen of *Escherichia coli* L-19, a candidate for a new O-serogroup. *Microbiology* 2014, *160*, 2102– 2107.

4. DISCOVERY AND FUNCTIONAL CHARACTERIZATION OF A CLANDESTINE ATP-DEPENDENT AMIDOLIGASE IN THE BIOSYNTHESIS OF THE CAPSULAR POLYSACCHARIDE FROM *CAMPYLOBACTER JEJUNI*[*]

4.1. Introduction

Campylobacter jejuni is a Gram-negative, zoonotic organism that is responsible for campylobacteriosis (1). *C. jejuni* is often found in poultry where it is a commensal organism, as well as in various farm animals such as cows, goats, and dogs (2). Campylobacteriosis is often contracted from unsanitary drinking water, undercooked or raw poultry, pet feces, and unpasteurized milk (3,4). *C jejuni* is the leading cause of gastroenteritis worldwide and an estimated 25,000 children die each year because of this infection. This disease primarily affects developing nations and is most prevalent in India and Africa (5). The infection can result in fever, cramps, bloody diarrhea, vomiting, and death (6). Additionally, campylobacteriosis can result in a rare autoimmune disease known as Guillain-Barré Syndrome (GBS) (7). It is estimated that 40% of all new GBS cases are preceded by a *Campylobacter* infection (8). Currently, there are no FDA-approved vaccines for the prevention of a *Campylobacter* infection. So far, the best candidates are vaccine conjugates, which mimic the surface-exposed capsular polysaccharides (CPSs) (5). CPSs protect bacteria from desiccation and from

[*] Reprinted with permission from "Discovery and Functional Characterization of a Clandestine ATP-Dependent Amidoligase in the Biosynthesis of the Capsular Polysaccharide from *Campylobacter jejuni*" by Alexander S. Riegert, Tamari Narindoshvili, and Frank M. Raushel, *Biochemistry* (2022) *2* (15), pp 117-124, Copyright 2022 American Chemical Society

complement mediated phagocytosis (9). Additionally, the CPS of *C. jejuni* plays an important role in colonization and invasion of the host organism (10).

The CPS of *Campylobacter jejuni* NCTC 11168 HS:2 is composed of a repeating unit of D-*glycero*-L-*gluco*-heptose, D-glucuronate, D-*N*-acetyl-galactosamine, and D-ribose (11,12). The CPS can be further modified with *O*-methyl phosphoramidylation, *O*-methylation, and amidylation. The amide bond at C6 of the D-glucuronate can be formed with either serinol or ethanolamine. A glucuronamide moiety is also found within the HS:19 serotype of *C. jejuni* RM1285 where the CPS is composed of *N*-acetyl-D-glucosamine and D-glucuronate that is amidated exclusively with serinol (13). The repeating structural units of the three capsules are shown in Figure 18.

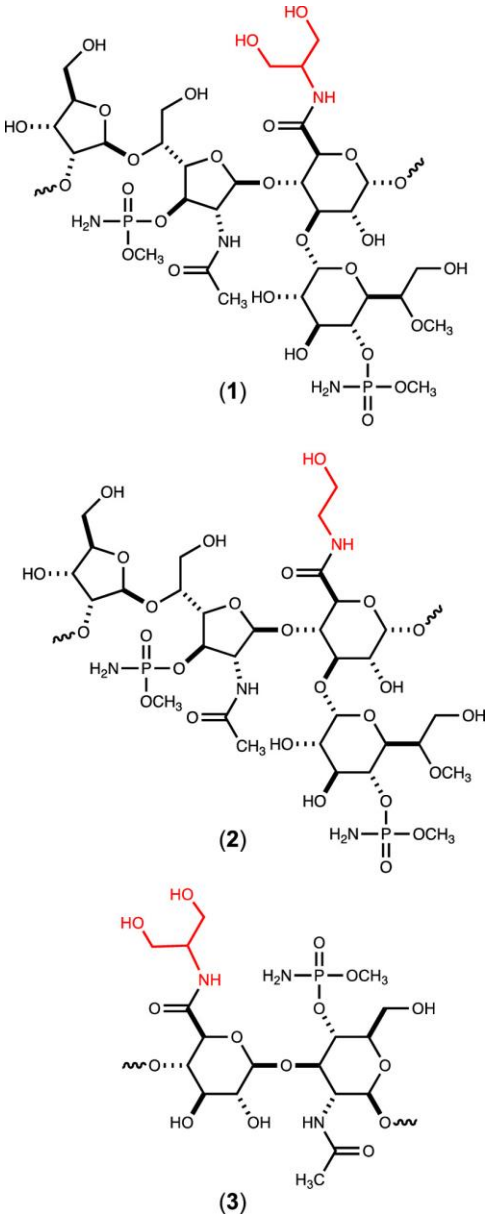


Figure 18 Structures of the repeating CPSs found in *C. jejuni* NCTC 11168 (HS:2) (**1** and **2**) and *C. jejuni* strain RM1285 (HS:19) (**3**). The amide moiety in each structure is highlighted in red. The wavy lines represent the linkage points for the repeating structural units.

Previously, we investigated the roles of the enzymes Cj1441, Cj1436, and Cj1437 in the biosynthetic pathway for the assembly of the CPS from *C. jejuni* NCTC

11168 (14,15). Cj1441 is a UDP-glucose 6-dehydrogenase that catalyzes the dual oxidation of UDP-D-glucose to UDP-D-glucuronate (**4**). The enzyme was crystallized and the protein structure determined to a resolution of 2.1 Å in the presence of UDP-glucose and NAD^+ (14). Cj1436 and Cj1437 were shown to catalyze the formation of the two amines that are ultimately used in amide bond formation. Cj1436 is a PLP-dependent enzyme that catalyzes the decarboxylation of L-serine phosphate to ethanolamine phosphate (**8**), whereas Cj1437 is a PLP-dependent transaminase that catalyzes the stereospecific transamination of dihydroxyacetone phosphate to (*S*)-serinol phosphate (**7**) (15). The gene clusters for the biosynthesis of the CPSs made by *C. jejuni* NCTC 11168 (HS:2) and *C. jejuni* RM1285 (HS:19) are shown in Appendix Figure S12.

Initially, attempts were made to catalyze amide bond formation using Cj1441, the UDP-glucose 6-dehydrogenase. These experiments attempted to capture the putative thioester intermediate, formed during the oxidation of UDP-D-glucose, by the primary amine. This proposal proved unsuccessful, but led to a more rigorous interrogation of the gene cluster for biosynthesis of the CPS. An exhaustive bioinformatics analysis identified Cj1438, an annotated glycosyltransferase, as the primary candidate for amide bond formation (16). The N-terminal domain of Cj1438 is homologous to the GT2 family of glycosyltransferases, while the C-terminal domain of the protein is homologous to proteins previously annotated as "TupA-like" with an ATP-grasp structural fold. TupA from *Bacillus halodurans* was proposed previously to catalyze amide bond formation between D-glucuronate and polyglutamyl-glutamate leading to the

formation of teichuronopeptide (17,18). An ATP-grasp domain is also present in the C-terminal region of *C. jejuni* HS19.11, the homologue to Cj1438 found in the HS:19 serotype. These two proteins (Cj1438 and HS19.11) share a 56% sequence identity across the terminal 378 amino acids. Here, we demonstrate that the C-terminal domain of Cj1438 catalyzes amide bond formation between D-glucuronate and either (*S*)-serinol phosphate or ethanolamine phosphate in the presence of MgATP.

4.2. Materials and Methods

4.2.1. Materials

Lactate dehydrogenase and pyruvate kinase were purchased from Sigma-Aldrich. Ethanolamine phosphate, D-glucuronate, ATP, NADH, $MgCl_2$, and 2′-deoxyribose were purchased from Sigma-Aldrich. Methyl-β-D-glucuronic acid (**6**) was acquired from Carbosynth. The chemical synthesis of (*R*/*S*) serinol phosphate (**7**) and (*S*)-serinol phosphate (**8**) was conducted as described previously (15). The structures of the substrates and products made for this investigation are presented in Figure 19.

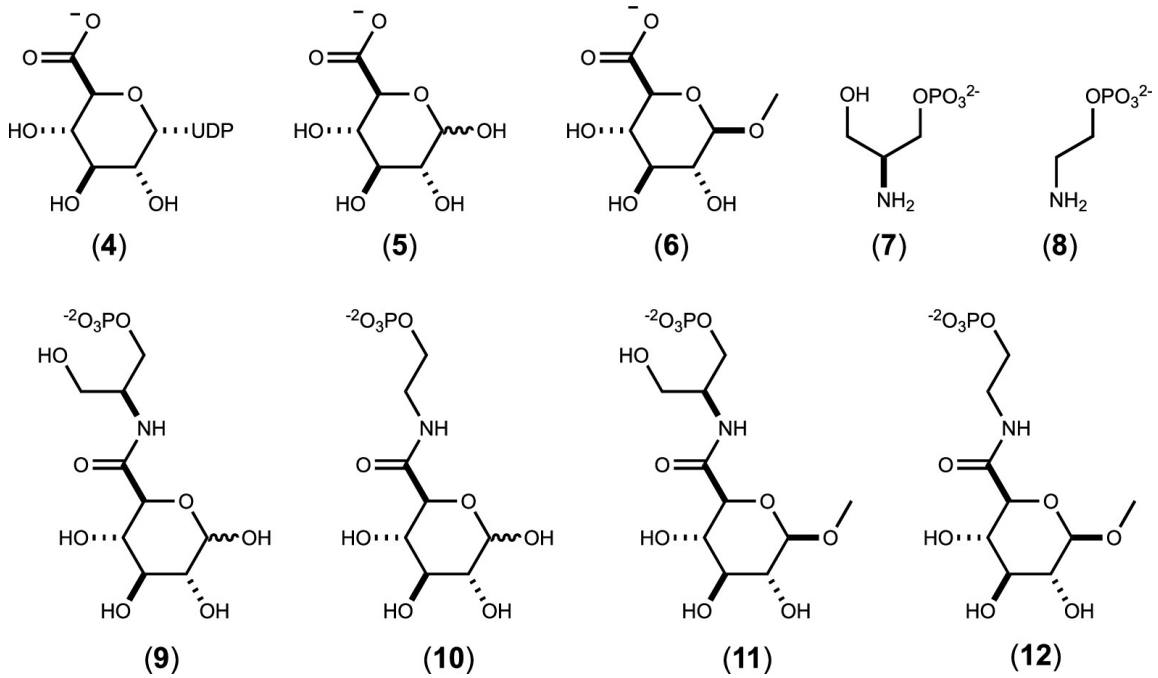


Figure 19 Structures of potential substrates and products for the reaction catalyzed by the C-terminal domain of Cj1438.

4.2.2. Preparation of Cj1438

The gene encoding Cj1438 (UniProt id: Q0P8H6) from serotype HS:2 of *C. jejuni* was purchased from Twist Bioscience with codon optimization for expression in *E. coli* K12. A fragment of this gene was subcloned into pET31b for expression of amino acids 399–776 with a C-terminal hexahistidine tag. The complete amino acid sequence of the expressed protein is shown in Appendix Figure S13a. The pET31b vector was used to transform BL21 *E. coli* cells (Novagen) via electroporation. Cells harboring the pET31b-Cj1438$_{(399–776)}$ plasmid were cultured in Terrific Broth (TB) with 100 mg/L ampicillin. The cells were grown at 37 °C with shaking and induced with 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the optical density reached ~0.8 at 600 nm. The cells were allowed to express protein at 16 °C for 18 h after

induction and then harvested by centrifugation at 15,000 rcf at 4 °C. The cell pellet was resuspended in loading buffer (50 mM HEPES/K^+, 300 mM KCl, 20 mM imidazole, pH 8.0, and 1% Triton X) and lysed by sonication. The sonicated cells were centrifuged at 25,000 rcf at 4 °C before the lysate was passed through a 0.45 μm filter (Whatman). The sample was loaded onto a prepacked 5 mL HisTrap (GE Healthcare) nickel affinity column. The protein was eluted from the column using 50 mM HEPES/K^+, pH 8.0, 300 mM KCl, and 250 mM imidazole over a gradient of 25 column volumes. The protein was pooled and dialyzed against 10 mM HEPES/K^+ pH 8.0 and 200 mM KCl. The C-terminal domain of Cj1438 was concentrated to 1.25 mg/mL and flash-frozen using liquid nitrogen before being stored at −80 °C. Approximately 2.0 mg of protein was obtained per liter of cell culture.

A modified expression plasmid for the production of a shorter C-terminal domain fragment was also constructed from the gene for Cj1438. This plasmid (pET31b-$Cj1438_{(453–776)}$) codes for the expression of amino acids 453 through 776. The protein ($Cj1438_{(453–776)}$) was purified in an identical manner as that for $Cj1438_{(399–776)}$. The amino acid sequence of the isolated protein is presented in Appendix Figure S13b. The $Cj1438_{(453–776)}$ construct increased the yield and solubility of the purified protein.

4.2.3. Bioinformatics Analysis of Cj1438

The amino acid sequence of the full-length Cj1438 gene was used as the input sequence for the online transmembrane predictor tool, TMpred (https://embnet.vital-it.ch/software/TMPRED_form.html) (19). The output of this analysis was the starting

point for where to make truncations. A truncated version of Cj1438 (residues 399 through 776) was used as the initial input sequence for the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST) (20,21). The search was limited to the 1000 closest protein sequences based on an all-by-all BLAST of the UniProt database. A sequence similarity network (SSN) was created for the 1000 genes at a sequence identity cutoff of 46% using the yFile organic layout in Cystoscape 3.8.2.

4.2.4. Reaction Product Confirmation Using Mass Spectrometry

$Cj1438_{(399–776)}$ (10 μM) was incubated with 5.0 mM ATP, 10 mM $MgCl_2$, 5.0 mM ethanolamine phosphate, 5.0 mM D-glucuruonic acid, and 50 mM ammonium bicarbonate at pH 8.0. The reaction was repeated substituting ethanolamine phosphate (**8**) with (*S*)-serinol phosphate (**7**). The reaction was allowed to incubate at 25 °C for 1 h. The reaction mixture was passed through a GE Healthcare Vivaspin 500 10 kDA MWCO filter and 100 μL of the flow-through was submitted for mass spectrometry analysis using a Thermo Scientific Q Exactive Focus mass spectrometer.

4.2.5. ^{31}P NMR Spectroscopy of Cj1438-Catalyzed Reaction Products

$Cj1438_{(399–776)}$ (2.0 μM) was incubated with 5.0 mM ATP, 5.0 mM D-glucuronate, 10 mM $MgCl_2$, 5.0 mM ethanolamine phosphate, and 50 mM HEPES/K^+, pH 8.0. The ^{31}P NMR spectra were collected over the course of 2 h using an Avance III 400 MHz NMR spectrometer. The reaction was repeated with the substitution of serinol phosphate (**7**) for ethanolamine phosphate (**8**).

4.2.6. Determination of Kinetic Constants for Cj1438

The kinetic constants for the reaction catalyzed by $Cj1438_{(453–776)}$ were determined spectrophotometrically using a lactate dehydrogenase/pyruvate kinase (LDH/PK) coupled assay. The oxidation of NADH to NAD^+ was monitored spectrophotometrically at 340 nm and 25 °C using a Spectramax340 UV–visible spectrophotometer. The concentration of NADH was calculated using an extinction coefficient of 6220 M^{-1} cm^{-1}. Initial rates were determined at fixed concentrations of ATP and D-glucuronate with ethanolamine phosphate (**8**) as the variable substrate. The assay contained 6.0 mM $MgCl_2$, 3.0 mM ATP, 2.0 mM phosphoenolpyruvate (PEP), 250 μM NADH, 10 mM D-glucuronate, 5.0 units of lactate dehydrogenase, 5.0 units of pyruvate kinase, 50 mM HEPES/K^+, and variable concentrations of ethanolamine phosphate (0.0025–2.0 mM). The reaction was initiated by the addition of 50 nM $Cj1438_{(453–776)}$. The assay was repeated with (*S*)-serinol phosphate (**7**) as the variable substrate (0.0025–8.0 mM). To determine if UDP-D-glucuronate was a substrate, the assay was conducted with variable amounts of UDP-D-glucuronate (0.01–2.0 mM). The kinetic parameters were determined by fitting the initial rates to eq 1 using GraphPad Prism, where ν is the initial velocity of the reaction, k_{cat} is the turnover number, [*A*] is the substrate concentration, and K_m is the Michaelis constant.

$$v / E_t = k_{cat} [A] / (A + K_m) \quad (1)$$

For reactions that exhibit partial substrate inhibition, the kinetic parameters were determined by fitting the initial rates to eq 2 using GraphPad Prism, where v is the initial velocity, k_{cat} is the turnover number, [A] is the substrate concentration, K_m is the Michaelis constant, and K_i is the inhibition constant.

$$v / E_t = k_{cat} [A] / (K_m + A (1 + A / K_i)) \quad (2)$$

4.2.7. Determination of Stereoselectivity of Cj1438

The stereoselectivity of $Cj1438_{(399–776)}$ was tested using ^{31}P NMR spectroscopy. The enzyme (4.0 μM) was incubated with 3.0 mM (*R*/*S*)-serinol phosphate, 10 mM ATP, 10 mM D-glucuronate, 20 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. The reaction was allowed to incubate at 25 °C for 30 min. The ^{31}P NMR spectra were collected using an Avance III 400 MHz NMR spectrometer.

4.3. Results

4.3.1. Bioinformatics Analysis of Cj1438

The full-length sequence of Cj1438 (residues 1–776) was used as the initial query for a BLAST search of the NCBI database. The first 172 residues of Cj1438 are homologous to the GT2 family of glycosyltransferases (pfam00535), whereas the C-terminal 217 amino acids are homologous to the TupA-like amidoligases (pfam14305) (22,23). TupA from *Bacillus halodurans* apparently catalyzes the ligation of poly-glutamyl glutamate and D-glucuronate (17,18). The full-length sequence of

Cj1438 was also used as the search query against the UniProtKB/SwissProt database. The protein with the best e-value ($5e^{-26}$) was EpsJ from *Bacillus subtilis* str. 168, which is 32% identical across the first 220 residues of the glycosyltransferase domain. There are no functionally characterized proteins which match the putative ATP-grasp amidoligase domain from the C-terminal end of Cj1438. When the full-length Cj1438 protein was used as the search query against the proteins in the PDB, no structures which match the ATP-grasp domain were identified. The closest structural match is CgT from *Streptococcus parasanguinis* FW213, which shares a 27% sequence identity across the first 211 residues of the glycosyltransferase domain.

To further investigate the ATP-grasp domain of Cj1438, an SSN was constructed. The 1000 closest protein sequences to $Cj1438_{(399–776)}$ from an all-by-all BLAST of UniProt is presented in Figure 20. The cutoff for edge formation was set at 46%. The *C. jejuni* serotypes HS:2, HS:19, HS:22, and HS:38 are shown as larger red, dark blue, green, and purple circles, respectively. Serotypes HS:2 and HS:19 have a known CPS structure that contains a D-glucuronamide molecule. Serotypes HS:22 and HS:38 do not have a known CPS structure, but their genomes have been sequenced and they contain genes which closely resemble those found in HS:2 and HS:19, including UDP-glucose dehydrogenase, a PLP-dependent transaminase, and a HAD phosphatase. Shown in yellow circles are proteins from other strains of *Campylobacter*. Enzymes that have Uniprot descriptions as TupA-like ATP-grasp enzymes are shown as orange circles. The blue nodes are representative of enzymes annotated as either glycosyl transferases or uncharacterized proteins. There are no functionally characterized

proteins in this network. From this network we can conclude that the enzyme necessary for amide bond formation found on the surface of *C. jejuni* NCTC 11168 (HS:2) is relatively common in other strains of *C. jejuni*.

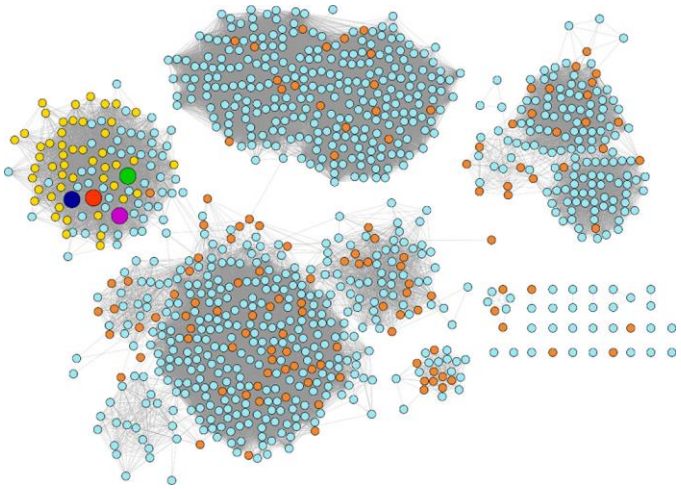


Figure 20 SSN for Cj1438. This SSN shows the 1000 closest sequences to the C-terminal domain of Cj1438 (residues 399–776) at a sequence identity of 46%. Cj1438 (HS:2) is shown as a large red circle. The corresponding genes from serotypes HS:19, HS:38, and HS:22 are shown as large dark blue, green, and purple circles, respectively. The orange nodes represent proteins described as TupA-like ATP-grasp enzymes. The yellow nodes represent other strains of *Campylobacter*. The light blue nodes are described as either glycosyltransferases or uncharacterized proteins.

4.3.2. Determination of Reaction Products Catalyzed by Cj1438

The products of the reaction catalyzed by the C-terminal domain of Cj1438 were determined by mass spectrometry. After incubation of $Cj1438_{(399–776)}$ with ATP, D-glucuronate (**5**), and ethanolamine phosphate (**8**), a new peak at an *m*/*z* of 316.04 appears in the mass spectrum (Figure 21a). This peak corresponds to that expected for the M-H anion of the glucuronamide with ethanolamine phosphate (**10**). The reaction was subsequently conducted in the presence of serinol phosphate (**7**) and a new peak at an *m*/*z* of 346.05 (Figure 21b), which corresponds to the formation of the glucuronamide

with serinol phosphate (**9**). The formation of ligated products was also tested using methyl-β-D-glucuronate (**6**) as a potential substrate. Incubation of Cj1438$_{(399–776)}$ with methyl-β-D-glucuronate and ethanolamine phosphate exhibited a new product peak at an *m*/*z* of 330.05 corresponding to the expected glucuronamide (**12**) (Figure 21c). When Cj1438 was tested with methyl-β-D-glucuronate (**5**) and (*S*)-serinol phosphate, a new product peak appears at an *m*/*z* of 360.07 (**11**), consistent with amide bond formation between the two substrates (Figure 21d).

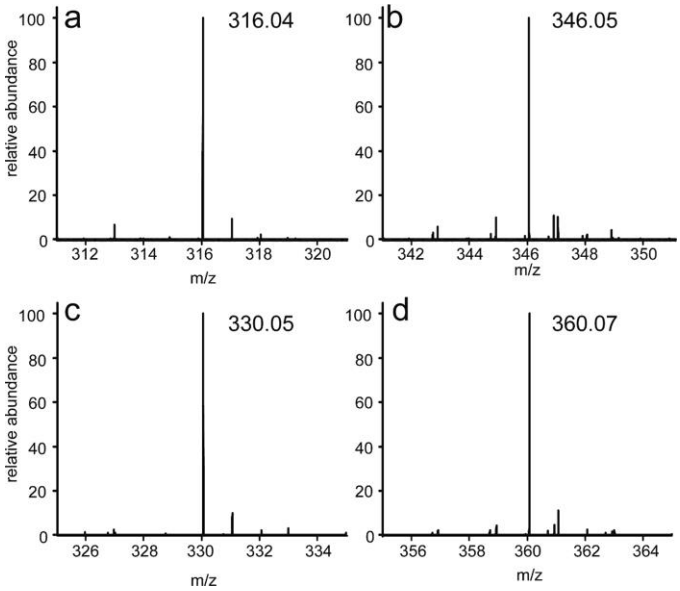


Figure 21 Mass spectrometry of the reaction products catalyzed by Cj1438$_{(399–776)}$. (a) Product (**10**) of the reaction catalyzed by Cj1438$_{(399–776)}$ using MgATP, d-glucuronate, and ethanolamine phosphate has an *m*/*z* of 316.04 for the M–H anion. (b) Product (**9**) of the reaction catalyzed by Cj1438 in the presence of MgATP, d-glucuronate, and serinol phosphate has an *m*/*z* of 346.05 for the M–H anion. (c) Product (**12**) of the reaction catalyzed by Cj1438 in the presence of methyl-β-d-glucuronate (**6**) and ethanolamine phosphate has an *m*/*z* of 330.05 for the M–H anion. (d) Product (**11**) of the reaction catalyzed by Cj1438 in the presence of **6**, MgATP, and serinol phosphate has an *m*/*z* of 360.07 for the M–H anion.

4.3.3. ^{31}P NMR Spectroscopic Analysis of the Cj1438 Reaction Products

The substrates and products of the reaction catalyzed by $Cj1438_{(399–776)}$ were analyzed with ^{31}P NMR spectroscopy. A control solution composed of D-glucuronate, MgATP, and ethanolamine phosphate, in the absence of added enzyme, is illustrated in Figure 22a. The resonance for (*S*)-serinol phosphate appears at 3.85 ppm. When enzyme is added to the reaction mixture, the serinol phosphate peak disappears and resonances for two new products appear at 3.97 and 2.15 ppm (Figure 22b). The resonance at 3.97 ppm represents the glucuronamide of serinol phosphate (**9**), while the resonance at 2.15 ppm is from inorganic phosphate. The control spectrum for the reaction mixture containing D-glucuronate, MgATP, and ethanolamine phosphate is shown in Figure 22c. A single resonance is observed at 3.78 ppm for the ethanolamine phosphate (**8**). After the addition of the enzyme, the resonances for two new products appear at 3.74 and 2.15 ppm (Figure 22d). The newly formed resonance at 3.74 ppm represents the glucuronamide of ethanolamine phosphate (**12**), while the resonance at 2.15 is inorganic phosphate.

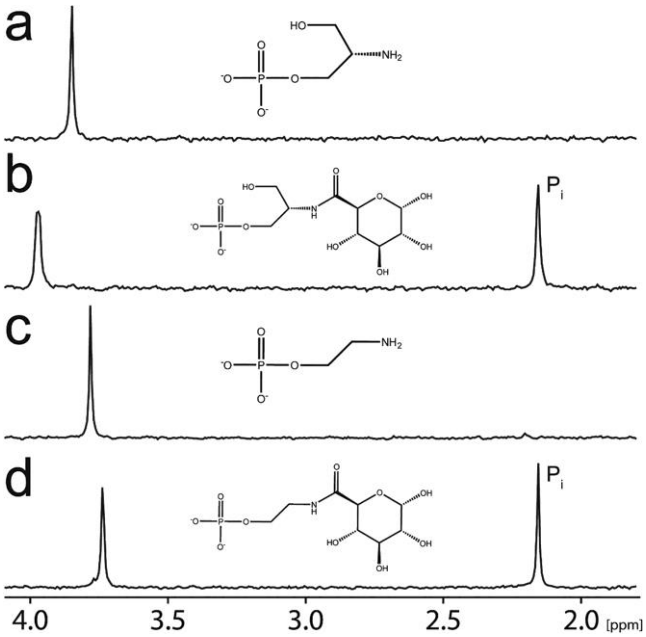


Figure 22 ^{31}P NMR spectra of the reactions catalyzed by Cj1438$_{(399–776)}$. (a) Resonance for (*S*)-serinol phosphate (**7**) appears at 3.85 ppm. (b) Enzyme-catalyzed reaction of (*S*)-serinol phosphate with d-glucuronate with the resonance for the glucuronamide phosphate (**9**) as a singlet at 3.97 ppm. (c) No-enzyme control reaction of ethanolamine phosphate. Ethanolamine phosphate (**8**) resonates at 3.78 ppm. (d) Enzyme-catalyzed reaction of Cj1438$_{(399–776)}$ using d-glucuronate and ethanolamine phosphate. The resonance for the glucuronamide phosphate (**10**) appears at 3.74 ppm. In panels b and d, inorganic phosphate is at 2.15 ppm.

4.3.4. Stereoselectivity of Cj1438

To determine if Cj1438 is stereoselective toward the utilization of either the (*R*)- or (*S*)-enantiomer of serinol phosphate, we used ^{31}P NMR spectroscopy. A control reaction in the absence of enzyme shows the ^{31}P resonance for (*R*/*S*)-serinol phosphate at 3.78 ppm (Figure 23a). After the addition of 4.0 μM enzyme and 3.0 mM (*R*/*S*)-serinol phosphate in the presence of an excess of ATP and D-glucuronate, two new resonances appeared at 3.90 ppm and 2.20 ppm for the newly formed glucuronamide of serinol phosphate (**9**) and phosphate (Figure 23b). Integration of the three resonances indicates

a molecular ratio of approximately 1.0:0.98:1.06. This result indicates that approximately one half of the initial substrate was consumed and converted to the glucuronamide of serinol phosphate, while the other half remained unreacted. To determine which enantiomer is consumed, we chemically synthesized the (*S*)-enantiomer of serinol phosphate and tested it in a similar manner. After the addition of Cj1438$_{(399–776)}$, two new two resonances appear at 3.97 and 2.15 ppm and all of the original serinol phosphate was consumed (Figure 22b). The resonance at 3.90 ppm represents the glucuronamide of (*S*)-serinol phosphate (**9**), while the resonance at 2.15 ppm represents inorganic phosphate. Upon integration, the two resonances are formed in a ratio of 1:1. These results demonstrate that Cj1438 is specific for the (*S*)-enantiomer of serinol phosphate.

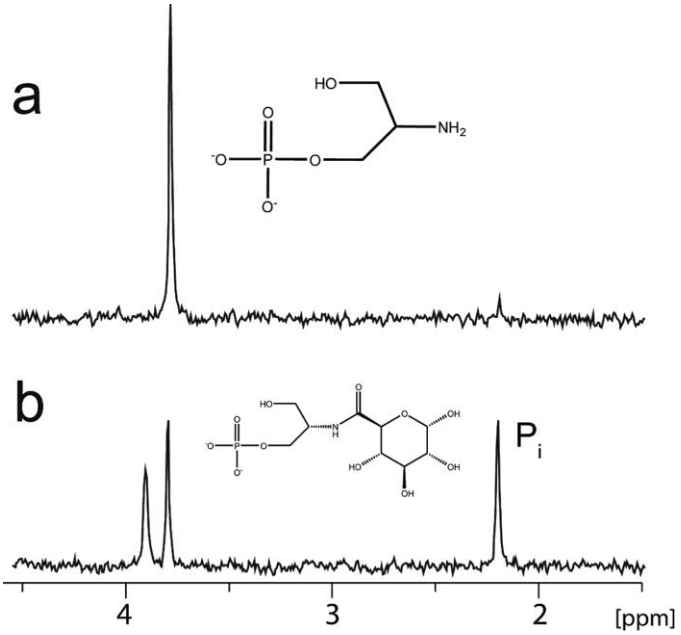


Figure 23 Stereoselectivity of Cj1438$_{(399–776)}$. (a) ^{31}P NMR spectrum of (*R*/*S*)-serinol phosphate. (b) Reaction catalyzed by Cj1438$_{(399–776)}$ in the presence of (*R*/*S*)-serinol phosphate and excess MgATP and D-glucuronate. The newly formed product appears at 3.90 ppm. The enantiomer that is not used remains at 3.78 ppm. The resonance for inorganic phosphate appears at 2.20 ppm.

4.3.5. Kinetic Analysis of Cj1438

The reaction catalyzed by Cj1438 was monitored spectrophotometrically at 340 nm using a lactate dehydrogenase/pyruvate kinase coupled assay. A summary of the kinetic constants is found in Table 6. Reactions containing fixed amounts of ATP and ethanolamine phosphate with variable levels of glucuronate gave an apparent k_{cat} of 2.7 ± 0.1 s^{-1}, a K_m of 1.3 ± 0.1 mM, and a k_{cat}/K_m of (2.1 ± 0.1) × 10^3 M^{-1} s^{-1}. Using fixed amounts of glucuronate (10 mM) and ethanolamine phosphate (3.0 mM) with variable amounts of ATP (0.0025–2.0 mM) the apparent k_{cat} is 2.9 ± 0.1 s^{-1}, the K_m is 47 ± 4 μM, and the k_{cat}/K_m is (6.2 ± 0.5) × 10^4 M^{-1} s^{-1}. The reaction of Cj1438 with variable concentrations of ethanolamine phosphate (0.0025–2.0 mM) and fixed concentrations of

ATP (3.0 mM) and glucuronate (10 mM) has an apparent k_{cat} of 3.0 ± 0.1 s^{-1}, K_m of 260 ± 12 μM, and k_{cat}/K_m of (1.2 ± 0.4) × 10^4 M^{-1} s^{-1}. The reaction of Cj1438 with variable concentrations of (*S*)-serinol phosphate (0.0025–8.0 mM) and fixed concentrations of ATP (3.0 mM) and glucuronate (10 mM) has an apparent k_{cat} of 2.9 ± 0.1 s^{-1}, a K_m of 57 ± 5, and a k_{cat}/K_m of (5.1 ± 0.5) × 10^4 M^{-1} s^{-1}. Substrate inhibition of the enzyme was observed for the (*S*)-serinol phosphate with a K_i of 16 ± 3 mM shown in Appendix Figure S14. The (*R*/*S*)-serinol phosphate also exhibited substrate inhibition with a K_i of 19 ± 2 mM shown in Appendix Figure S15.

When the kinetic assays were conducted with UDP-D-glucuronate (**4**) instead of D-glucuronate (**5**), Cj1438 did not catalyze at a rate above the background hydrolysis of ATP. We also tried the addition of 2′-D-deoxyribose to mimic a putative disaccharide substrate, (α-D-glucuronate-(1 → 2) β-D-ribose). The addition of a fixed concentration of 2-deoxy-D-ribose (10 mM) did not enhance the rate of ADP formation.

Table 6 Kinetic Constants for the Reactions Catalyzed by Cj1438[a]

fixed substrates	**variable substrate**	**app k_{cat} (s^{-1})**	**app K_m (μM)**	**app k_{cat}/K_m (M^{-1} s^{-1})**
ATP (3.0 mM)[b] EtnP (3.0 mM)[b]	D-glucuronate	2.7 ± 0.1	1300 ± 110	(2.1 ± 0.1) × 10^3
GlcA (10 mM)[b] EtnP (3.0 mM)	ATP	2.9 ± 0.1	47 ± 4	(6.2 ± 0.5) × 10^4

ATP (3.0 mM) GlcA (10 mM)	ethanolamine-P	3.0 ± 0.1	260 ± 12	(1.2 ± 0.4) × 104
ATP (3.0 mM) GlcA (10 mM)	(*S*)-serinol-P	2.9 ± 0.1	57 ± 5	(5.1 ± 0.5) × 10^4
ATP (3.0 mM) GlcA (10 mM)	(*R*/*S*)-serinol-P[c]	2.8 ± 0.1	120 ± 5	(2.3 ± 0.2) × 10^4
ATP (3.0 mM) EtnP (3.0 mM) 2′-deoxyribose (10 mM)	D-glucuronate	2.4 ± 0.1	1300 ± 80	(1.8 ± 0.1) × 10^3
ATP (3.0 mM) MeGlcA (10 mM)[b]	ethanolamine-P	2.5 ± 0.1	230 ± 15	(1.1 ± 0.1) × 10^4
ATP (3.0 mM) MeGlcA (10 mM)	(*S*)-serinol-P[c]	2.6 ± 0.1	46 ± 6	(5.7 ± 0.8) × 10^4

[a]Reactions conducted at 25 °C, pH 8.0.
[b]ATP, adenosine triphosphate; EtnP, ethanolamine phosphate; GlcA, D-glucuronate; MeGlcA, methyl-β-D-glucuronate.
[c]Substrate inhibition observed and fit to eq 2.

4.4. Discussion

Campylobacter jejuni is an emerging global pathogen, which is the leading cause of gastroenteritis worldwide (5). The bacterium is commensal in chickens and can easily be spread to other farm animals and humans through contaminated drinking water, uncooked meat, unpasteurized milk, and pet feces (3,4). There is currently no FDA-

approved vaccine for campylobacteriosis, but the most promising vaccine candidates have been conjugate vaccines which utilize surface-exposed sugars (5).

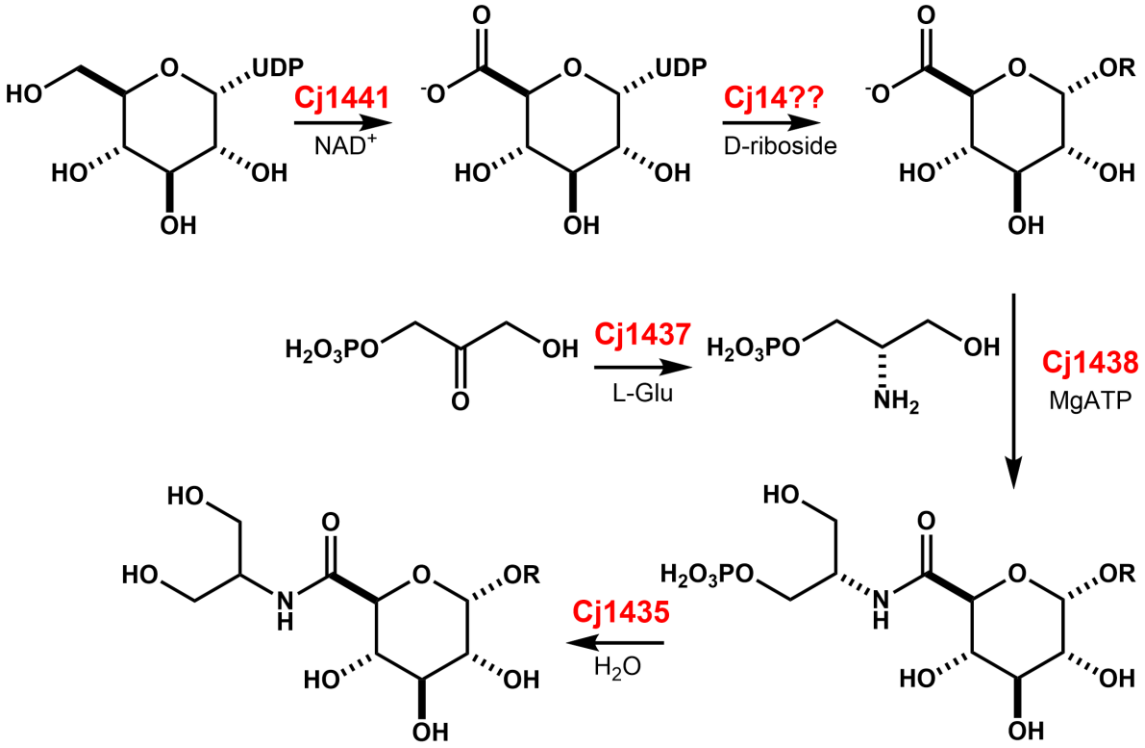
C. jejuni is coated with a thick layer of sugar molecules known as the CPS. Understanding how the sugars of the polysaccharide are made and modified may prove useful for vaccine synthesis and development. *C. jejuni* NCTC 11168 (HS:2) has a CPS containing an unusual amide bond consisting of α-D-glucuronate with either ethanolamine or serinol. Only one other strain of *C. jejuni* has a known capsule with a similar modification (serotype HS:19), although other strains will certainly bear this modification based on the known DNA sequences of other serotypes of *C. jejuni*. Similar glucuronamides are found in lipopolysaccharides, CPSs, and O-antigens of many bacteria, including *Proteus penneri* strain 14, *Providencia stuartii* O25, *Bacillus halodurans*, *Vibrio vulnificus* strain 6353, *Proteus mirabilis* O27, *Pseudomonas aeruginosa*, *Shigella boydii* type 8, *Vibrio cholerae* H11 (non-O1), *E. coli* L-19, *E. coli* O143, and *Acinetobacter baumannii* G7 (16,17,24−32).

Using bioinformatics, nuclear magnetic resonance, and mass spectrometry, we have now shown that the amide bond found within the CPS of *C. jejuni* is synthesized using the TupA-like ATP-grasp domain from the C-terminal end of Cj1438. Cj1438 is apparently composed to two functional domains: an N-terminal glycosyltransferase domain and a C-terminal amidoligase domain. Expression of the C-terminal truncated version of Cj1438 enabled us to demonstrate that the amide bond is formed using ATP, D-glucuronate, and either ethanolamine phosphate or (*S*)-serinol phosphate. The products of these reactions were confirmed by high resolution mass spectrometry and ^{31}P

NMR spectroscopy. Additionally, we demonstrated that D-glucuronate can be substituted with methyl-β-D-glucuronate (**6**); however, UDP-glucuronate (**4**) is not a substrate for the reaction. This result suggests that the amide bond in the CPS of *C. jejuni* is not made until after the D-glucuronate moiety is transferred to the growing polysaccharide chain. It is unclear which glycosyltransferase in the CPS gene cluster is responsible for the formation of the glycosidic bond between D-glucuronate and D-ribose in the growing polysaccharide chain. However, we predict that the N-terminal domain of Cj1438 is responsible for the binding of UDP-D-glucuronate and the nucleophilic attack of the C2 hydroxyl group from D-ribose. Other possible candidates include Cj1442, Cj1440, Cj1434, and Cj1432.

The reaction diagram shown in Scheme 19 illustrates the probable path of formation for D-glucuronamide in *C. jejuni*. We have shown that Cj1441 acts as a UDP-glucose 6-dehydrogenase to make UDP-glucuronate (**4**) (14). This molecule is the most likely substrate for the glycosyltransferase that creates the glycosidic bond between D-glucuronate and the C2-hydroxyl from D-ribose at the growing end of the developing polysaccharide chain. The C-terminal domain of Cj1438 catalyzes amide bond formation between the D-glucuronate moiety and (*S*)-serinol phosphate to make the corresponding glucuronamide phosphate. Cj1435, a putative HAD phosphatase, will subsequently hydrolyze the phosphate ester bond creating the glucuronamide molecule found on the CPS. Experiments to test this hypothesis are currently in progress.

Scheme 19 Proposed Reaction Pathway for the Formation of the D-Glucuronamide Found on the Surface of *C. jejuni* NCTC 11168 (HS:2)



4.5. Conclusions

We report the functional characterization of the TupA-like amidoligase from the C-terminal domain of Cj1438 from *C. jejuni* NCTC 11168 (HS:2). This enzyme is responsible for the formation of an amide bond formed between D-glucuronate and either (*S*)-serinol phosphate or ethanolamine phosphate. These results suggest that other enzyme activities are likely embedded within the large multidomain proteins currently annotated as glycosyl transferases in the CPS gene clusters of *C. jejuni*.

4.6. References

1. Epps, S. V.; Harvey, R. B.; Hume, M. E.; Phillips, T. D.; Anderson, R. C.; Nisbet, D. J. Foodborne *Campylobacter*: infections, metabolism, pathogenesis and reservoirs. *Int. J. Environ. Res. Public Health* 2013, *10*, 6292– 6304.
2. Burnham, P. M.; Hendrixson, D. R. *Campylobacter jejuni*: collective components promoting a successful enteric lifestyle. *Nat. Rev. Microbiol.* 2018, *16*, 551– 565.
3. Kaakoush, N. O.; Castaño-Rodríguez, N.; Mitchell, H. M.; Man, S. M. Global Epidemiology of *Campylobacter* Infection. *Clin. Microbiol. Rev.* 2015, *28*, 687– 720.
4. Jaakkonen, A.; Kivistö, R.; Aarnio, M.; Kalekivi, J.; Hakkinen, M. Persistent contamination of raw milk by *Campylobacter jejuni* ST-883. *PLoS One* 2020, *15*, e0231810.
5. Poly, F.; Noll, A. J.; Riddle, M. S.; Porter, C. K. Update on *Campylobacter* vaccine development. *Hum. Vaccines Immunother.* 2019, *15*, 1389– 1400.
6. Acheson, D.; Allos, B. M. *Campylobacter jejuni* Infections: update on emerging issues and trends. *Clin. Infect. Dis.* 2001, *32*, 1201– 1206.
7. Nachamkin, I.; Allos, B. M.; Ho, T. *Campylobacter* species and Guillain-Barré syndrome. *Clin. Microbiol. Rev.* 1998, *11*, 555– 567.
8. Yuki, N. Molecular mimicry between gangliosides and lipopolysaccharides of *Campylobacter jejuni* isolated from patients with Guillain-Barré Syndrome and Miller Fisher Syndrome. *J. Infect. Dis.* 1997, *176*, S150– S153.

9. Guerry, P.; Poly, F.; Riddle, M.; Maue, A. C.; Chen, Y.-H.; Monteiro, M. A. *Campylobacter* polysaccharide capsules: virulence and vaccines. *Front. Cell. Infect. Microbiol.* 2012, *2*, 7.
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.; Szymanski, C. M. Biological roles of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. *PLoS One* 2014, *99*, e87051.
11. St. Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. The structures of the lipooligosaccharide and capsule polysaccharide of *Campylobacter jejuni* genome sequenced strain NCTC 11168. *Eur. J. Biochem.* 2002, *269*, 5119– 5136.
12. Young, K. T.; Davis, L. M.; Dirita, V. J. *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* 2007, *5*, 665– 679.
13. McNally, D. J.; Jarrell, H. C.; Khieu, N. H.; Li, J.; Vinogradov, E.; Whitfield, D. M.; Szymanski, C. M.; Brisson, J. R. The HS:19 serostrain of *Campylobacter jejuni* has a hyaluronic acid-type capsular polysaccharide with a nonstoichiometric sorbose branch and O-methyl phosphoramidate group. *FEBS J.* 2006, *273*, 3975– 3989.
14. Riegert, A. S.; Raushel, F. M. Functional and Structural Characterization of the UDP-Glucose Dehydrogenase Involved in Capsular Polysaccharide Biosynthesis from *Campylobacter jejuni*. *Biochemistry* 2021, *60*, 725– 734.

15. Riegert, A. S.; Narindoshvili, T.; Coricello, A.; Richards, N. G. J.; Raushel, F. M. Functional Characterization of Two PLP-Dependent Enzymes Involved in Capsular Polysaccharide Biosynthesis from *Campylobacter jejuni*. *Biochemistry* 2021, *60*, 2836– 2843.
16. Iyer, L. M.; Abhiman, S.; Maxwell Burroughs, A.; Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol. BioSyst.* 2009, *5*, 1636– 1660.
17. Aono, R. The poly-alpha- and -beta-1,4-glucuronic acid moiety of teichuronopeptide from the cell wall of the alkalophilic Bacillus strain C-125. *Biochem. J.* 1990, *270*, 363– 367.
18. Aono, R.; Ito, M.; Machida, T. Contribution of the cell wall component teichuronopeptide to pH homeostasis and alkaliphily in the alkaliphile Bacillus lentus C-125. *J. Bacteriol.* 1999, *181*, 6600– 6606.
19. Hofmann, K.; Stoffel, W. TMbase – A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* 1993, *374*, 166.
20. Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta, Proteins Proteomics* 2015, *1854*, 1019– 1037.

21. Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One* 2009, *44*, e4345.

22. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M. K.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Marchler, G. H.; Song, J. S.; Thanki, N.; Yamashita, R. A.; Yang, M.; Zhang, D.; Zheng, C.; Lanczycki, C. J.; Marchler-Bauer, A. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020, *48*, D265– D268.

23. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2013, *42*, D490– D495.

24. Zdorovenko, E. L.; Varbanets, L. D.; Liu, B.; Valueva, O. A.; Wang, Q.; Shashkov, A. S.; Garkavaya, E. G.; Brovarskaya, O. S.; Wang, L.; Knirel, Y. A. Structure and gene cluster of the O antigen of *Escherichia coli* L-19, a candidate for a new O-serogroup. *Microbiology* 2014, *160*, 2102– 2107.

25. Vinogradov, E.; Sidorczyk, Z.; Knirel, Y. A. Structure of the core part of the lipopolysaccharides from *Proteus penneri* strains 7, 8, 14, 15, and 21. *Carbohydr. Res.* 2002, *337*, 643– 649.

26. Kocharova, N. A.; Ovchinnikova, O. G.; Bialczak-Kokot, M.; Shashkov, A. S.; Knirel, Y. A.; Rozalski, A. Structure of the O-polysaccharide of *Providencia alcalifaciens* O25 containing an amide of D-galacturonic acid with N(ε)-[(R)-1-carboxyethyl]-L-lysine. *Biochemistry* 2011, *76*, 707– 712.

27. Reddy, G. P.; Hayat, U.; Xu, Q.; Reddy, K. V.; Wang, Y.; Chiu, K. W.; Morris, J. G., Jr.; Bush, C. A. Structure determination of the capsular polysaccharide from *Vibrio vulnificus* strain 6353. *Eur. J. Biochem.* 1998, *255*, 279– 288.
28. Vinogradov, E. V.; Krajewska-Pietrasik, D.; Kaca, W.; Shashkov, A. S.; Knirel, Y. A.; Kochetkov, N. K. Structure of *Proteus mirabilis* O27 O-specific polysaccharide containing amino acids and phosphoethanolamine. *Eur. J. Biochem.* 1989, *185*, 645– 650.
29. Vinogradov, E. V.; Knirel, Y. A.; Shashkov, A. S.; Kochetkov, N. K. Determination of the degree of amidation of 2-deoxy-2-formamido D-galacturonic acid in O-specific polysaccharides of *Pseudomonas aeruginosa* O4 and related strains. *Carbohydr. Res.* 1987, *170*, c1– c4.
30. Vinogradov, E. V.; Holst, O.; Thomas-Oates, J. E.; Broady, K. W.; Brade, H. The structure of the O-antigenic polysaccharide from lipopolysaccharide of *Vibrio cholerae* strain H11 (non-O1). *Eur. J. Biochem.* 1992, *210*, 491– 498.
31. Landersjö, C.; Weintraub, A.; Widmalm, G. Structure determination of the O-antigen polysaccharide from the enteroinvasive *Escherichia coli* (EIEC) O143 by component analysis and NMR spectroscopy. *Carbohydr. Res.* 1996, *291*, 209– 216.
32. Kenyon, J. J.; Senchenkova, S.; Shashkov, A. S.; Shneider, M. M.; Popova, A. V.; Knirel, Y. A.; Hall, R. M. K17 capsular polysaccharide produced by *Acinetobacter baumannii* isolate G7 contains an amide of 2-acetamido-2-deoxy-D-galacturonic acid with D-alanine. *Int. J. Biol. Macromol.* 2020, *144*, 857– 862.

5. FUNCTIONAL CHARACTERIZATION OF A HAD PHOSPHATASE INVOLVED IN CAPSULAR POLYSACCHARIDE BIOSYNTHESIS IN *CAMPYLOBACTER JEJUNI* [*]

5.1. Introduction

The zoonotic pathogen *Campylobacter jejuni* is the leading cause of gastroenteritis worldwide (1). Infections caused by this Gram-negative bacterium are known as campylobacteriosis and have symptoms of nausea, fever, cramps, bloody diarrhea, and death (*2, 3*). Long term effects of campylobacteriosis includes reactive arthritis, Miller-Fischer Syndrome, and Guillain-Barré Syndrome (GBS) (4, 5). It is estimated that ~40% of all GBS cases are preceded by a *Campylobacter* infection (6). *C. jejuni* is a commensal organism in chickens, and can be found in the intestinal tracts of most farm animals such as cows, goats, and dogs. The most common route of human infection is through contact with raw or improperly cooked chicken, pet feces, and unpasteurized milk (7, 8). *Campylobacter* infections most commonly affect the elderly and very young children. It is estimated that 25,000 children die every year from campylobacteriosis worldwide (9). The disease is most prevalent in Asia and Africa. Efforts to combat the bacterium with antibiotics have been hampered due to the development of multi-drug resistance (10, 11, 12). Currently, there are no known FDA-

[*] Reprinted with permission from "Functional Characterization of a HAD Phosphatase Involved in Capsular Polysaccharide Biosynthesis in Campylobacter jejuni NCTC 11168" by Alexander S. Riegert, Tamari Narindoshvili, Nicole Platzer, and Frank M. Raushel, *Biochemistry* (Accepted 10/2022), Copyright 2022 American Chemical Society

approved vaccines to prevent *Campylobacter* infections. The best vaccine candidates to date have been glycoconjugate vaccines, which enable the immune system to identify surface exposed sugars on the bacterium (13).

Surrounding *C. jejuni* is a thick carbohydrate layer known as the capsular polysaccharide (CPS). In *C. jejuni* NCTC 11168 (serotype HS:2), the CPS is composed of a repeating unit of D-*glycero*-L-*gluco*-heptose, D-glucuronate, *N*-acetyl-D-galactosamine, and D-ribose (14, 15). These sugars can be further modified by *O*-methylation, phosphoramidylation, and amidation. The D-glucuronate moiety is modified as an amide with either serinol (**1**) or ethanolamine (**2**). The gene cluster for the biosynthesis of the capsular polysaccharide in *C. jejuni* NCTC 11168 (serotype HS:2) is presented in Figure 24 (14, 16).

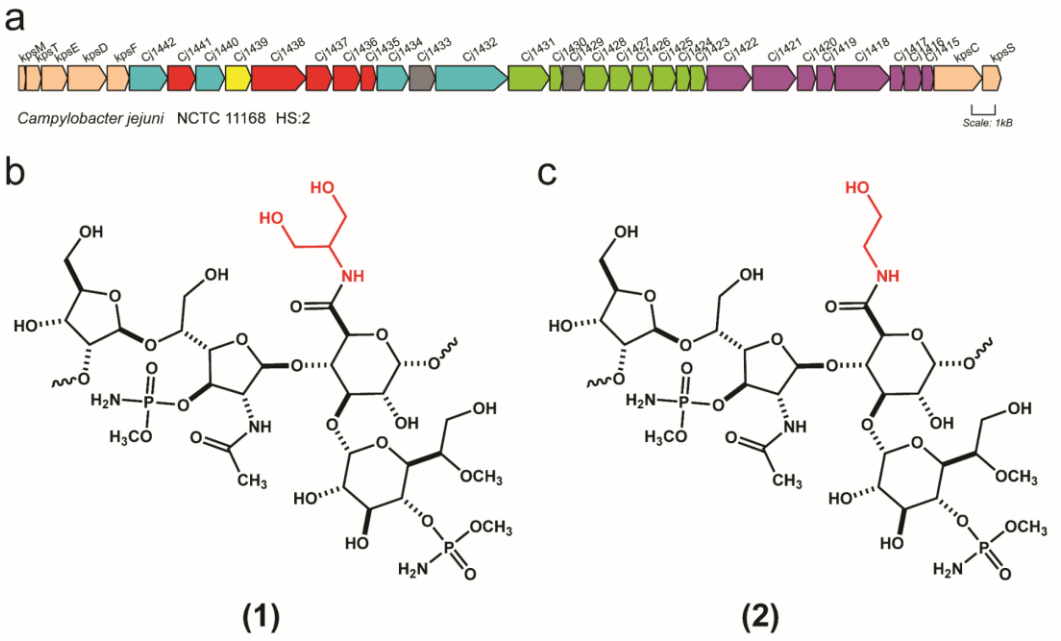


Figure 24 (a) Gene cluster for CPS formation in *C. jejuni* NCTC 11168 (serotype HS:2). Shown in red are genes responsible for D-glucuronamide formation. Genes shown in green are responsible for the biosynthesis of D-*glycero*-L-*gluco*-heptose. Genes displayed in purple are used to construct the phosphoramidate modification. Blue-colored genes are annotated as sugar transferases and grey-colored genes are hypothetical/uncharacterized. Cj1439, shown in yellow, is a pyranose/furanose mutase (17); wheat-colored genes are likely used for capsular polysaccharide export. (b) Serotype HS:2 capsular polysaccharide with the amide of serinol shown in red. (c) Serotype HS:2 capsular polysaccharide with the amide of ethanolamine shown in red.

We have previously shown that Cj1441 catalyzes the double NAD-dependent oxidation of UDP-D-glucose to UDP-D-glucuronate (18). Cj1436 is a PLP-dependent enzyme that catalyzes the decarboxylation of L-serine phosphate to ethanolamine phosphate (19), and Cj1437 is a PLP-dependent transaminase that catalyzes the conversion of dihydroxyacetone phosphate to *S*-serinol phosphate using L-glutamate as the amine donor (19). The C-terminal domain of Cj1438 is an ATP-dependent amidoligase (20). This enzyme was shown to use MgATP, D-glucuronate (or the methyl

glycoside of D-glucuronate) and either ethanolamine phosphate or *S*-serinol phosphate to synthesize the phosphorylated glucuronamide product (20). The unphosphorylated compounds are not substrates for the amidoligase enzyme (20). In this investigation, we interrogate the ability of the putative HAD phosphatase, Cj1435, to catalyze the ultimate step in the biosynthesis of the D-glucuronamide moiety in the CPS of *C. jejuni*.

Cj1435 is a member of the haloalkanoic acid dehalogenase (HAD) superfamily. This large and diverse superfamily is found in prokaryotes, eukaryotes, and archaea. There are roughly 18,000 enzymes in this superfamily and a vast majority of these enzymes catalyze the hydrolysis of phosphate esters, but some have been shown to be ATPases or phosphotransferases (21). Enzymes in this superfamily share a common catalytic scaffold that facilitates phosphate ester hydrolysis through two partial reactions (22). In the first step, the sidechain carboxylate of an active site aspartate residue attacks the phosphorus center of the substrate and forms an acyl-phosphate enzyme intermediate that is subsequently hydrolyzed by the activation of water from a second aspartate residue. Magnesium is most often used as a metal cofactor to help position the nucleophile and facilitate catalysis. Here, we demonstrate that Cj1435 is responsible for catalyzing the final step of glucuronamide biosynthesis in *C. jejuni* NCTC 11168 (serotype HS:2). We demonstrate that Cj1435 catalyzes the hydrolysis of phosphate from the glucuronamide of ethanolamine phosphate and of serinol phosphate.

5.2. Materials and Methods

5.2.1. Materials

Escherichia coli BL21 (DE3) was obtained from New England Biolabs. Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance III 400 MHz system equipped with a broadband probe and sample changer. Spectrophotometric data were collected on a SpectraMax$_{340}$ UV-visible plate reader. Unless otherwise noted, all compounds were purchased from commercially available sources. The chemical structures of the compound used for this investigation are presented in Figure 25. The chemical syntheses of compounds **7**, *R*/*S*-**8**, *R*/*S*-**9**, **10**, and **13** are presented in the Appendix Supplementary Information.

5.2.2. Cloning, Expression, and Purification of Cj1435

The gene encoding Cj1435 (Uniprot ID: Q0P8H9) from *C. jejuni* NCTC 11168 genomic DNA (ATCC 700819D-5) served as the starting template for amplification by PCR using Phusion High Fidelity polymerase (New England Biolabs). Primers were designed to incorporate the NdeI and XhoI restriction sites. The amplified DNA was ligated into a pET31b expression vector. This construct contained a C-terminal hexahistidine tag and the amino acid sequence of the purified enzyme is presented in Appendix Figure S16.

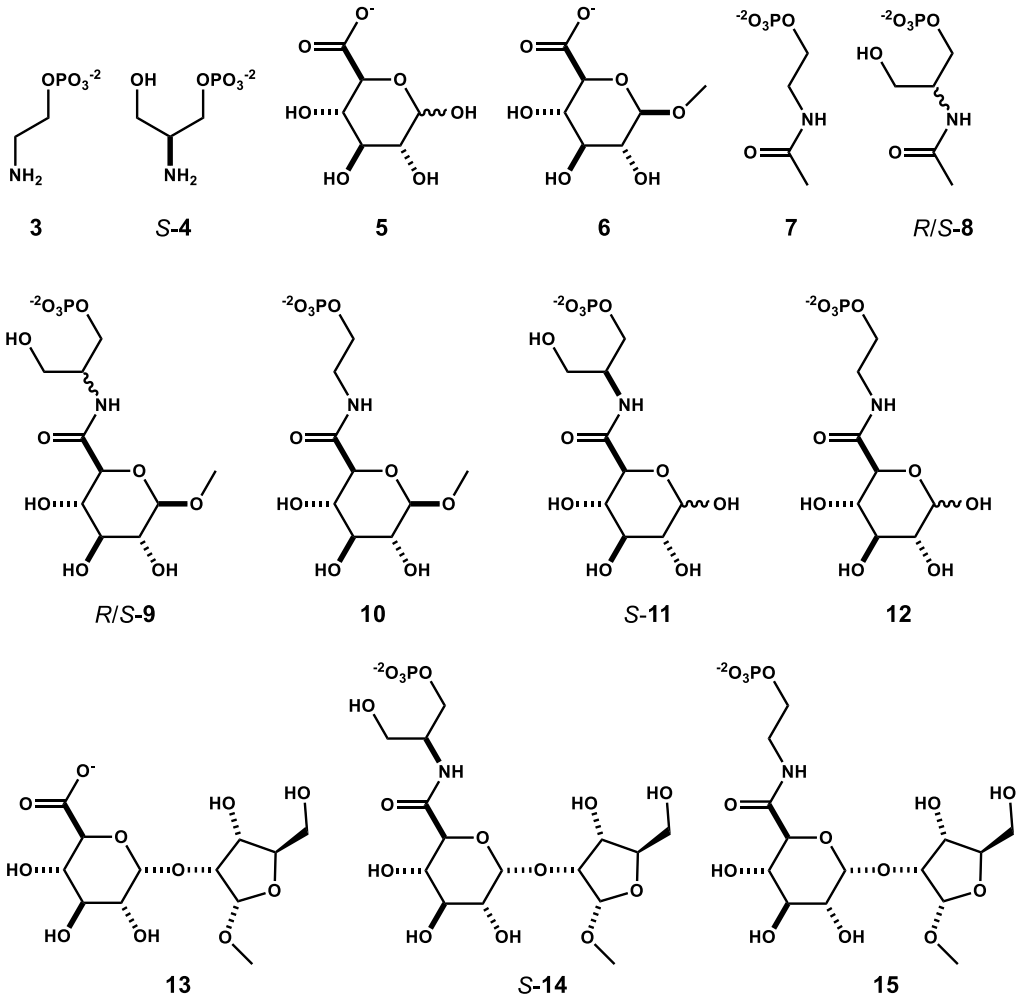


Figure 25 Structures of compounds used in this investigation.

The pET31b vector was used to transform BL21 (DE3) *E. coli* cells via electroporation. The cells harboring the pET31b-*Cj1435* plasmid were cultured in lysogeny broth with 100 mg/L ampicillin. The cells were grown at 37 °C until they reached an OD_{600} of ~0.8, whereupon the cells were induced with 1.0 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were allowed to grow at 21 °C for 18 h before centrifugation at 15,000 rcf at 4 °C. The cell pellet was resuspended in 50 mM HEPES/K^+ (pH 8.0), 300 mM KCl, 20 mM imidazole, and 1.0 mg DNAse from bovine

liver. The cells were placed in a beaker on ice and lysed by sonication. The lysate was clarified by centrifugation at 25,000 rcf at 4 °C for 30 min. The clarified lysate was passed through a 0.45 µm filter (Whatman) and then applied to a 5.0 mL prepacked HisTrap nickel affinity purification column (Cytivia). The protein was eluted using 50 mM HEPES/K^+, 300 mM KCl, 500 mM imidazole (pH 8.0) over a gradient of 25 column volumes. Fractions were pooled and dialyzed against 10 mM HEPES/K^+, 200 mM KCl, pH 8.0. The dialyzed protein was concentrated using 10 kDa MWCO filters (Cytivia) to a final concentration of 5.0 mg/mL before being flash-frozen in liquid nitrogen and stored at -80 °C.

5.2.3. Stereoselectivity of the Reaction Catalyzed by Cj1435

The stereoselectivity of Cj1435 was investigated with ^{31}P NMR spectroscopy using the chemically synthesized glucuronamide (10 mM) with racemic serinol phosphate (*R*/*S*-**9**), 3.0 µM Cj1435, 10 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. The reaction was incubated at 25 °C for 4 h before the ^{31}P NMR spectrum was obtained. The stereoselectivity of Cj1435 was also investigated by ^{31}P NMR spectroscopy using the product of the reaction catalyzed by Cj1438. Cj1438 catalyzes the formation of compound *S*-**11** using D-glucuronate, MgATP, and *S*-serinol phosphate (*S*-**4**). The reaction was initiated by the addition of 10 µM Cj1438, 10 mM *S*-serinol phosphate (*S*-**4**) with 10 mM ATP, 20 mM $MgCl_2$, 10 mM D-glucuronate (**5**), and 50 mM HEPES/K^+, pH 8.0 to make the chiral glucuronamide (*S*-**11**). The reaction mixture was incubated at 25 °C for 2 h and then passed through a Vivaspin 10 kDa MWCO filter (Cytivia) to

remove Cj1438. Cj1435 (3.0 μM) was added to the flowthrough and the reaction mixture was allowed to incubate for 30 min before the ^{31}P NMR spectrum was obtained.

5.2.4. Determination of Substrate Profile by ^{31}P NMR Spectroscopy

The substrate specificity of Cj1435 was interrogated by ^{31}P NMR spectroscopy. The reaction mixture contained 5.0 μM Cj1438, 5.0 mM D-glucuronate (**5**), 10 mM *S*-serinol phosphate (*S*-**4**), 10 mM ATP, 20 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. The reaction was allowed to incubate at 25 °C for 2 h before an NMR spectrum was obtained to ensure the complete formation of compound *S*-**11**. The amidoligase, Cj1438, was removed using a 10 kDa MWCO filter and the subsequent phosphatase reaction was initiated by the addition of 3.0 μM Cj1435 to hydrolyze *S*-**11**. The ^{31}P NMR spectrum was collected after an incubation period of 15 min at 25 °C. This procedure was repeated by substituting *S*-serinol phosphate with ethanolamine phosphate (**3**) to form compound **12**, which was then dephosphorylated after the addition of Cj1435.

The reaction catalyzed by Cj1435 was also investigated using ^{31}P NMR spectroscopy using the *N*-acetyl derivatives of ethanolamine phosphate (**7**) and racemic serinol phosphate (*R*/*S*-**8**) as potential substrates. The enzyme (3.0 μM) was incubated at 25 °C with 10 mM *N*-acetyl ethanolamine phosphate (**7**), 10 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0 for 4 h. Cj1435 (3.0 μM) was also incubated with 10 mM of racemic *N*-acetyl serinol phosphate (*R*/*S*-**8**), 10 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. Additionally, Cj1435 (3.0 μM) was incubated with either 10 mM ethanolamine phosphate (**3**) or 10 mM *S*-serinol phosphate (*S*-**4**), in the presence of 10 mM $MgCl_2$, and

50 mM HEPES/K^+, pH 8.0. ^{31}P NMR spectra were recorded after an incubation period of 4 h at 25 °C.

The reaction of Cj1435 was also tested with a phosphorylated disaccharide consisting of D-ribose and D-glucuronamide (*S*-**14**) and (**15**). To produce the phosphorylated amides attached to a disaccharide, Cj1438 (20 μM) was incubated with 10 mM D-glucuronate (1→2)-β-1-*O*-methyl ribose (**13**), 10 mM ATP, 2.5 mM *S*-serinol phosphate (*S*-**4**), 20 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. The enzyme was incubated at 25 °C for 30 min before the ^{31}P NMR spectrum was obtained showing the reaction had gone to completion with formation of compound *S*-**14**. This process was repeated with ethanolamine phosphate (**3**) to obtain compound **15**. Cj1438 was removed by filtration using a 10 kDa MWCO filter (Cytivia). To the flowthrough containing compound *S*-**14**, 3.0 μM Cj1435 was added and the reaction mixture allowed to incubate for 30 min at 25 °C before the ^{31}P NMR spectrum was obtained. In the same fashion, compound **15** was incubated with 3.0 μM Cj1435 and allowed to incubate for 30 min at 25 °C before the ^{31}P NMR spectrum was obtained.

5.2.5. Determination of Kinetic Constants

The rate of product formation was investigated using ^{31}P NMR at 25 °C. The following substrates were tested with Cj1435: D-glucuronamide of ethanolamine phosphate (**12**), glucuronamide of *S*-serinol phosphate (*S*-**11**), ethanolamine phosphate (**3**), *S*-serinol phosphate (*S*-**4**), *N*-acetyl ethanolamine phosphate (**7**), racemic *N*-acetyl *R*/*S*-serinol phosphate (*R*/*S*-**8**), 1-*O*-methyl β-glucuronamide of *R*/*S*-serinol phosphate

(*R*/*S*-**9**), 1-*O*-methyl β-glucuronamide of ethanolamine phosphate (**10**), and *O*-phospho-L-serine. Compounds *S*-**11** and **12** were synthesized enzymatically. The reaction containing 25 μM Cj1438, 10 mM D-glucuronate, 10 mM ATP, 20 mM $MgCl_2$, 5.0 mM ethanolamine phosphate (**3**), and 50 mM HEPES/K^+, pH 8.0 was allowed to incubate at 25 °C for 1 h before the Cj1438 enzyme was removed using a 10 kDa MWCO filter (Cytivia). The resulting glucuronamide of ethanolamine phosphate (**12**) was added to 50 nM Cj1435 and the ^{31}P NMR spectra were recorded over the first 25 min of the reaction. This procedure was repeated for compound *S*-**11**.

For compounds **3**, *S*-**4**, **7**, **10** and *O*-phospho-L-serine, the reactions were conducted using 5.0 mM substrate, 2.0 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. The reactions were initiated by the addition of enzyme and the ^{31}P NMR spectra were collected as a function of time. Since compounds *R*/*S*-**8**, and *R*/*S*-**9** are racemic mixtures, 10 mM of each substrate was used with 2.0 mM $MgCl_2$, and 50 mM HEPES/K^+, pH 8.0. For compounds **3**, *S*-**4** and *O*-phospho-L-serine, 20 μM of Cj1435 was added to the reaction mixtures. For compounds **7**, *R*/*S*-**8**, *R*/*S*-**9**, and **10**, 250 nM Cj1435 was added to the reaction mixtures.

5.2.6. Bioinformatic Analysis of Cj1435 and Amide Gene Cluster

A sequence similarity network (SSN) of Cj1435 was created using the amino acid sequence downloaded from Uniprot (Q0P8H9) as the input for a BLAST search of the Uniprot database using the EFI-EST webtool (23, 24). The e-value cutoff was set to 10^{-5} and the maximum number of returned sequences was set to 5000. The network was

generated using Cytoscape (version 3.8.2) and the clusters were created using the yFile organic layout with a cutoff value of 50% sequence identity. A genome neighborhood network (GNN) was created using the 3686 genes of the Cj1435 SSN as the input. Using the Pfam identifiers for UDP-glucose 6-dehydrogenase (PF00984), the PLP-dependent transaminase/decarboxylase (PF00155), the TupA-like amidoligase (PF14305), and the putative HAD-phosphatase (PF12710), a list of putative amide bond forming gene clusters was generated. A sequence alignment of HAD-phosphatases similar to Cj1435 was generated using Clustal Omega (25). The genes used for the alignment were identified as the phosphatase genes in the genome neighborhood network that likely participate in hydrolysis of a phosphate bond after amide bond formation.

5.2.7. Predicted Structure of Cj1435 using Alphafold2

The amino acid sequence of Cj1435 was used as the input for the three-dimensional structure prediction performed by Alphafold2 through use of the Colab fold webserver developed by Mirdita *et. al* (26). The predicted three-dimensional structure of Cj1435 was then used to create an alignment with the closest known structurally characterized protein, the phosphoserine phosphatase from *Methanocaldococcus jannaschii* (PDB id: 1L7P), in Pymol (27, 28).

5.3. Results

5.3.1. Bioinformatic Analysis of Cj1435

Cj1435 is 211 amino acids in length and is a member of the haloalkanoic dehalogenase (HAD) superfamily. The amino acid sequence for Cj1435 was obtained from Uniprot and used as the input for the EFI-EST BLAST retrieval option. The webtool retrieved 3686 sequences and the SSN was generated at a 50% sequence identity (Figure 26). The nodes colored blue are uncharacterized proteins. Shown as a large orange circle is a putative HAD-phosphatase from *Bordetella pertussis* Tohama I (PDB id: 3FVV). This enzyme shares 26% sequence identity across 89% of the query length. Cj1435 is shown as a large red circle. The yellow nodes are annotated as *Campylobacter* strains and tend to cluster together near Cj1435.

The phosphatases likely involved in the formation of the glucuronamide modification appear to cluster together away from the main clusters in the SSN (Figure 26). To investigate this further, we created a sequence alignment with all of the phosphatases identified in the genome neighborhood network shown in Appendix Figure S17. Strikingly, the catalytic motif found in the HAD-phosphatases involved in glucuronamide formation are different than what is usually found in the HAD-phosphatase superfamily. Traditionally, the catalytic motif of HAD-phosphatases is composed of a conserved DXD motif (29). In the sequence alignment of the phosphatases involved in glucuronamide phosphate hydrolysis, this motif is changed to DXCE/D. It is unclear why these proteins have a different catalytic motif from those previously investigated.

In an effort to more clearly understand the genes necessary for glucuronamide formation across various organisms, a genome neighborhood network (GNN) was generated using the 3686 genes identified in the SSN as the input (Appendix Figure S18). The genome neighborhood was filtered by gene clusters which contained Pfam identifiers for a UDP-glucose 6-dehydrogenase (PF00984), PLP-dependent transaminase/decarboxylase (PF00155), TupA-like amidoligase (PF14305), and the putative HAD-phosphatase (PF12710). A total of 35 gene clusters were identified to contain the genes necessary for the formation of a glucuronamide modification. All 35 gene clusters had enzymes homologous to Cj1435, Cj1438, and Cj1437/Cj1436. Only 19 of the 35 gene clusters contained an enzyme homologous to Cj1441, the UDP-glucose 6-dehydrogenase. The list of organisms contains bacteria such as *Capnocytophaga canimorsus*, *Campylobacter avium* LMG 24591, *Helicobacter fennelliae*, *Vibrio parahaemolyticus*, *Helicobacter bilis* WiWa, *Shigella boydii*, and *Photobacterium phosphoreum*, among others. A complete list of the Uniprot accession IDs and organism information is found in Appendix Table S1.

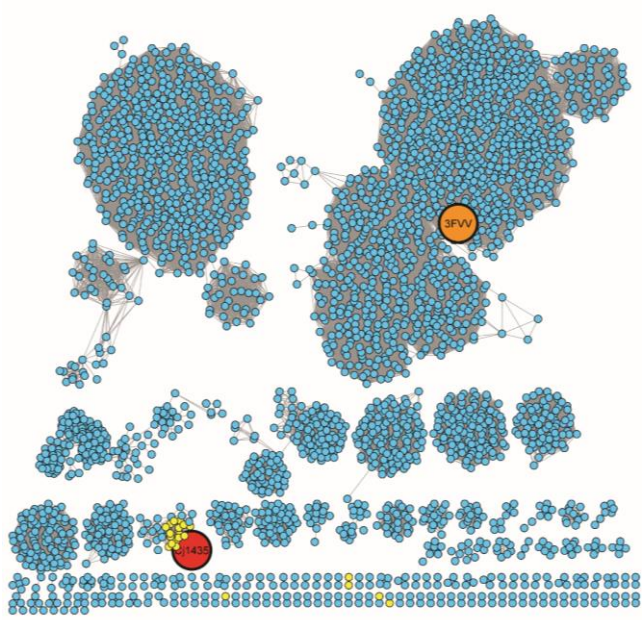


Figure 26 Sequence similarity network for Cj1435. Sequence similarity network of Cj1435 and 3686 sequences at 50% sequence identity. Cj1435 is displayed as a large red circle. The large orange circle is a protein of unknown function from *Bordetella pertussis* Tohama I (PDB id: 3FVV). Yellow circles are genes from *Campylobacter* strains. Blue circles are uncharacterized proteins.

5.3.2. Alphafold2 Structure Prediction of Cj1435

The predicted three-dimensional structure of Cj1435 was determined using Alphafold2 and is shown in Figure 27a (*26*). Cj1435 adopts a βαβ Rossmann fold with a type C1 capping domain as well as a C2 β-hairpin cap insertion (30). The predicted structure of Cj1435 was subsequently aligned with that of a phosphoserine phosphatase from *Methanocaldococcus jannaschii* (PDB id: 1L7P) shown in Figure 27b. An alignment of the two structures revealed an RMSD of 3.0 Å across 855 atoms. The two structures align well across the Rossmann fold and the active site residues, despite their low sequence identity of 23%. Where the structures differ is in their capping domains. Both contain a type C1 cap, which closes on top of the substrate upon binding (30). The

phosphoserine phosphatase from *M. jannaschii* was co-crystallized with *O*-phospho-L-serine and the cap is shown in the "closed" form. The Alphafold2 predicted structure depicts the cap region of Cj1435 in an "open" conformation. The capping region of the phosphoserine phosphatase from *M. jannaschii* is defined by two hinge points at residues N18 and P77. An alignment of the primary sequences of Cj1435 and the structurally characterized phosphoserine phosphatase is presented in Appendix Figure S19. This comparison, along with the structural alignment, was used to estimate the C1 capping domain of Cj1435 to extend from residue N17 to N83. In addition to the C1 cap found in Cj1435, it appears Cj1435 also contains a C2 β-hairpin cap insertion extending from residue E123 to Q148. A C2 β-hairpin cap insertion is also found in the phosphoserine phosphatase of *M. jannaschii* extending from Y117 to E141. These β-hairpin cap insertions are commonly seen in phosphoserine phosphatases as well as pyrimidine 5'-nucleotidase families (30). The C2 β-hairpin cap insertion is not believed to be the dominant capping domain of these proteins, but is instead used to sequester the active site away from the solvent by packing against the C1 cap (30).

The catalytic motif of Cj1435 is found at the end of the β1 strand like other HAD superfamily members, however it is composed of an unusual DXXE motif. Most HAD superfamily members have a DXD catalytic motif where the first aspartate residue is used to coordinate a Mg^{2+} cofactor while the second catalytic aspartate is responsible for activating a water molecule which attacks the acyl-phosphate intermediate formed with the first aspartate residue (31-33). It is unclear why the catalytic motif for Cj1435 is different from other characterized HAD-phosphatases.

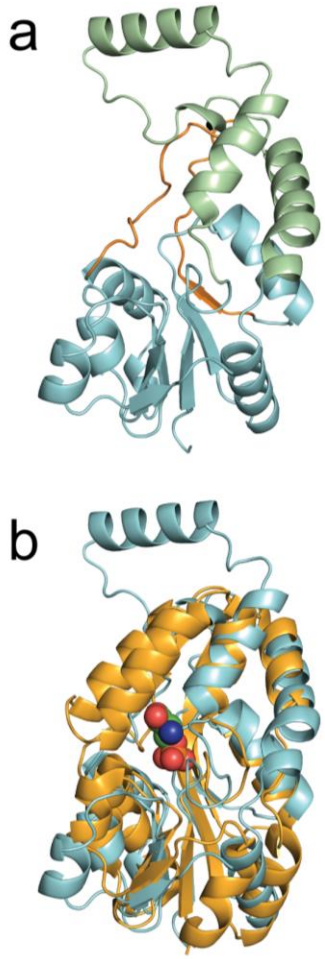


Figure 27 Structural prediction of Cj1435 using Alphafold2. (a) Alphafold2 predicted structure of Cj1435. The type C1 capping domain is colored green. The type C2 capping domain is shown in orange. The Rossmann fold is colored teal. (b) Alignment of the predicted structure of Cj1435 and the phosphoserine phosphatase from *M. jannaschii* (PDB id:1L7P). Shown in teal is the Alphafold2 prediction of the structure of Cj1435. Shown in gold is the X-ray crystal structure of the phosphoserine phosphatase from *M. jannaschii*. The *O*-phospho-L-serine bound in the active site of the phosphoserine phosphatase from *M. jannaschii* is shown as space filing spheres occupying the active site pocket.

5.3.3. Stereoselectivity for Phosphate Hydrolysis

We have previously demonstrated that the C-terminal domain of Cj1438 catalyzes the ATP-dependent amidation of D-glucuronate (**5**) and the *O*-methyl glycoside of D-glucuronate (**6**) with either *S*-serinol phosphate (*S*-**4**) or ethanolamine phosphate (**3**), but will not utilize UDP-D-glucuronate as a substrate (20). These results imply that the physiological substrate for the phosphatase, Cj1435, is the phosphorylated D-glucuronamide moiety after it has been added to the growing polysaccharide chain. To more firmly establish the substrate profile of Cj1435, the chemically synthesized amide (*R*/*S*-**9**) was incubated with the phosphatase Cj1435. The ^{31}P NMR spectrum of the initial substrate is presented in Figure 28a and the products are shown in Figure 28b. The phosphorylated D-glucuronamide (*R*/*S*-**9**) resonates at 3.93 ppm, while the inorganic phosphate product resonates at 2.30 ppm. Figure 28b demonstrates that only 50% of the diastereomeric substrate (either *S*-**9** or *R*-**9**) is capable of being dephosphorylated by Cj1435 and thus only one of the two diastereomers is a substrate for Cj1435.

To determine the specific stereoisomer that is hydrolyzed by Cj1435 we enzymatically synthesized the phosphorylated D-glucuronamide (*S*-**11**) using *R*/*S*-serinol phosphate (*R*/*S*-**4**), D-glucuronate (**5**), and MgATP using the amidoligase Cj1438. The ^{31}P NMR spectrum of the Cj1438-catalyzed products is presented in Figure 28c. The phosphorylated glucuronamide (*S*-**11**) has a chemical shift of 3.85 ppm and one equivalent of inorganic phosphate appears at 2.10 ppm. The unreacted *R*-serinol phosphate (*R*-**4**) resonates at 3.78 ppm. After incubation with Cj1435, all of the initial substrate (*S*-**11**) is now hydrolyzed and converted quantitatively to inorganic phosphate, while the *R*-serinol phosphate remains at 3.78 ppm (Figure 28d). This result clearly

demonstrates that Cj1435 is specific for hydrolysis of the amide made from *S*-serinol phosphate (*S*-**11**). These findings are consistent with the reaction observed when Cj1435 is incubated with *N*-acetyl *R*/*S*-serinol phosphate (*R*/*S*-**8**) shown in Appendix Figure S21.

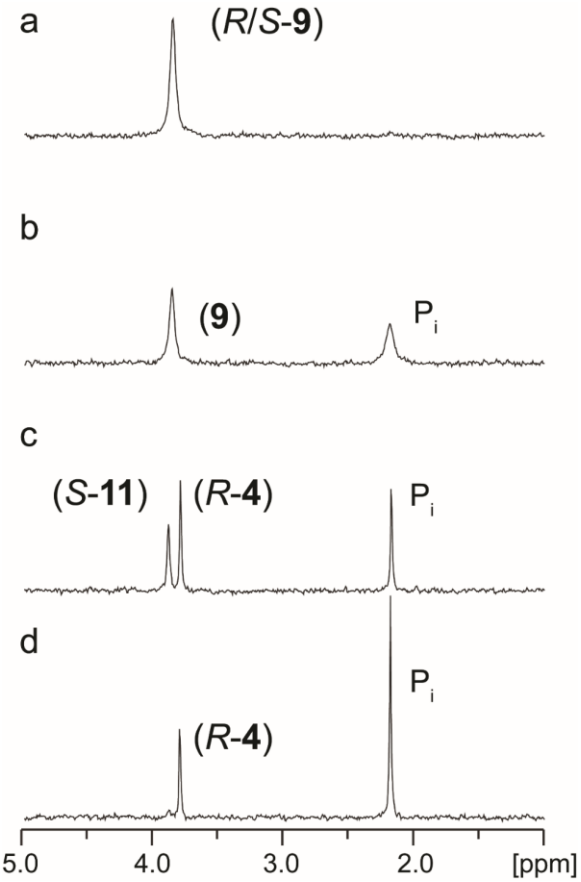


Figure 28 ^{31}P NMR spectra of β-1-*O*-methyl-D-glucuronamide with racemic serinol phosphate (*R*/*S*-**9**) and D-glucuronamide with *S*-serinol phosphate (*S*-**11**). (a) The ^{31}P NMR spectrum of the chemically synthesized 1-*O*-methyl-glucuronamide with *R*/*S*-serinol phosphate (*R*/*S*-**9**). The resonance for *R*/*S*-**9** is at 3.93 ppm. (b) The ^{31}P NMR spectrum of the reaction catalyzed by Cj1435 and (*R*/*S*-**9**). The unreacted substrate resonates at 3.93 ppm while the inorganic phosphate product appears at 2.30 ppm. (c) The ^{31}P NMR spectrum of the products of the reaction catalyzed by Cj1438 with *R*/*S*-serinol phosphate (*R*/*S*-**4**), MgATP, and D-glucuronate (**5**). The product (*S*-**11**) appears at 3.85 ppm, while the unreacted *R*-serinol phosphate (*R*-**4**) appears at 3.78 ppm. Inorganic phosphate is at 2.10 ppm. (d) The products of the reaction catalyzed by Cj1435 when incubated with the reaction mixture shown in **Figure 28c**. The resonance at 3.78 ppm is the unreacted *R*-serinol phosphate (*R*-**4**) while the phosphate resonance appears as a singlet at 2.10 ppm.

5.3.4. Glucuronamide Phosphate Hydrolysis

The HAD phosphatase Cj1435 was tested as a catalyst for the hydrolysis of phosphate esters formed from the amidoligase activity of Cj1438. The amidoligase Cj1438 was incubated with MgATP, D-glucuronate (**5**), and *S*-serinol phosphate (*S*-**4**) and the ^{31}P NMR spectrum of the products are shown in Figure 29a. The glucuronamide of *S*-serinol phosphate (*S*-**11**) appears as a singlet at 3.85 ppm while inorganic phosphate, formed from the activation of the carboxylate group with ATP appears as a singlet at 2.08 ppm. After addition of Cj1435, the resonance for the glucuronamide of *S*-serinol phosphate completely disappears and is replaced with the resonance for phosphate at 2.12 ppm (Figure 29b). To determine whether Cj1435 can use the glucuronamide of ethanolamine phosphate (compound **12**), a separate reaction was conducted. In this experiment, Cj1438 was incubated with D-glucuronate (**5**), MgATP, and ethanolamine phosphate (**3**). The glucuronamide of ethanolamine phosphate (**12**) appears as a singlet at 3.56 ppm, while inorganic phosphate resonates at 2.00 ppm (Figure 29c**)**. After Cj1435 is added, the resonance at 3.56 ppm disappears, while the inorganic phosphate resonance at 2.06 ppm has increased in size (Figure 29d).

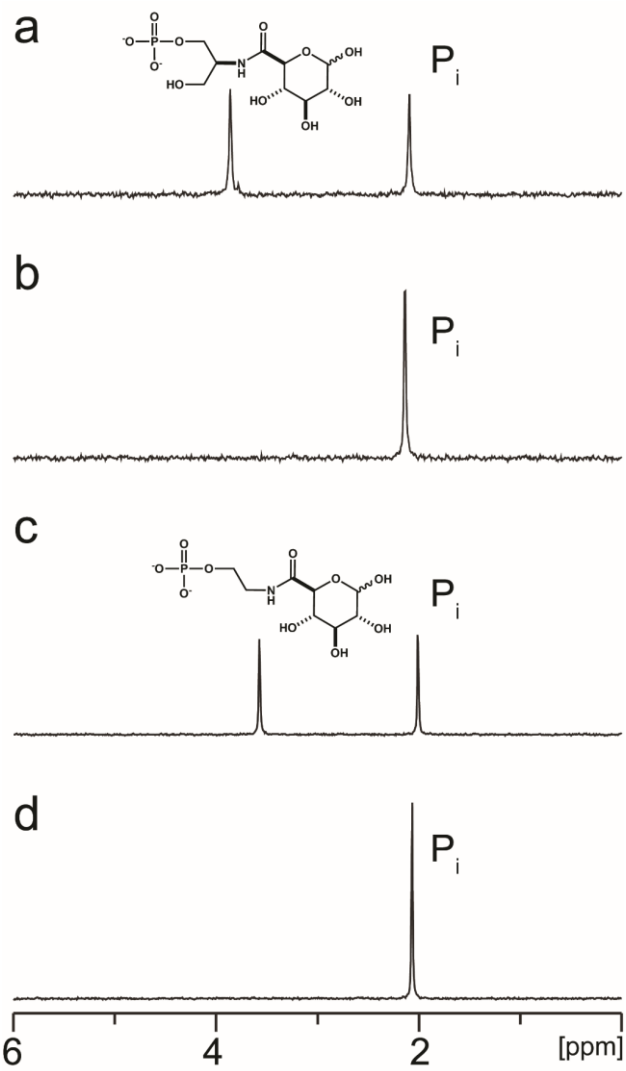


Figure 29 ^{31}P NMR spectra of the reaction products catalyzed by Cj1435. (a) The enzymatically prepared D-glucuronamide of *S*-serinol phosphate (*S*-**11**) appears as a singlet at 3.85 ppm and inorganic phosphate appears at 2.08 ppm. (b) The products of the reaction after the addition of Cj1435 to the enzymatically prepared *S*-**11**. The substrate was completely hydrolyzed and the phosphate appears as a singlet at 2.12 ppm. (c) The enzymatically prepared D-glucuronamide of ethanolamine phosphate (**12**) is present at 3.56 ppm and phosphate resonates at 2.00 ppm. (d) Products of the reaction after the hydrolysis of **12** by Cj1435. The substrate is completely converted to inorganic phosphate.

It is known from our previous investigation that Cj1438 will not accept UDP-D-glucuronate as a substrate for the formation of the phosphorylated D-glucuronamide

(20). This suggests that the glycosyltransferase for polysaccharide formation likely functions before the amidation reaction. Given that UDP-D-glucuronate is activated at C1, and the CPS structure of *C. jejuni* HS:2 contains a D-glucuronamide (1→2)-D-ribose, we set out to determine if Cj1438 and Cj1435 would use a disaccharide of D-glucuronate and 1-*O*-methyl-D-ribose as a substrate for their respective reactions. Cj1438 was incubated with the disaccharide of D-glucuronate (1→2)-1-*O*-methyl-D-ribose (**13**), MgATP, and ethanolamine phosphate (**3**) and formation of inorganic phosphate and (**15**) are shown in the ^{31}P NMR spectrum in Figure S20a. The products of the reaction of Cj1438 subsequently served as the substrate for Cj1435. After incubation with the HAD phosphatase, Cj1435, the resonance at 3.46 ppm disappeared and the inorganic phosphate resonance peak at 1.88 ppm increased (Figure S20b). The reaction was repeated with the substitution of *S*-serinol phosphate (*S*-**4**) for ethanolamine phosphate. Cj1438 catalyzed the synthesis of the phosphorylated glucuronamide disaccharide and the product (*S*-**14**) resonates at 3.65 ppm and inorganic phosphate at 1.68 ppm (Figure S20c). Upon addition of Cj1435, the resonance at 3.65 ppm completely disappeared and the inorganic phosphate peak at 1.73 ppm increased (Figure S20d).

5.3.5. Determination of Kinetic Constants for Cj1435

The rate constants for the reactions catalyzed by Cj1435 were determined using ^{31}P NMR spectroscopy with a single fixed concentration of substrate. The products of the reaction catalyzed by Cj1438 were used as substrates for the reaction catalyzed by

Cj1435. For the reaction of Cj1435 with the D-glucuronamide of ethanolamine phosphate (**12**) the rate constant at a substrate concentration of 5.0 mM was determined to be 22 ± 1.0 s^{-1}. With *S*-serinol phosphate glucuronamide (*S*-**11**) the rate constant for substrate hydrolysis was determined to be 37 ± 5.0 s^{-1}.

To obtain a better understanding of the substrate specificity of Cj1435, the enzyme was tested against the following primary amines: ethanolamine phosphate (**3**), *S*-serinol phosphate (*S*-**4**), and *O*-phospho-L-serine and the rate constants were determined to be 0.060 ± 0.010 s^{-1}, 0.16 ± 0.030 s^{-1}, and 0.037 ± 0.002 s^{-1} for the hydrolysis of **3**, *S*-**4**, and *O*-phospho-L-serine, respectively. To address the preference of Cj1435 for substrates which contain an amide bond, the enzyme was tested using *N*-acetyl ethanolamine phosphate (**7**) where the rate constant was determined to be 7.1 ± 1.3 s^{-1}. The enzyme was next tested with the chemically synthesized *N*-acetyl *R*/*S*-serinol phosphate (*R*/*S*-**8**) where the rate constant was found to be 16 ± 1.0 s^{-1}.

To determine the rate constant for the hydrolysis of the chemically synthesized 1-*O*-methyl β-glucuronamide of *R*/*S*-serinol phosphate (*R*/*S*-**9**) and 1-*O*-methyl β-glucuronamide of ethanolamine phosphate (**10**), the substrates were incubated with Cj1435. The rate constant for the 1-*O*-methyl β-glucuronamide of *R*/*S*-serinol phosphate (*R*/*S*-**9**) was determined to be 4.0 ± 1.0 s^{-1}. The enzyme was tested with 1-*O*-methyl β-glucuronamide of ethanolamine phosphate (**10**) where the rate constant was observed to be 7.1 ± 1.0 s^{-1}. The values are summarized in **Table 7**.

Table 7 Rate constant for the reactions catalyzed by Cj1435.

Substrate [5.0 mM]	Rate constant (s^{-1})
(12)	22 ± 1.0
(11)	37 ± 5.0
(3)	0.060 ± 0.010
(4)	0.16 ± 0.030
(7)	7.1 ± 1.3
(*R*/*S*-**8) ***	16 ± 1.0
(*R*/*S*-**9) ***	4.0 ± 1.0
(10)	7.1 ± 1.0

Reactions conducted at 25 °C, pH 8.0
*10 mM substrate used

5.4. Discussion

The biosynthesis of the D-glucuronamide modification present in the CPS of *C. jejuni* is not completely understood. Previously, we have demonstrated that the enzymes Cj1441, Cj1436, Cj1437, and the C-terminal domain of Cj1438 from *C. jejuni* NCTC 11168 (serotype HS:2) are necessary for formation of the D-glucuronamides of *S*-serinol phosphate and ethanolamine phosphate (18-20). The ultimate enzyme in the overall biosynthetic pathway is apparently Cj1435, which has been previously annotated as a HAD phosphatase. Here we demonstrated that Cj1435 is responsible for the hydrolysis of the phosphate group from the D-glucuronamide of either ethanolamine phosphate or *S*-serinol phosphate.

The catalytic substrate profile for the hydrolysis of various substrate analogs shows a marked preference of the enzyme toward substrates that contain an amide bond

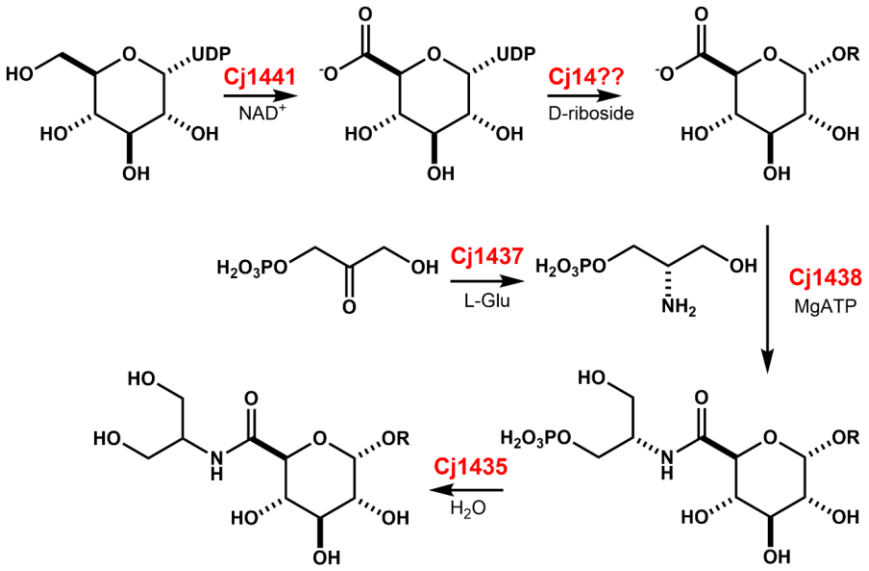
instead of a primary amine. For example, the rate of hydrolysis of compounds *S*-**11** and **12** are more than two orders of magnitude faster than the hydrolysis of compounds **3** and *S*-**4**. This observation is fully consistent with the fact that the enzyme responsible for amide bond formation does not utilize ethanolamine or serinol as substrates (20). The amide bond forming enzyme Cj1438 does not utilize UDP-D-glucuronate as a substrate, but can utilize D-glucuronate (**5**), the methyl glycoside of D-glucuronate (**6**), and the disaccharide with 1-*O*-methyl-D-ribose (**13**) (20). These results demonstrate quite clearly that amide bond formation follows polysaccharide formation and that phosphate hydrolysis is the ultimate step in formation of the glucuronamide moiety in the HS:2 serotype of *C. jejuni*.

The glucuronamide modification is present in the capsular polysaccharides of *C. jejuni* serotype HS:2 and HS:19 (14, 34). Based on previous bioinformatic analysis conducted in our laboratory, we also identified serotypes HS:22 and HS:38 as likely containing a glucuronamide moiety (35). While these four serotypes contain a D-glucuronamide moiety within their CPS, it should be noted that HS:2 is the only serotype which has a D-glucuronamide moiety formed with ethanolamine. The other serotypes lack an enzyme homologous to Cj1436, the PLP-dependent L-serine phosphate decarboxylase, which is necessary to create the ethanolamine phosphate required to form the D-glucuronamide of ethanolamine phosphate.

The biosynthetic pathway begins with the dual oxidation of UDP-D-glucose to UDP-D-glucuronate catalyzed by Cj1441 (18). This is followed by the reaction of UDP-D-glucuronate with the ribosyl end of the growing capsular polysaccharide catalyzed by

an unidentified glycosyltransferase to form the D-glucuronate (1→2)-D-ribosyl polysaccharide. Possible candidates for the glycosyltransferase are Cj1432, Cj1434, Cj1438, Cj1440, and Cj1442. The N-terminal domain of Cj1432 is an intriguing candidate because it contains an apparent GT4 (retaining) glycosyltransferase contained within residues 1 through 356. The other possible candidates are all GT2 inverting enzymes, which would form a glycosidic bond of the opposite stereochemistry. Investigations are ongoing to determine which enzyme is responsible for the glycosyltransferase reaction. The phosphorylated primary amines necessary for the biosynthesis of the glucuronamide are synthesized using Cj1437, a PLP-dependent transaminase, which uses DHAP and L-glutamate and produces *S*-serinol phosphate (19). The other primary amine is formed by Cj1436, a PLP-dependent L-serine phosphate decarboxylase, which catalyzes the formation of ethanolamine phosphate (19). These phosphorylated primary amines are then used as substrates along with MgATP and the glucuronate (1→2)-D-ribosyl end of the growing polysaccharide chain to form the phosphorylated D-glucuronamide. This molecule is dephosphorylated by Cj1435 to give the final product of the pathway. The reaction scheme is summarized in Scheme 20.

Scheme 20 Proposed reaction pathway for the formation of the D-glucuronamide found within the capsular polysaccharide of *C. jejuni* NCTC 11168 (serotype HS:2)



To our knowledge, this work represents the first time a glucuronamide modification of a capsular polysaccharide has been functionally characterized. Based on known capsular polysaccharides, this modification is also present in *E. coli* L19, *Acinetobacter baumannii* G7, *Shigella boydii* type 8, *Vibrio cholerea* H11 (non-O1), *Proteus mirabilis* O27, and *Vibrio vulnificus* strain 6353 (36-40). We propose naming the biosynthetic pathway for making this modification to the CPS of *C. jejuni*, the Capsule Amide Ligase pathway. Enzymes in this pathway will adopt the following naming convention: Cj1438, the glucuronate amide ligase, will be named CalA; Cj1437, the DHAP transaminase, is to be named CalB; Cj1436, the L-serine phosphate decarboxylase, will be CalC; Cj1435, the glucuronamide phosphate phosphatase, will be

called CalD. Cj1441, the UDP-glucose 6-dehydrogenase, is not unique to this pathway and thus will retain its designator, KfiD.

5.5. CONCLUSIONS

Campylobacter jejuni is a pathogenic organism that is surrounded by a capsular polysaccharide. This CPS is modified with a D-glucuronoamide moiety formed from either ethanolamine or serinol. We have now shown that a phosphorylated glucuronamide moiety is formed by the combined catalytic activities of Cj1441, Cj1436, Cj1437, Cj1438 (and possibly Cj1432). We demonstrate that the final enzyme is the pathway is Cj1435 which acts to hydrolyze the phosphate ester bond. These studies were used to formulate a biosynthetic pathway for the formation of the D-glucuronamide moiety contained within the capsular polysaccharide of *C. jejuni*.

5.6. References

1. Johnson, T. J., Shank, J. M., and Johnson, J. G. (2017). Current and Potential Treatments for Reducing Campylobacter Colonization in Animal Hosts and Disease in Humans. *Frontiers in microbiology, 8,* 487.
2. Allos B. M. (2001). Campylobacter jejuni Infections: update on emerging issues and trends. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *32*(8), 1201–1206.
3. Ruiz-Palacios G. M. (2007). The health burden of Campylobacter infection and the impact of antimicrobial resistance: playing chicken. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *44*(5), 701–703.
4. Pope, J. E., Krizova, A., Garg, A. X., Thiessen-Philbrook, H., and Ouimet, J. M. (2007). Campylobacter reactive arthritis: a systematic review. *Seminars in arthritis and rheumatism*, *37*(1), 48–55.
5. Nachamkin, I., Allos, B. M., and Ho, T. (1998). Campylobacter species and Guillain-Barré syndrome. *Clinical microbiology reviews*, *11*(3), 555–567.
6. Yuki N. (1997). Molecular mimicry between gangliosides and lipopolysaccharides of Campylobacter jejuni isolated from patients with Guillain-Barré syndrome and Miller Fisher syndrome. *The Journal of infectious diseases*, *176 Suppl 2,* S150–S153.
7. Kaakoush, N. O., Castaño-Rodríguez, N., Mitchell, H. M., and Man, S. M. (2015). Global Epidemiology of Campylobacter Infection. *Clinical microbiology reviews*, *28*(3), 687–720.

8. Jaakkonen, A., Kivistö, R., Aarnio, M., Kalekivi, J., and Hakkinen, M. (2020). Persistent contamination of raw milk by Campylobacter jejuni ST-883. *PloS one*, *15*(4), e0231810.
9. Poly, F., Noll, A. J., Riddle, M. S., and Porter, C. K. (2019). Update on Campylobacter vaccine development. *Human vaccines & immunotherapeutics*, *15*(6), 1389–1400.
10. Luangtongkum, T., Jeon, B., Han, J., Plummer, P., Logue, C. M., and Zhang, Q. (2009). Antibiotic resistance in Campylobacter: emergence, transmission and persistence. *Future microbiology*, *4*(2), 189–200.
11. Gupta, A., Nelson, J. M., Barrett, T. J., Tauxe, R. V., Rossiter, S. P., Friedman, C. R., Joyce, K. W., Smith, K. E., Jones, T. F., Hawkins, M. A., Shiferaw, B., Beebe, J. L., Vugia, D. J., Rabatsky-Ehr, T., Benson, J. A., Root, T. P., Angulo, F. J., and NARMS Working Group (2004). Antimicrobial resistance among Campylobacter strains, United States, 1997-2001. *Emerging infectious diseases*, *10*(6), 1102–1109.
12. Engberg, J., Aarestrup, F. M., Taylor, D. E., Gerner-Smidt, P., and Nachamkin, I. (2001). Quinolone and macrolide resistance in Campylobacter jejuni and C. coli: resistance mechanisms and trends in human isolates. *Emerging infectious diseases*, *7*(1), 24–34.
13. Riddle, M. S., and Guerry, P. (2016). Status of vaccine research and development for Campylobacter jejuni. *Vaccine*, *34*(26), 2903–2906.
14. 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. *European journal of biochemistry*, *269*(21), 5119–5136.

15. Young, K. T., Davis, L. M., and Dirita, V. J. (2007). Campylobacter jejuni: molecular biology and pathogenesis. *Nature reviews. Microbiology*, *5*(9), 665–679.
16. Harrison, K. J., Crécy-Lagard, V., and Zallot, R. (2018). Gene Graphics: a genomic neighborhood data visualization web application. *Bioinformatics (Oxford, England)*, *34*(8), 1406–1408.
17. 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. *The Journal of biological chemistry*, *285*(1), 493–501.
18. Riegert, A. S., and Raushel, F. M. (2021). Functional and structural characterization of the UDP-glucose dehydrogenase involved in capsular polysaccharide Biosynthesis from Campylobacter jejuni. *Biochemistry*, *60*(9), 725–734.
19. Riegert, A. S., Narindoshvili, T., Coricello, A., Richards, N., and Raushel, F. M. (2021). Functional characterization of two PLP-dependent enzymes involved in capsular polysaccharide biosynthesis from Campylobacter jejuni. *Biochemistry*, *60*(37), 2836–2843.
20. Riegert, A. S., Narindoshvili, T., and Raushel, F. M. (2022). Discovery and functional characterization of a clandestine ATP-dependent amidoligase in the

biosynthesis of the capsular polysaccharide from Campylobacter jejuni. *Biochemistry*, *61*(2), 117–124.

21. Allen, K. N., and Dunaway-Mariano, D. (2009). Markers of fitness in a successful enzyme superfamily. *Current opinion in structural biology*, *19*(6), 658–665.
22. Lu, Z., Dunaway-Mariano, D., and Allen, K. N. (2008). The catalytic scaffold of the haloalkanoic acid dehalogenase enzyme superfamily acts as a mold for the trigonal bipyramidal transition state. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 5687–5692.
23. Zallot, R., Oberg, N., and Gerlt, J. A. (2019). The EFI web resource for genomic enzymology tools: Leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry*, *58*(41), 4169–4182.
24. Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009). Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PloS one*, *4*, e4345.
25. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A., Potter, S. C., Finn, R. D., and Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research*, *47*(W1), W636–W641.
26. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682.

27. Wang, W., Cho, H. S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H. H., Grigoriev, I. V., Wemmer, D. E., and Kim, S. H. (2002). Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states. *Journal of Molecular Biology*, *319*, 421–431.

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

29. Collet, J. F., Stroobant, V., and Van Schaftingen, E. (1999). Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases. *The Journal of Biological Chemistry*, *274*, 33985–33990.

30. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006). Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *Journal of Molecular Biology*, *361*, 1003–1034.

31. Morais, M. C., Zhang, W., Baker, A. S., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2000). The crystal structure of Bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. *Biochemistry*, *39*, 10385–10396.

32. Baker, A. S., Ciocci, M. J., Metcalf, W. W., Kim, J., Babbitt, P. C., Wanner, B. L., Martin, B. M., and Dunaway-Mariano, D. (1998). Insights into the mechanism of catalysis by the P-C bond-cleaving enzyme phosphonoacetaldehyde hydrolase

derived from gene sequence analysis and mutagenesis. *Biochemistry*, *37*, 9305–9315.

33. Collet, J. F., Gerin, I., Rider, M. H., Veiga-da-Cunha, M., and Van Schaftingen, E. (1997). Human L-3-phosphoserine phosphatase: sequence, expression and evidence for a phosphoenzyme intermediate. *FEBS letters*, *408*, 281–284.
34. McNally, D. J., Jarrell, H. C., Khieu, N. H., Li, J., Vinogradov, E., Whitfield, D. M., Szymanski, C. M., and Brisson, J. R. (2006). The HS:19 serostrain of Campylobacter jejuni has a hyaluronic acid-type capsular polysaccharide with a nonstoichiometric sorbose branch and O-methyl phosphoramidate group. *The FEBS journal*, *273*, 3975–3989.
35. Poly, F., Serichantalergs, O., Kuroiwa, J., Pootong, P., Mason, C., Guerry, P., and Parker, C. T. (2015). Updated Campylobacter jejuni capsule PCR multiplex typing system and its application to clinical isolates from south and southeast Asia. *PloS one*, *10*, e0144349.
36. Landersjö, C., Weintraub, A., and Widmalm, G. (1996). Structure determination of the O-antigen polysaccharide from the enteroinvasive Escherichia coli (EIEC) O143 by component analysis and NMR spectroscopy. *Carbohydrate Research*, *291*, 209–216.
37. Kenyon, J. J., Senchenkova, S., Shashkov, A. S., Shneider, M. M., Popova, A. V., Knirel, Y. A., and Hall, R. M. (2020). K17 capsular polysaccharide produced by Acinetobacter baumannii isolate G7 contains an amide of 2-acetamido-2-deoxy-D-

galacturonic acid with D-alanine. *International Journal of Biological Macromolecules*, *144*, 857–862.

38. Vinogradov, E. V., Holst, O., Thomas-Oates, J. E., Broady, K. W., and Brade, H. (1992). The structure of the O-antigenic polysaccharide from lipopolysaccharide of Vibrio cholerae strain H11 (non-O1). *European journal of biochemistry*, *210*, 491–498.
39. Vinogradov, E. V., Krajewska-Pietrasik, D., Kaca, W., Shashkov, A. S., Knirel, Y. A., and Kochetkov, N. K. (1989). Structure of Proteus mirabilis O27 O-specific polysaccharide containing amino acids and phosphoethanolamine. *European Journal of Biochemistry*, *185*, 645–650.
40. Reddy, G. P., Hayat, U., Xu, Q., Reddy, K. V., Wang, Y., Chiu, K. W., Morris, J. G., Jr, and Bush, C. A. (1998). Structure determination of the capsular polysaccharide from Vibrio vulnificus strain 6353. *European journal of biochemistry*, *255*, 279–288.

6. CONCLUSIONS

6.1. Biosynthesis Capsular Polysaccharide Glucuronamide Modification

The enzymes necessary for the biosynthesis of the capsular polysaccharide glucuronamide modification in *Campylobacter jejuni* NCTC 11168 (HS:2) have been determined. This represents the first time the enzymes responsible for this modification have been functionally characterized. There are five enzymes required for the biosynthesis of the glucuronamide modification. The first enzyme in the pathway is Cj1441, a UDP-glucose dehydrogenase which catalyzes the dual oxidation of UDP-glucose to UDP-glucuronic acid. The X-ray crystal structure of this enzyme has been determined at 2.09 Å. Bound in the active site of this enzyme was NAD(H) and UDP-glucose. The phosphorylated primary amines necessary for biosynthesis of the glucuronamide modification are made by Cj1436 and Cj1437. Cj1436 catalyzes the PLP-dependent decarboxylation of *O*-phospho-L-serine to ethanolamine phosphate. Cj1437 is responsible for the catalysis of dihydroxyacetone phosphate to (*S*)-serinol phosphate using PLP and L-glutamate. To create the amide bond, Cj1438 uses D-glucuronate, MgATP, and either ethanolamine phosphate or (*S*)-serinol phosphate. This enzyme is stereospecific for the (*S*)- enantiomer of serinol phosphate. Additionally, this enzyme will not catalyze the reaction with UDP-glucuronate implying that a glycosyltransferase reaction occurs before the amide bond is formed. Finally, Cj1435 will hydrolyze the phosphodiester bond found on the glucuronamide of ethanolamine phosphate or the glucuronamide of (*S*)-serinol phosphate.

The genes necessary for this modification are found in 3 other known *Campylobacter jejuni* serotypes: HS:19, HS:22, and HS:38. However, none of these other serotypes contain a homolog to Cj1436. This suggests that none of the capsular polysaccharides contain a glucuronamide with ethanolamine attached. The glucuronamide modification is found in numerous other bacteria including *E. coli* L19, *A. baumanii* G7, *S. Boydii* type 8, *V. cholerae* H11 (non-O1), *Proteus mirabilis* O27, *V. vulnificus* strain 6353. It is unclear what the function of this modification is at this time.

6.2. Future work

The future work for this project will need to focus on how the capsular polysaccharide is assembled. This will require the functional characterization of the glycosyltransferases required for glycosidic bond formation between the sugars in the polysaccharide. Preliminary findings suggest that Cj1432 is responsible for the catalysis of glycosidic bond formation between the C1 of UDP-glucuronate and C2 of ribose. This enzyme contains a GT4 glycosyltransferase fold which will retain the stereochemistry of the anomeric carbon of the glucuronate. This is the only retaining glycosyltransferase found in the gene cluster. A challenge to determining the functions of the glycosyltransferases necessary for CPS elongation will be their substrate specificities. The donor substrates will need to be an activated sugar (for example: UDP-glucuronate), however it is unclear whether the acceptor substrates for the enzymes will be monosaccharides, disaccharides, trisaccharides or even polysaccharides. Further complicating matters is the fact that many of these glycosyltransferases have regions that

appear to be associated with a membrane. Special care will need to be taken to either purify these proteins with detergents, or truncate them to remove the membrane associated regions.

Another area that needs attention, is the elucidation of the initiator sugar that attaches to the poly-Kdo linker on the surface of the bacterium. Currently, it is unknown which glycosyltransferase or glycosyltransferases are responsible for this transformation. Furthermore, it is also unknown what terminates the polysaccharide elongation. In order to answer these questions, a different serotype than HS:2 should be used. HS:2 is one of the largest and most complicated capsular polysaccharides in *C. jejuni*. A simpler capsule structure with less genes in the gene cluster would help reduce the number of substrates that are needed to be tested. Some possible serotypes to use instead of HS:2 would be HS:19 and HS:15. Both of these capsules are composed of a repeating disaccharide.

The ability to characterize genes that create modifications on the capsular polysaccharide of *C. jejuni* is limited by the number of known capsular polysaccharide structures. There are currently only 12 fully characterized CPS structures. In order to more fully understand the genes necessary for capsular polysaccharide production in pathogenic bacteria, more CPS structures are needed. A collaboration with one of the groups that determines the CPS structures may prove fruitful.

Finally, the enzymes in the glucuronamide modification pathway should be structurally characterized. A collaboration is ongoing with the Holden lab at UW-

Madison to structurally characterize these enzymes. Determining the structures of these enzymes will help elucidate the necessary amino acids for catalysis and may provide a foundation for inhibitor design.

APPENDIX A

SUPPLEMENTARY INFORMATION

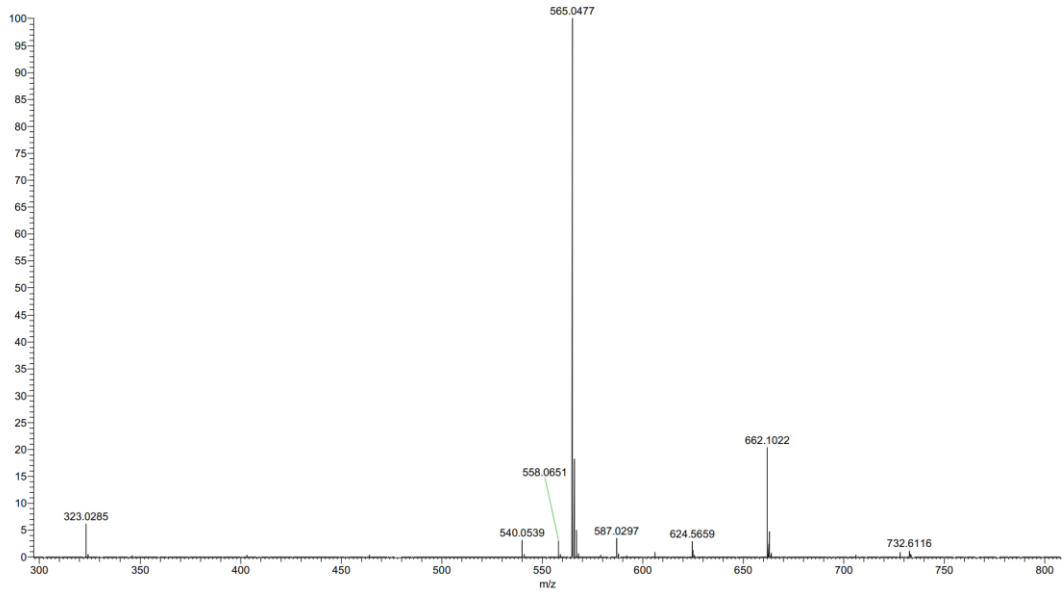


Figure S1 Mass spectrum of the no enzyme control for the Cj1441 catalyzed reaction. UDP-glucose is identified at an m/z of 565.0 and NAD^+ is identified at an m/z of 662.1

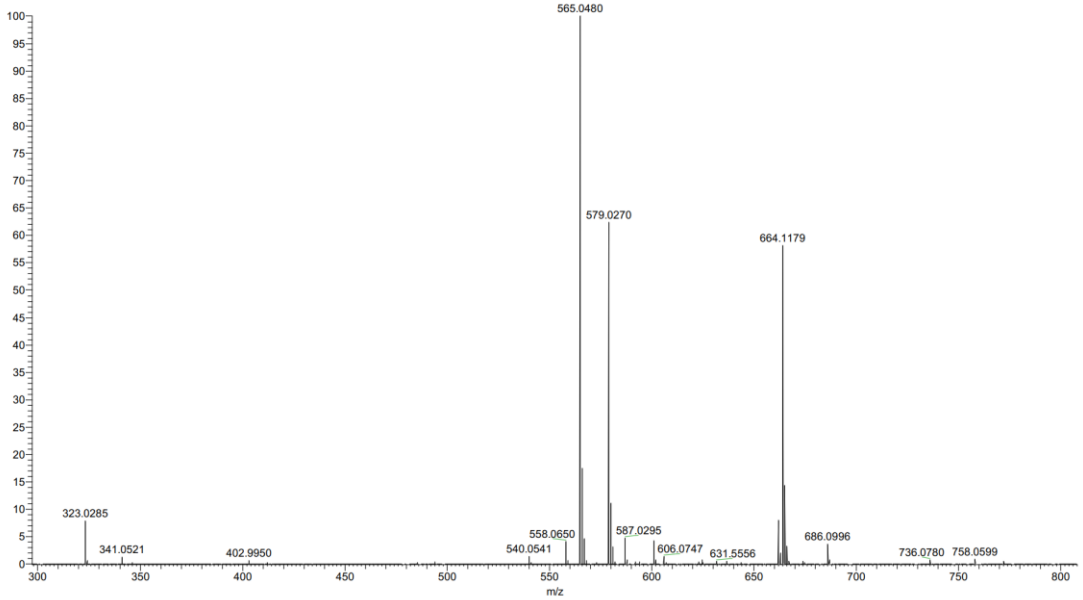


Figure S2 Mass spectrum of the reaction catalyzed by Cj1441 in the presence of UDP-glucose and NAD^+. UDP-glucose is identified at an m/z of 565.05, UDP-glucuronic acid is identified at an m/z of 579.03, and NADH is identified at an m/z of 664.12.

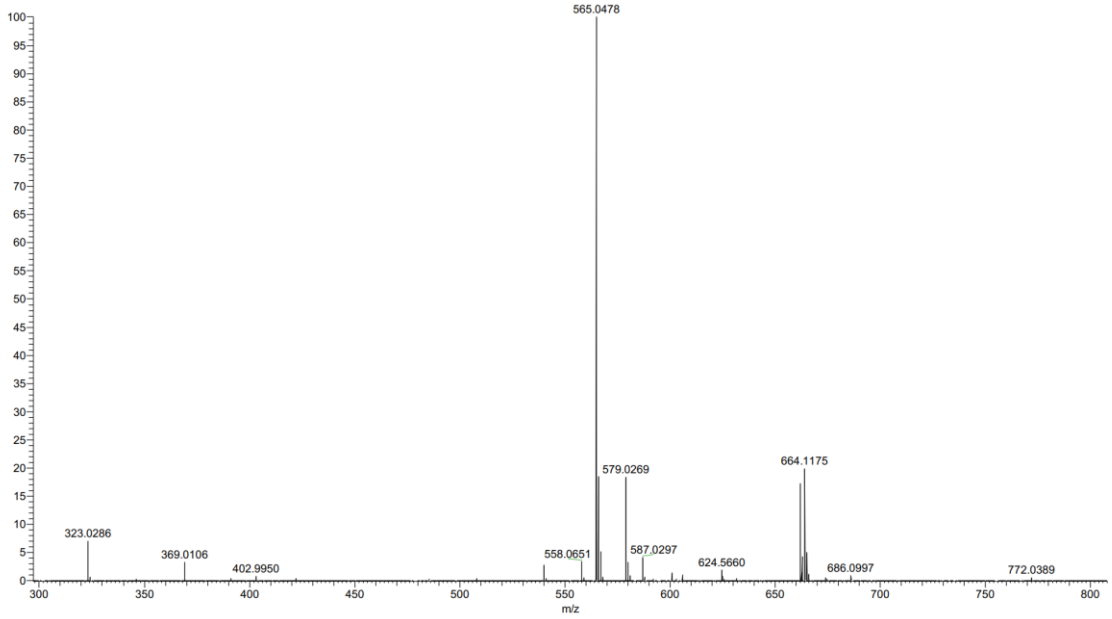


Figure S3 Mass spectrum for the reaction catalyzed by Cj1441 in the presence of serinol phosphate. UDP-glucose is identified at an m/z of 565.0, UDP-glucuronic acid is identified at an m/z of 579.0, and NADH is identified at an m/z of 664.1. There was no support for the formation of an amide with serinol phosphate (expected m/z of 729.0).

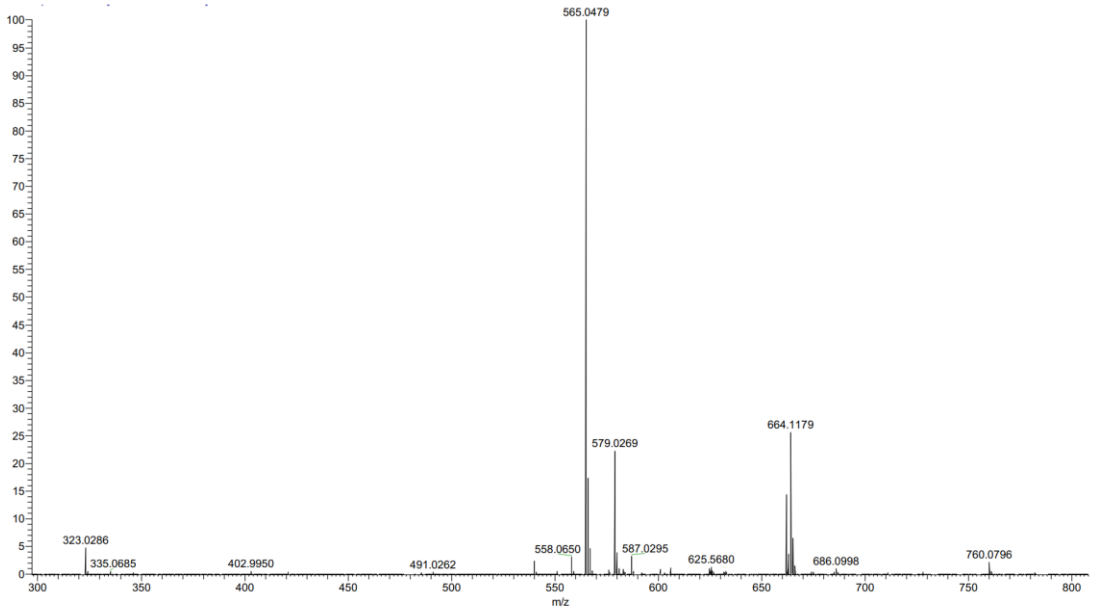


Figure S4 Mass spectrum of the reaction catalyzed by Cj1441 in the presence of added serinol. UDP-glucose is identified at 565.0, UDP-glucuronic acid is identified at 579.0, and NADH is identified at 664.1. There was no support for the formation of an amide with serinol at an m/z of 651.0.

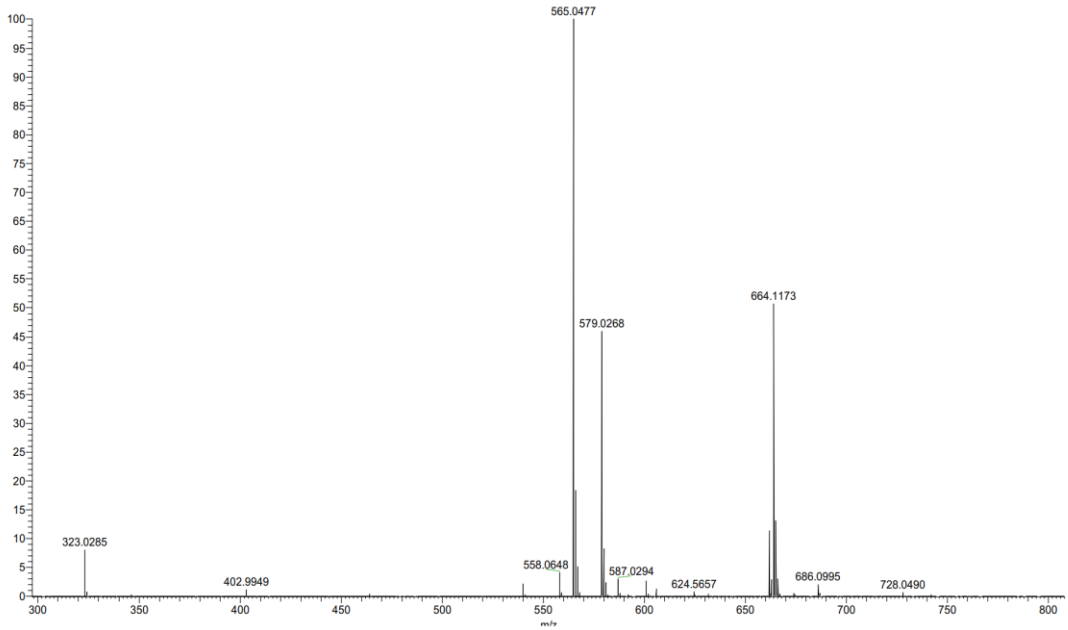


Figure S5 Mass spectrum of the reaction catalyzed by Cj1441 in the presence of ethanolamine phosphate. UDP-glucose is identified at 565.0, UDP-glucuronic acid is identified at 579.0, and NADH is identified at 664.1. There was no support for the formation of an amide with ethanolamine phosphate at an m/z of 699.0.

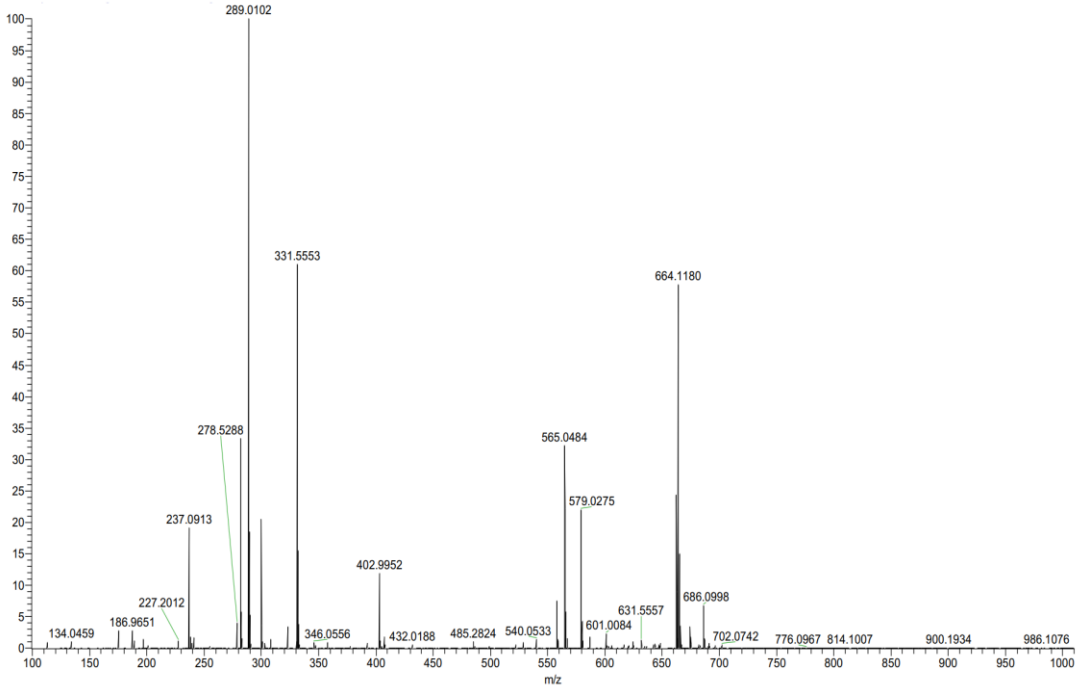


Figure S6 Mass spectrum of the reaction catalyzed by Cj1441 in the presence of ethanolamine. UDP-glucose is identified at 565.0, UDP-glucuronic acid is identified at 579.0, and NADH is identified at 664.1. There was no support for the formation of an amide with ethanolamine at an m/z of 621.0.

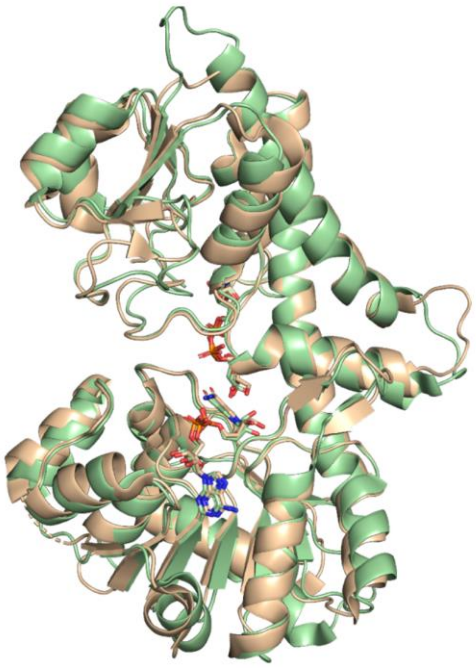


Figure S7 Structural alignment of Cj1441 shown in wheat with the UDP-glucose 6-dehydrogenase from *S. pyogenes* (PDB id: 1DLJ) shown in green. The two proteins align with an RMSD of 0.75 Å.

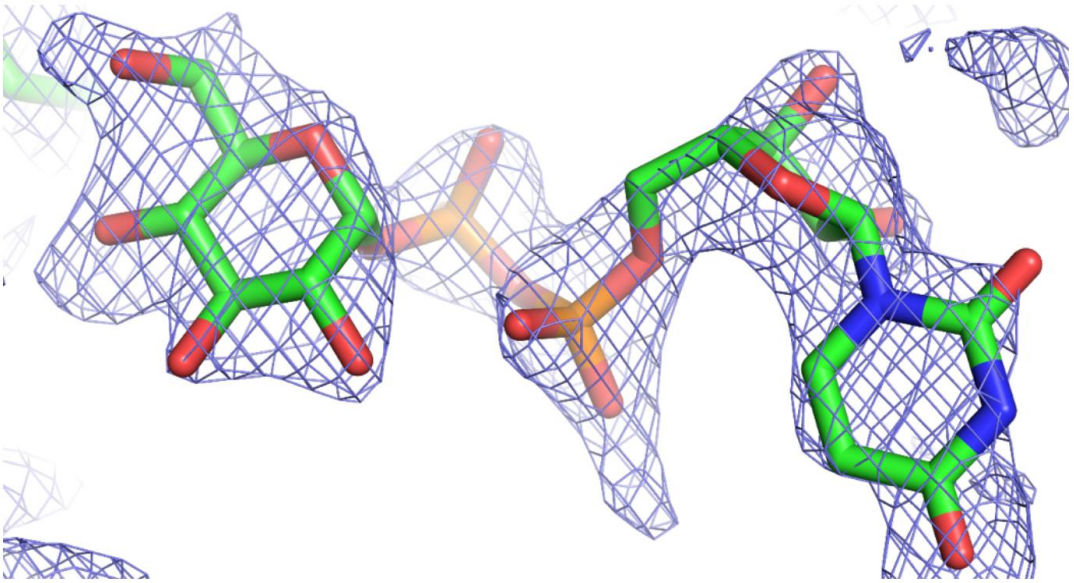


Figure S8 The density shown is from an omit map contoured at 2σ for UDP-glucose. The map was calculated with coefficients of the form (F_o - F_c), where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude.

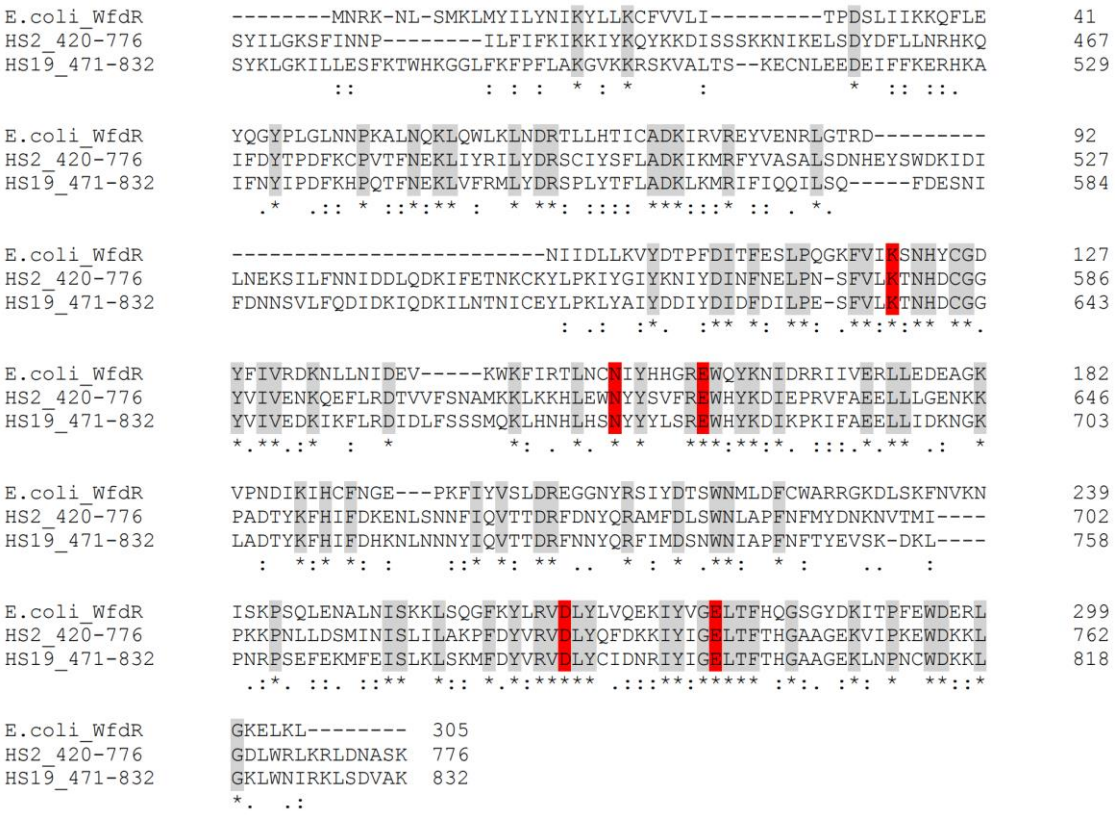


Figure S9 Protein sequence alignment of the C-terminal domain (residues 420-776) of Cj1438 (Uniprot id: Q0P8H6) with WfdR (Uniprot id: B5L389) from *E. coli* O143 and the C-terminal domain (residues 471-832) of HS:19.11 (Uniprot id: Q5M6M2). Residues highlighted in grey are conserved among all three proteins. Residues highlighted in red are likely responsible for ATP binding. These assignments are based on structurally characterized ATP-grasp enzymes. (1)

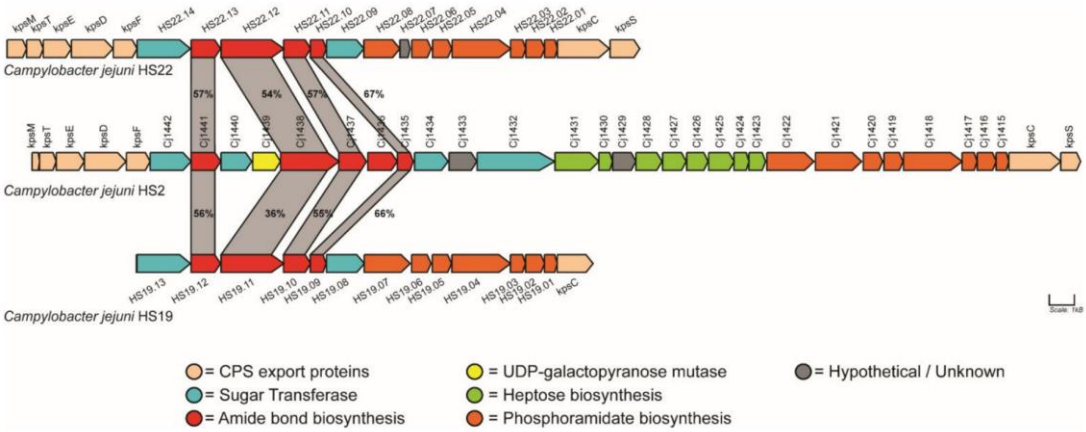


Figure S10 Gene cluster alignment of *Campylobcter jejuni* HS:2, *Campylobacter jejuni* HS:19, and *Campylobacter jejuni* HS:22. Genes highlighted in red are proposed to be responsible for the formation of the amide bond containing molecules on the capsular polysaccharide. Sequence identities for the genes proposed to be involved in amide bond biosynthesis are highlighted in grey boxes, with the sequence identity in bold. (2-5)

Figure S11 Pair-wise sequence alignment of Cj1436, SeCobD (GenBank: AAC79515.1), and SMUL_1544 (NCBI Reference Sequence: WP_025344678.1). Residues highlighted in red are involved in coordinating the phosphate group of the substrate, L-serine phosphate. Lysine 242, highlighted in teal, is likely involved in formation of the internal aldimine with PLP. (6, 7)

Chemical Synthesis of (*R*/*S*)-serinol-P.

A solution of *N*-Cbz protected serinol (**1**) (0.45 g, 2.0 mmol, 1.0 eqv.) and imidazole (0.27 g, 4.0 mmol, 2.0 equiv.) in 15 mL of anhydrous dichloromethane was chilled to 0 ^{o}C and dibenzyl chlorophosphonate (0.63 g, 2.2 mmol, 1.1 equiv.) was added. The reaction mixture was allowed to warm to ambient temperature (23 ^{o}C) and stirred for 16 h. After filtration and concentration, the residue was purified by silica gel column chromatography (hexanes/ethyl acetate, 1:3) to afford **2** (0.21 g, 0.43 mmol) as a colorless oil. (22% yield).

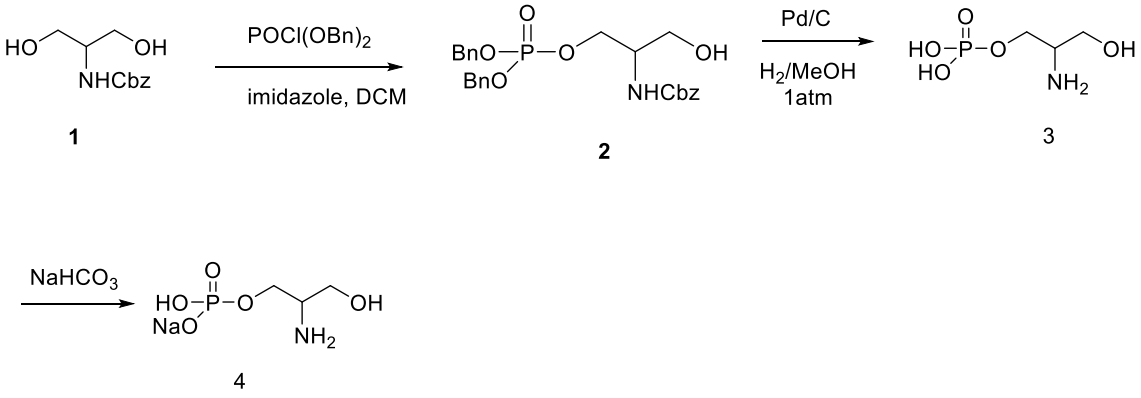
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.13 (br. s, 1H), 7.41 – 7.28 (m, 15H), 5.28-5.18 (m, 1H), 5.10 (s, 2H), 5.08-5.00 (m, 4H), 4.07-3.93 (m, 1H), 3.91-3.82 (m, 1H), 3.73 (dd, J_1 = 11.4 Hz, J_2 = 3.6 Hz, 1 H), 3.55 (dd, J_1 = 11.4 Hz, J_2 = 5.4 Hz, 1 H), ^{31}P NMR (160 MHz, $CDCl_3$) δ 1.11 (s). HRMS (ESI^-) m/z $[M - H]^-$ calc. for $C_{25}H_{27}NO_7P$: 484.1525, found: 484.1521.

To the solution of **2** (0.20 g, 0.41 mmol) in 10 mL of methanol, Pd/C (10%) (0.10 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtration and concentration 0.068 g (98%) of (*R*/*S*)-serinol phosphate (**3**) was obtained.

^{1}H NMR (400 MHz, D_2O) δ 4.03 - 3.94 (m, 1H), 3.94 - 3.86 (m, 1H), 3.79 (dd, J_1 = 12.4 Hz, J_2 = 4.8 Hz, 1 H), 3.71 (dd, J_1 = 12.4 Hz, J_2 = 6.6 Hz, 1 H), 3.55- 3.46 (m, 1H). ^{31}P NMR (160 MHz, D_2O) δ 2.97 (s). HRMS (ESI^-) m/z $[M - H]^-$ calc. for C_3NO_5P: 170.0218, found: 170.0214.

A solution of **3** in 2 mL of water was neutralized by $NaHCO_3$ and evaporated to dryness to afford 75 mg (97%) of the sodium salt of racemic serinol-phosphate (**4**). The synthetic transformation is summarized in **Scheme S1**.

Scheme S1. Synthesis of racemic serinol phosphate (**4**).



Chemical Synthesis of (*S*)-serinol-P.

N-Cbz protected serinol (**1**) was converted into enantiomerically pure (*R*)-serinol mono acetyl product **5** using an enzymatic desymmetrization procedure described previously (8).

A solution of (*R*)-2-(((benzyloxy)carbonyl)amino)-3-hydroxypropyl acetate **5** (0.80 g, 3.0 mmol, 1.0 equiv.) and imidazole (0.41 g, 6.0 mmol, 2.0 equiv.) in 20 mL of anhydrous dichloromethane, was chilled to 0 ^{o}C and dibenzyl chlorophosphonate (0.95 g, 3.3 mmol, 1.1 equiv.) was added. The reaction was allowed to warm to ambient temperature (23 ^{o}C) and stirred for 16 h. After filtration and concentration, the residue

was purified by silica gel column chromatography (hexanes/ethyl acetate, 1:1) to afford **6** (0.80 g, 1.52 mmol) as a colorless oil. (51% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38 – 7.32 (m, 15H), 5.18 – 5.08 (m, 3H), 5.07 – 5.01 (m, 4H), 4.13 – 3.99 (m, 5H), 2.01 (s, 3H). ^{31}P NMR (160 MHz, $CDCl_3$) δ – 0.62 (s). HRMS (ESI^-) m/z $[M - H]^-$ calcd. for $C_{27}H_{29}NO_8P$: 526.1631, found: 526.1635.

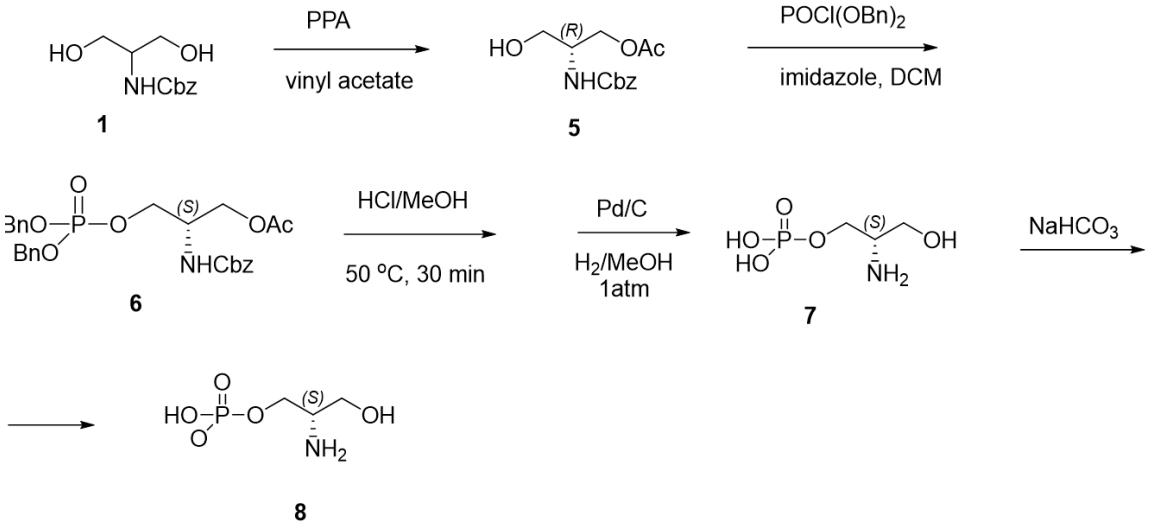
(*S*)-2-amino-3-hydroxypropyl dihydrogen phosphate (7).

To a solution of **6** (0.30 g, 0.57 mmol) in 5 mL of MeOH, 5 mL of conc. HCl (37%) was added and the reaction mixture was heated at 50 ^{o}C for 30 min. After concentration to dryness, the residue was dissolved in 15 mL of MeOH, Pd/C (10%) (0.145 g) was added and the reaction was subjected to H_2 (1 atm) for 16 h. After filtration and solvent evaporation, product **7** (0.089 g, 0.52 mmol, 91 %) was obtained.

^{1}H NMR (400 MHz, MeOH-d_4) δ 4.27- 4.10(m, 2H), 3.82 (dd, J_1 = 11.6 Hz, J_2 = 4.6 Hz, 1 H), 3.73 (dd, J_1 = 11.6 Hz, J_2 = 6.2 Hz, 1 H), 3.54 - 3.46 (m, 1H). ^{31}P NMR (160 MHz, MeOH-d_4) δ – 0.33 (s). HRMS (ESI^-) m/z $[M - H]^-$ calc. for C_3NO_5P: 170.0218, found: 170.0214.

Solution of **7** in 3 mL of water was neutralized by $NaHCO_3$ and evaporated to dryness to afford 100 mg (98%) of **8**. The reaction scheme is summarized in **Scheme S2**.

Scheme S2: Reaction scheme for synthesis of (*S*)-serinol phosphate.



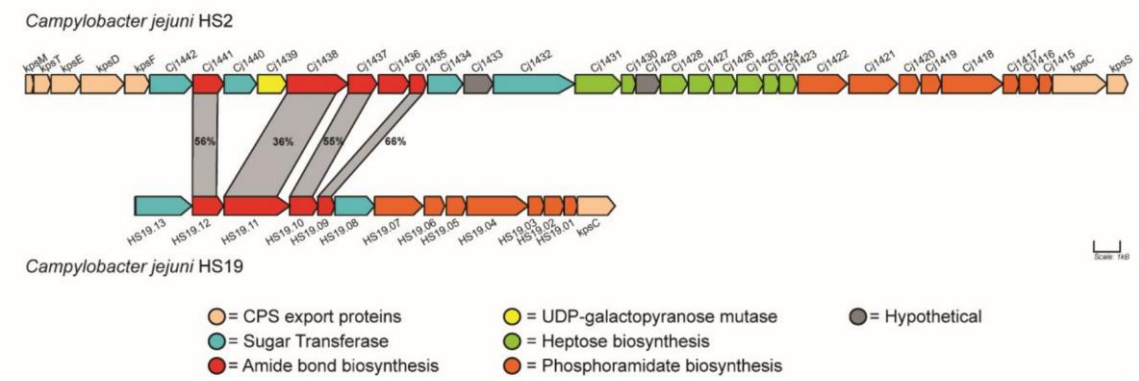


Figure S12 Gene cluster alignment of *C. jejuni* from HS:2 serotype and *C. jejuni* from HS:19 serotype. Highlighted in red are genes proposed to be responsible for the formation of the amide bond containing molecules on the capsular polysaccharide. The sequence identity between the genes proposed to be involved in amide bond biosynthesis is shown in bold. (2, 3, 5)

(a)

MLEEYSDYEQALKTKKHLSYILGKSFINNPILFIFKIKKIYKQYKKDISSSKKNIKELSDYDFL
LNRHKQIFDYTPDFKCPVTFNEKLIYRILYDRSCIYSFLADKIKMRFYVASALSDNHEYSWDKI
DILNEKSILFNNIDDLQDKIFETNKCKYLPKIYGIYKNIYDINFNELPNSFVLKTNHDCGGYVI
VENKQEFLRDTVVFSNAMKKLKKHLEWNYYSVFREWHYKDIEPRVFAEELLLGENKKPADTYKF
HIFDKENLSNNFIQVTTDRFDNYQRAMFDLSWNLAPFNFMYDNKNVTMIPKKPNLLDSMINISL
ILAKPFDYVRVDLYQFDKKIYIGELTFTHGAAGEKVIPKEWDKKLGDLWRLKRLDNASKLEHHH
HHH

(b)

MKELSDYDFLLNRHKQIFDYTPDFKCPVTFNEKLIYRILYDRSCIYSFLADKIKMRFYVASALS
DNHEYSWDKIDILNEKSILFNNIDDLQDKIFETNKCKYLPKIYGIYKNIYDINFNELPNSFVLK
TNHDCGGYVIVENKQEFLRDTVVFSNAMKKLKKHLEWNYYSVFREWHYKDIEPRVFAEELLLGE
NKKPADTYKFHIFDKENLSNNFIQVTTDRFDNYQRAMFDLSWNLAPFNFMYDNKNVTMIPKKPN
LLDSMINISLILAKPFDYVRVDLYQFDKKIYIGELTFTHGAAGEKVIPKEWDKKLGDLWRLKRL
DNASKLEHHHHHH

Figure S13 (a) Amino acid sequence for Cj1438(399-776). The methionine at the beginning is shown in red font and the added polyhistidine tag at the C-terminus is also shown in red font. (b) Amino acid sequence for Cj1438(453-776). The methionine at the N-terminus is shown in red font and the added polyhistidine tag at the C-terminal end is also shown in red font.

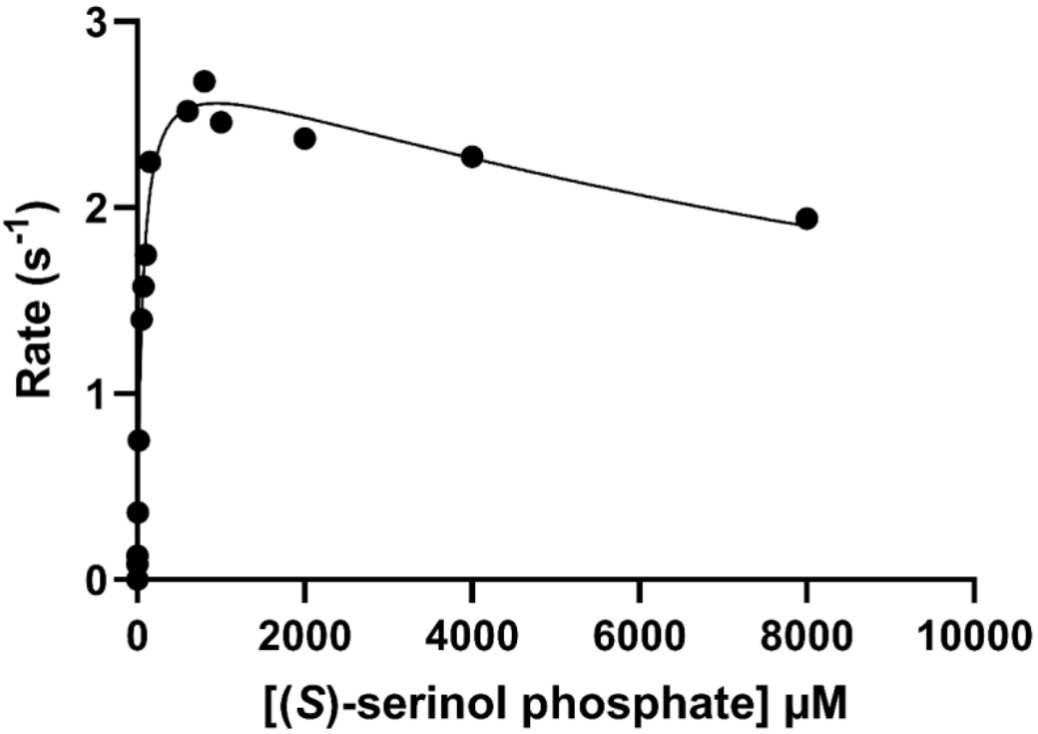


Figure S14 Variation of (*S*)-Serinol Phosphate. Substrate inhibition curve of the reaction of 50 nM Cj1438(453-776), variable concentrations of (*S*)-serinol phosphate (0.0025-8.0 mM), 3.0 mM ATP, 6.0 mM MgCl2, 2.0 mM PEP, 10 mM D-glucuronate, 250 µM NADH, 5.0 units lactate dehydrogenase, 5.0 units pyruvate kinase, and 50 mM HEPES/K+ pH 8.0. Inhibition constant (K_i) for the reaction is 16 mM.

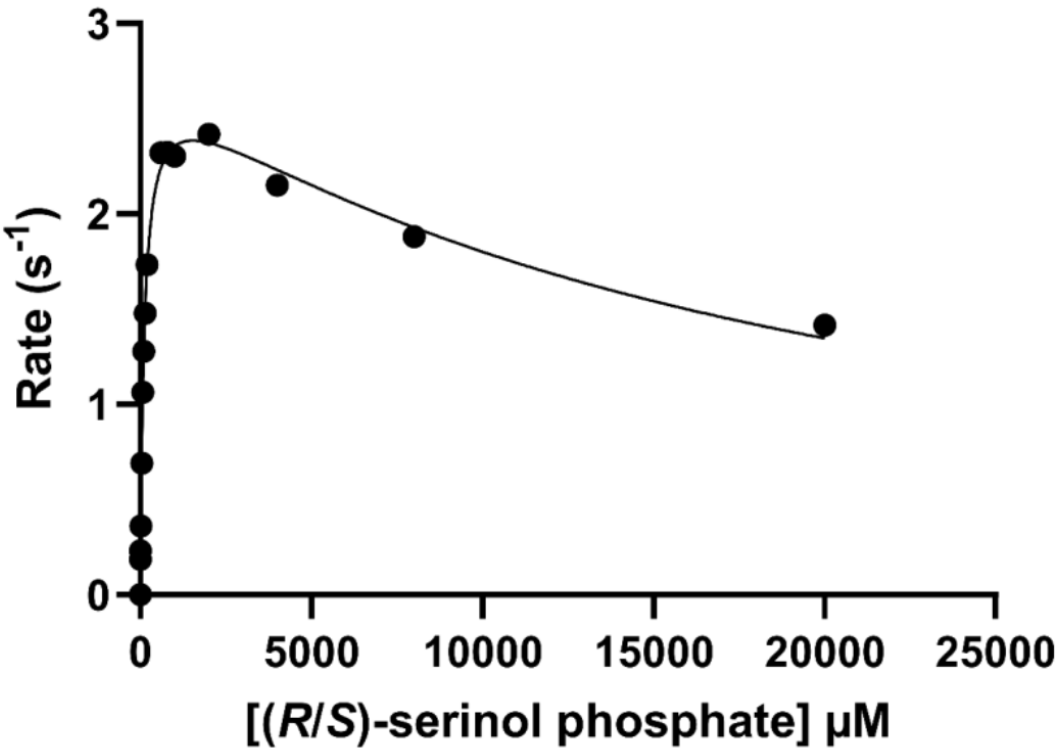


Figure S15 Variation of (*R*/*S*)-Serinol Phosphate. Substrate inhibition curve of the reaction of 50 nM Cj1438, variable (*R*/*S*)-serinol phosphate (0.0025-20 mM), 3.0 mM ATP, 6.0 mM MgCl2, 2.0 mM PEP, 10 mM glucuronate, 250 µM NADH, 5.0 units lactate dehydrogenase, 5.0 units pyruvate kinase, and 50 mM HEPES/K+ pH 8.0. Inhibition constant (K_i) for the reaction is 19 mM.

MSKEILALFDFCETLTNFQTLDRYLPLAGSKNINYTQSKNLARRERFQRENLPYPRYEW
LIDLDVDLAEEIAQEFVYTDVMANLNQNVMDRLFWHQDEGHTIVIVSGGLTIYIKEFAR
IYNIENIVAVDLEIYKNKLTGNIDGIHTMQERKLYKLAQKFNLKQFDLKNSYAYSDCVS
DIPLLSLVGNPNVIECGKDLQWARILGFNILLKY

Figure S16 Amino acid sequence of Cj1435 with a C-terminal hexahistidine tag shown in red font.

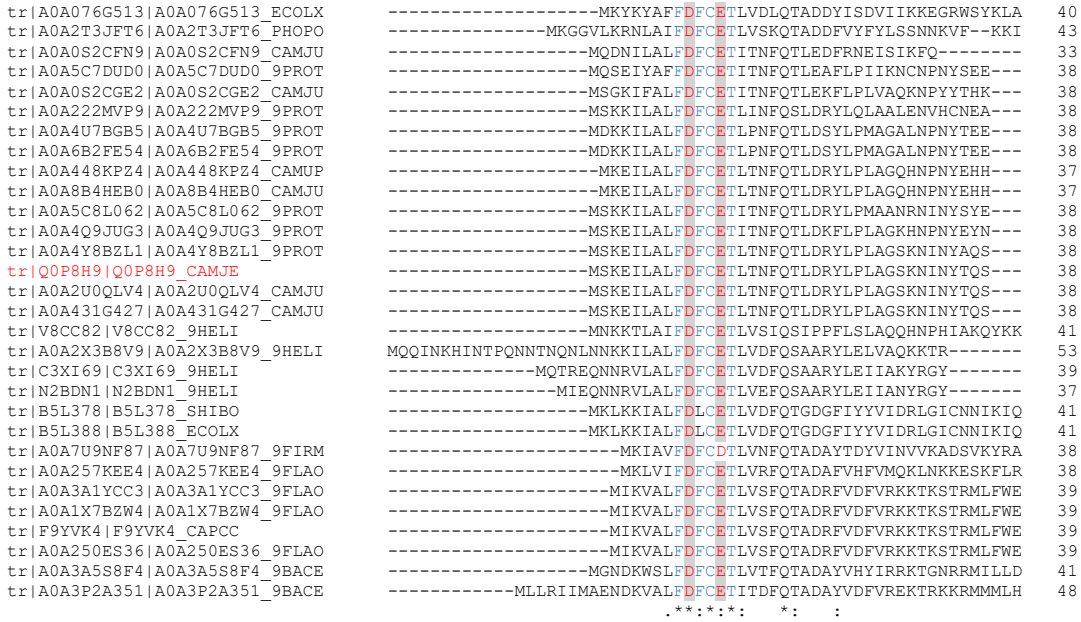


Figure S17 Alignment of the catalytic motif of Cj1435 with other HAD phosphatases involved in the biosynthesis of glucuronamide modifications. Cj1435 (UniProt id: Q0P8H9) is shown red on the left. Catalytic residues are shown in red and the DXXE motif is highlighted in grey. Other conserved residues are shown in blue.

Figure S18 Genome neighborhood network for the capsule amide ligase pathway. Network is aligned on Cj1435 homologs colored yellow (Pfam12710). The UDP-glucose dehydrogenases are colored orange (Pfam00984). The ATP-grasp enzymes are colored green (Pfam14305). The aminotransferases are colored dark blue (Pfam00155).

Table S1 Putative Amide Bond Forming Enzymes

Organism	Genome	Homolog to Cj1435	Homolog to Cj1436	Homolog to Cj1437	Homolog to Cj1438	Homolog to Cj1441
Campylobacter jejuni HS:2	AL111168	Q0P8H9	Q0P8H8	Q0P8H7	Q0P8H6	Q0P8H3
Escherichia coli	CP021726	A0A076G513		A0A076G300	A0A2Y0NZX7	
Capnocytophaga canimorsus	CDOK01000115	A0A0B7IGS5		A0A0B7ID17	A0A0B7ID22	A0A0B7IIH2
Campylobacter jejuni	KT893430	A0A0S2CFN9		A0A0S2CFS2	A0A0S2CGU1	A0A0S2CFN0
Campylobacter jejuni subsp. jejuni str. RM3421	KT893439	A0A0S2CGE2		A0A0S2CGI6	A0A0S2CHH3	A0A0S2CGH8
Capnocytophaga canimorsus	LR134513	A0A1X7BZW4		A0A1X7BZ63	A0A0B7ID22	A0A448U6C1
Campylobacter avium LMG 24591	CP022347	A0A222MVP9		A0A222MV91	A0A222MVJ5	
Capnocytophaga sp. H2931	CP022381	A0A250ES36		A0A250ENY4	A0A250ESA5	A0A250ENX2
Flavobacterium sp. BFFFF2	NKJD01000007	A0A257KEE4		A0A257KCP9	A0A257KCG9	
Photobacterium phosphoreum	PYMP01000023	A0A2T3JFT6		A0A2T3JFU8	A0A2T3JFT1	A0A2T3JFV6
Campylobacter jejuni	AABNCU010000001	A0A2U0QLV4		A0A2U0QLW6	A0A825NF14	A0A2U0QQW0
Helicobacter fennelliae	UAWL01000006	A0A2X3B8V9		A0A2X3AZ54	A0A2X3DFB5	
Capnocytophaga canis	NSDI01000008	A0A3A1YCC3		A0A3A1YIC1	A0A3A1YI56	A0A3A1YF59
Bacteroides sp. AM30-16	QTNI01000033	A0A3A5S8F4		A0A3A5S8H6	A0A3A5SM71	A0A3A5S9Z6
Bacteroides heparinolyticus	RQYF01000054	A0A3P2A351		A0A3P2A1N8	A0A3P2A2V9	
Campylobacter jejuni	PRCL01000002	A0A431G427	A0A431G425	A0A431G5E0	A0A431G448	A0A2U0QQW0
Campylobacter upsaliensis	LR134372	A0A448KPZ4	A0A3S4SU52	A0A3S4SPC0	A0A448KQ00	A0A448KPZ9
Campylobacter novaezeelandiae	QPGR01000006	A0A4Q9JUG3		A0A4Q9JUF2	A0A4V6MUU8	A0A4Q9JUW4
Campylobacter sp. MIT 12-5580	NXLK01000003	A0A4U7BGB5	A0A4U7BHD1	A0A4U7BH66	A0A4U7BDJ0	
Campylobacter sp. US18a	RYYO01000042	A0A4Y8BZL1		A0A4Y8C200	A0A4Y8BZK6	
Campylobacter volucris	VOWJ01000010	A0A5C7DUD0		A0A5C7DWZ3	A0A5C7E2X9	
Campylobacter volucris	VRMB01000006	A0A5C8L062		A0A5C8L607	A0A5C8L0A1	A0A5C8L3U0
Vibrio parahaemolyticus	MK473649	A0A5P5X5L8		A0A5P5X5H6	A0A5P5X5K6	
Campylobacter jejuni	AANOUJ010000006	A0A623SU25		A0A3Z8SM41	A0A430VAU7	A0A2U0QQW0
Campylobacter sp. MIT 19-121	JAADJR010000001	A0A6B2FE54		A0A6B2FAC3	A0A6B2FFD9	
Lachnospiraceae bacterium	BLLY01000214	A0A7U9NF87		A0A7U9NF68	A0A7U9R9P4	
Campylobacter coli	AACSZU010000001	A0A825YW84		A0A825Y3X3	A0A825YCH7	A0A825YI01

Campylobacter jejuni subsp. jejuni	SZUF01000003	A0A851H9Y9		A0A851HF95	A0A851HGL4	A0A851HE12
Campylobacter jejuni subsp. doylei	LS483295	A0A8B4HEB0	A0A8B4HDX1	A0A8B4HEB3	A0A8B4HDH3	A0A8B4HDG6
Shigella boydii	UAUR01000007	B5L378		A0A2X2J7W5	B5L379	
Escherichia coli	CP065613	B5L388		A0A377E9B2	A0A377E982	
Helicobacter bilis ATCC 43879	KI392040	C3XI69		T5LPM0	T5LS62	
Capnocytophaga canimorsus (strain 5)	CP002113	F9YVK4		F9YVK3	F9YVK5	F9YVJ9
Helicobacter bilis WiWa	AQFW01000014	N2BDN1		N2BIV3	N2B9M1	
Helicobacter macacae MIT 99-5501	AZJI01000001	V8CC82		V8CE12	V8CD75	

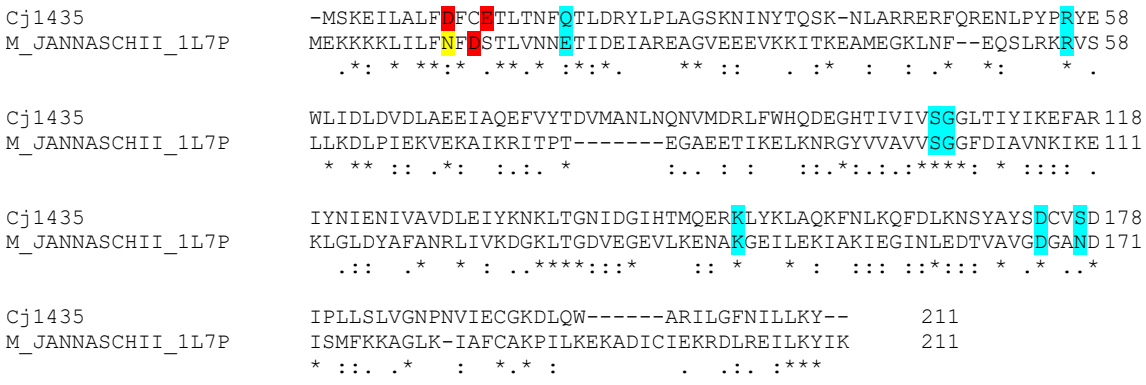


Figure S19 Amino acid sequence alignment of Cj1435 with the HAD phosphatase from *M. jannaschii* (PDB id: 1L7P). The residues shown in light blue are likely responsible for substrate binding. The residues shown in red are necessary for catalysis. Residue 11 of the HAD phosphatase from *M. jannaschii* is highlighted in yellow to illustrate this residue was mutated from an aspartate to asparagine in order to allow the substrate L-serine phosphate to bind.

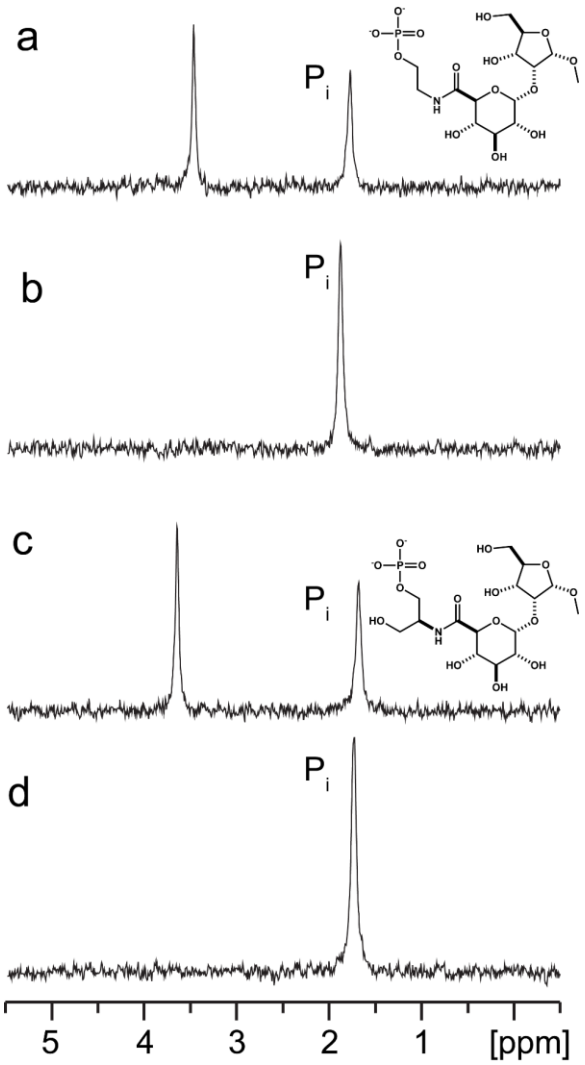


Figure S20 ^{31}P NMR spectra of the reaction products catalyzed by Cj1435 and the disaccharide of D-glucuronate (1→2) 1-*O*-methyl-D-ribose. (a) The products of the reaction catalyzed by Cj1438, MgATP, the disaccharide of D-glucuronate (1→2) 1-*O*-methyl-D-ribose (**13**), and ethanolamine phosphate (**3**). The resonance for the disaccharide of the D-glucuronamide of ethanolamine phosphate (1→2) 1-*O*-methyl-D-ribose (**15**) is a singlet at 3.46 ppm and the inorganic phosphate resonance appears as a singlet at 1.76 ppm. (b) The addition of Cj1435 results in complete hydrolysis of the substrate to inorganic phosphate shown as a singlet at 1.88 ppm. (c) The products of the reaction catalyzed by Cj1438, MgATP, the disaccharide of D-glucuronate (1→2) 1-*O*-methyl-D-ribose (**13**), and *S*-serinol phosphate (*S*-**4**). The resonance for the disaccharide of the glucuronamide of *S*-serinol phosphate (1→2) 1-*O*-methyl-D-ribose is a singlet at 3.65 ppm and the inorganic phosphate resonance appears as a singlet at 1.68 ppm. (d)

The addition of Cj1435 results in complete hydrolysis of the substrate to inorganic phosphate shown as a single resonant peak at 1.73 ppm. The compound designations are for those structures that appear in Figure 25 of the main text.

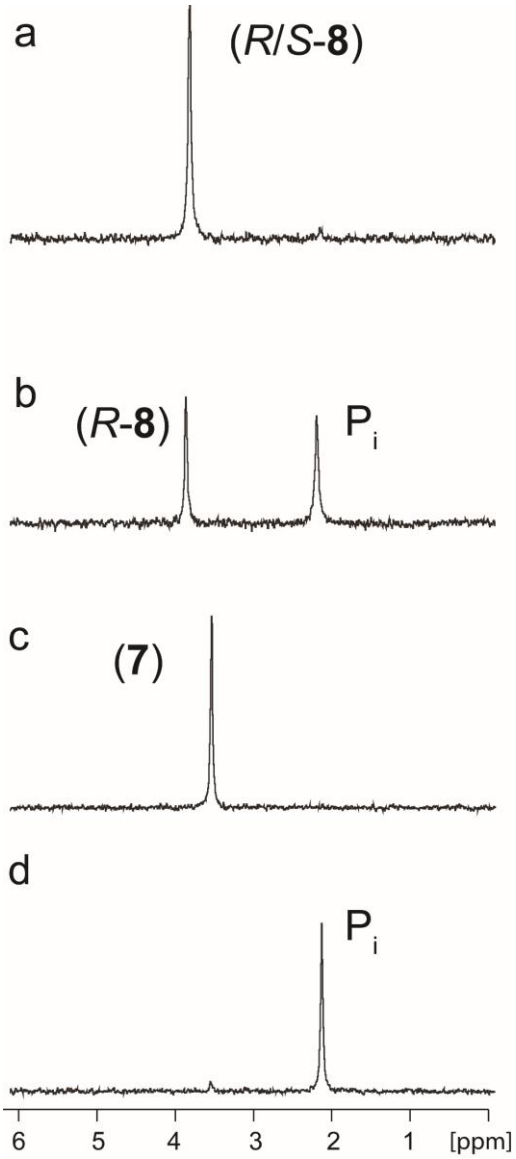
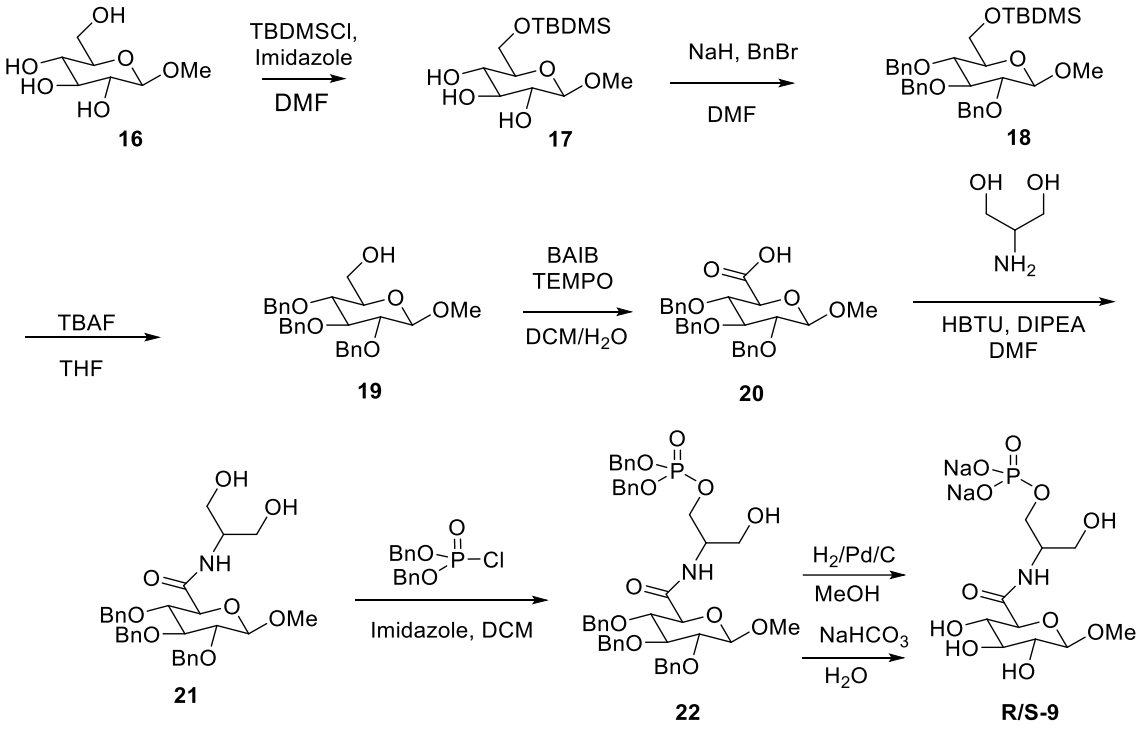


Figure S21 ^{31}P NMR spectra of the reaction products catalyzed by Cj1435 and *N*-acetyl *R*/*S*-serinol phosphate or *N*-acetyl ethanolamine phosphate. (a) A no-enzyme control of *N*-acetyl *R*/*S*-serinol phosphate (*R*/*S*-**8**) which appears as a singlet at 3.82 ppm. (b) The addition of Cj1435 results in hydrolysis of the (*S*)- enantiomer to inorganic phosphate, shown as a singlet at 2.20 ppm. The (*R*)- enantiomer (*R*-**8**) remains as a singlet at 3.86 ppm. (c) A no-enzyme control of *N*-acetyl ethanolamine phosphate (**7**) appears as a singlet at 3.54 ppm. (d) The addition of Cj1435 results in complete hydrolysis of the substrate to inorganic phosphate shown as a singlet at 2.13 ppm. The compound designations are for those structures that appear in Figure 25 of the main text.

Synthesis of (*R*/*S*-9) from Scheme S3.

Methyl beta-D-glucopyranoside **16** was converted to benzyl protected D-glucuronic acid **20** according literature procedures (9**,** 10). **20** was coupled with serinol to obtain amide **21**, which was monophosphorylated with dibenzyl phosphorochloridate to obtain compound **22**. Compound (***R*/*S*-9**) was obtained by catalytic hydrogenation of **22** and subsequent neutralization.

Scheme S3



Synthesis of Compound 21 from Scheme S3.

To a solution of **20** (0.13 g, 0.27 mmol, 1 equiv.) in anh. DMF (2.0 mL) was added: TBTU (0.13 g, 0.40 mmol, 1.5 equiv.), DIEA (0.14 mL) and serinol (0.036 g, 0.40 mmol, 1.5 equiv). Reaction mixture was stirred for 16 h at ambient temperature (23 ^{o}C). After concentrating, the residue was purified by silica gel column chromatography (EtOAc) to afford 0.10 g (67 %) **21** as a white solid.

^{1}H NMR (400 MHz, DMSO-d$_{6}$) δ 7.81-7.66 (m, 15H), 7.47 (d, J = 7.2 Hz, 1H), 5.28 (d, J = 11.4 Hz, 1H), 5.21(d, J = 11.3 Hz, 1H), 5.18-5.09 (m, 4H), 4.93 (d, J = 7.6 Hz, 1H), 4.44 (d, J = 7.8 Hz, 1H), 4.39-4.32 (m, 1H), 4.27-4.11 (m, 8H), 4.00 (s, 3H), 3.91 (t, J = 7.6 Hz, 1H). ^{13}C NMR (125 MHz, DMSO-d$_{6}$) δ 164.9, 164.9, 133.5, 133.4, 133.0, 123.6, 123.4, 123.2, 123.1, 123.0, 123.0, 99.4, 78.2, 76.5, 74.7, 72.4, 70.7, 70.3, 69.9, 69.6, 58.3, 52.6. HRMS (ESI^{+}) m/z [M + H]$^{+}$ calcd. for $C_{31}H_{38}NO_{8}$: 552.2592, found: 552,2592.

Synthesis of Compound 22 from Scheme S3.

A solution of **21** (0.42 g, 0.76 mmol, 1.0 equiv.) and imidazole (0.10 g, 1.5 mmol, 2.0 equiv.) in 20 mL of anh. dichloromethane was chilled to 0 ^{o}C and dibenzyl phosphorochloridate (0.26 g, 0.91 mmol, 1.2 equiv.) was added. Reaction was allowed to warm up to ambient temperature (23 ^{o}C) and stirred for 16 h. After filtering and concentrating, the residue was purified by silica gel column chromatography (hexanes/ethyl acetate, 1:3) to afford **22** (0.16 g, 0.20 mmol) as a colorless oil. (26 % yield).

^{1}H NMR (400 MHz, MeOH-d$_{4}$) δ 7.41-7.17 (m, 25H), 5.06-4.97 (m, 4H), 4.88-4.81 (m, 2H), 4.72 (dd, J_1 = 11.4 Hz, J_2 = 4.2 Hz, 1H), 4.69-4.63 (m, 3H), 4.43 (dd, J_1 = 7.8 Hz, J_2 = 4.2 Hz, 1H), 4.24-4.08 (m, 3H), 3.90 (dd, J_1 = 9,2 Hz, J_2 = 7.2 Hz, 1H), 3.75 (t, J = 8.7 Hz, 1H), 3.68-3.52 (m, 3H), 3.50 (s, 1.5H), 3.48 (s, 1.5H), 3.44-3.37 (m, 1H). ^{31}P

NMR (160 MHz, MeOH-d_4) δ -1.23 (s), -1.35 (s). HRMS (ESI^+) m/z $[M + H]^+$ calcd. for $C_{45}H_{51}NO_{11}P$: 812.3200, found: 812.3207.

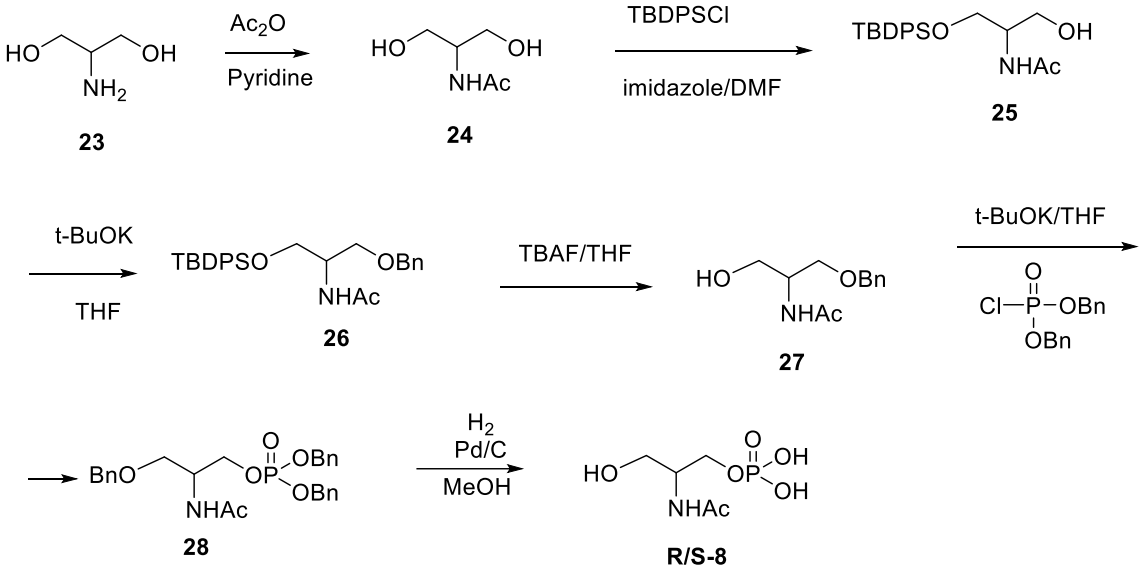
Synthesis of Compound (*R*/*S*-9) from Scheme S3.

To the solution of **22** (0.16 g, 0.20 mmol) in 20 mL of methanol, Pd/C (10%) (0.12 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtering and concentrating, the residue was dissolved in 3 mL of water, neutralized with $NaHCO_3$ sat. solution and lyophilized to obtain white solid of (*R*/*S*-**9**) (0.08g, quantitative yield).

NMR spectra and mass data are reported for acidic form of the final compound. ^1H NMR (400 MHz, MeOH-d_4) δ 4.30 (d, J = 7.9 Hz, 1H), 4.21-4.14 (m, 1H), 4.13-4.07 (m, 2H), 3.80 (d, J = 9.5 Hz, 1 H), 3.75-3.66 (m, 2H), 3.56 (s, 3H), 3.55-3.49 (m, 1H), 3.44 (t, J = 9.04, 1H), 3.28-3.22 (m, 1H). ^{31}P NMR (160 MHz, MeOH-d_4) δ -0.11 (s), -0.13 (s). HRMS (ESI^-) m/z $[M\ \ H]^-$ calcd. for $C_{10}H_{19}NO_{11}P$: 360.0696, found: 360.0705.

Synthesis of *N*-Acetyl Serinol Monophosphate (*R*/*S*-8) from Scheme S4.

Scheme S4



Serinol **23** was converted into *N*-acetyl serinol **24** by described procedure (11). Direct mono phosphorylation of **24** was not successful due to its poor solubility in conventional solvents (DCM, THF, ethyl ether). To increase solubility for the phosphory-lation step, **24** was converted into **25** (11), which was further benzylated into **26** and TBDPS was removed obtain **27**. Phosphorylation of **27** affored **28** and catalytic hydrogenation of **28** afforded N-acetyl serinol monophosphate (*R*/*S*-**8**).

Synthesis of Compound 26 from Scheme S4.

Solution of **25** (0.47 g, 1.26 mmol, 1 equiv.) in 10 mL of anh. THF was chilled to 0^{o}C and potassium t-butoxide (1.4 mL, 1.4 mmol, 1.1 equiv.; 1M in THF) was added. Solution was stirred for 30 min at the same temperature and benzyl bromide (1.2 mL, 1.51 mmol, 1.2 equiv) was added. Reaction was allowed to warm up to ambient temperature (23 ^{o}C) and stirred for 16 h. After filtering and concentrating, the residue was purified by silica gel column chromatography (hexanes/ethyl acetate, 2.5:2) to afford **26** (0.40 g, 0.87 mmol) as a colorless oil (68 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.49 - 7.44 (m, 4H), 7.29-7.10 (m, 11H), 5.54 (d, J = 8.4 Hz, 1H), 4.34 (s, 2H), 4.12 - 4.05 (m, 1H), 3.68 (dd, J_1 = 9.8, J_2 = 4.2 Hz, 1H), 3.60 – 3.52 (m, 2H), 3.40 (dd, J_1 = 9.2, J_2 = 5.4 Hz, 1H), 1.73 (s, 3H), 0.89 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 169.5, 138.1, 135.5, 135.5, 133.4, 133.2, 73.2, 68.2, 62.2, 50.0, 26.8, 23.6, 19.3.

HRMS (ESI^{+}) m/z $[M + H]^{+}$ calcd. for $C_{28}H_{36}NO_3Si$: 462.2464, found: 462.2456.

Synthesis of Compound 27 from Scheme S4.

A solution of **26** (0.40 g, 0.87 mmol, 1 equiv.) in anh. THF (10 mL) was chilled to 0^{o}C and TBAF (2.6 mL, 2.6 mmol, 3 equiv., 1M THF) was added. Solution was stirred for 3 hours at ambient temperature (23 ^{o}C), then concentrated and the residue

purified by silica gel column chromatography (DCM:MeOH, 20:1) to afford **27** (0.16 g, 0.72 mmol) as a colorless oil (82 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.25 - 7.10 (m, 5H), 6.02 - 5.39 (m, 1H), 4.34 (s, 2H), 3.90 (quin., J = 4.1 Hz, 1H), 3.65 (dd, J_1 = 11.2, J_2 = 4.3 Hz, 1H), 3.52 - 3.42 (m, 3H), 1.81 (s, 9H). ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.5, 137.5, 134.8, 128.5, 128.0, 127.7, 127.6, 73.6, 70.6, 63.9, 50.8, 23.3.

(ESI$^+$) m/z $[M + H]^+$ calcd. for $C_{12}H_{18}NO_3$: 224.1287, found: 224.1281.

Synthesis of Compound 28 from Scheme S4.

A solution of **27** (0.216 g, 0.97 mmol, 1 equiv.) in anh. THF (10 mL) was chilled to 0^{o}C, potassium t-butoxide (1.4 mL, 1.4 mmol, 1.1 equiv.; 1M in THF) added and stirred for 30 min. Dibenzyl chlorophosphonate (0.55 g, 1.94 mmol, 2.0 equiv.) was added and reaction allowed to warm up to ambient temperature (23 ^{o}C) and stirred for 18 h. After filtering and concentrating, the residue was purified by silica gel column chromatography (DCM:MeOH, 20:0.5) to afford **28** (0.21 g, 0.43 mmol) as a colorless oil (45 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 - 7.27 (m, 15H), 6.06 (d, J = 8.4 Hz, 1H), 5.10 – 4.99 (m, 4H), 4.51, 4.47 (ABq, 2H, J_{AB} = 12.0 Hz, 2H), 4.34-4.25 (m, 1H), 4.23-4.16 (m, 1H), 4.11-4.03 (m, 1H), 3.56 (dd, J_1 = 9.4, J_2 = 4.2 Hz, 1H), 3.45 (dd, J_1 = 9.4, J_2 = 6.4 Hz,1H), 1.91 (s, 3H). ^{31}P NMR (160 MHz, $CDCl_3$) δ – 0.44 (s).

(ESI$^+$) m/z $[M + H]^+$ calcd. for $C_{26}H_{31}NO_6P$: 484.1889, found: 484.1881

Synthesis of Compound (*R*/*S*-8) from Scheme S4.

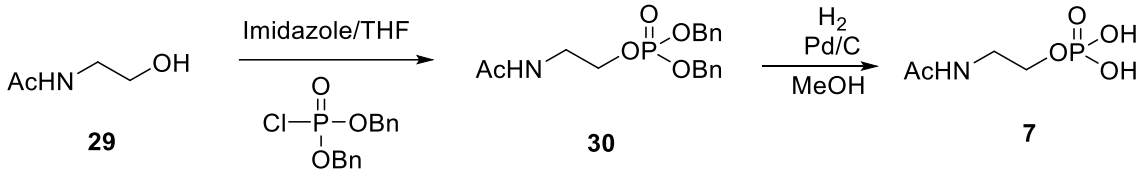
To the solution of **28** (0.21 g, 0.43 mmol) in 20 mL of methanol, Pd/C (10%) (0.12 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtering and concentrating, (*R*/*S*-**8**) (0.09 g, 0.42 mmol) was obtained as colorless oil (98% yield).

^{1}H NMR (400 MHz, MeOH-d_4) δ 4.16-4.03 (m, 3H), 3.65 (d, J = 5.6 Hz, 2H), 1.99 (s, 3H). ^{13}C NMR (125 MHz, MeOH-d_4) δ 172.1, 64.5 (d, J = 5.4 Hz, 1C), 59.97, 51.4 (d, J = 8.2 Hz, 1C), 21.3. ^{31}P NMR (160 MHz, MeOH-d_4) δ – 0.08 (s).

HRMS (ESI$^-$) m/z [M - H]$^-$ calcd. for $C_5H_{11}NO_6P$: 212.0324, found: 212.0322

Synthesis of *N*-acetyl ethanolamine phosphate (7) from Scheme S5.

Scheme S3



Synthesis of Compound 30 from Scheme S5.

To the solution of N-acetylethanolamine **29** (0.34 g, 3.00 mmol, 1.0 equiv., 90% purity) and imidazole (0.34 g, 5.0 mmol, 1.67 equiv.) in 15 mL of anh. dichloromethane, dibenzyl chlorophosphonate (0.86 g, 3.0 mmol, 1.0 equiv.) was added. Reaction stirred for 16 h at ambient temperature (23 ^{o}C). After filtering and concentrating, the residue was purified by silica gel column chromatography (ethyl acetate/MeOH, 10:1) to afford **30** (0.35 g, 0.96 mmol) as a colorless oil. (32 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 - 7.23 (m, 10H), 5.98 (brs, 1H), 5.06-4.91 (m, 4H), 3.98-3.91 (m, 2H), 3.38-3.32 (m, 2H), 1.82 (s, 3H). ^{31}P NMR (160 MHz, $CDCl_3$) δ – 0.40 (s).

(ESI^{+}) m/z $[M + H]^{+}$ calcd. for $C_{18}H_{23}NO_5P$: 364.1314, found: 364.1311.

Synthesis of Compound 7 from Scheme S5.

To the solution of **30** (0.35 g, 0.96 mmol) in 15 mL of methanol, Pd/C (10%) (0.10 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtering and concentrating, **7** (0.17 g, 0.93 mmol) was obtained as colorless oil (97 % yield).

^{1}H NMR (400 MHz, D_2O) δ 4.03-3.96 (m, 2H), 3.44 (t, J = 5.4 Hz, 2H), 1.97 (s, 3H). ^{31}P NMR (160 MHz, D_2O) δ + 0.40 (s).

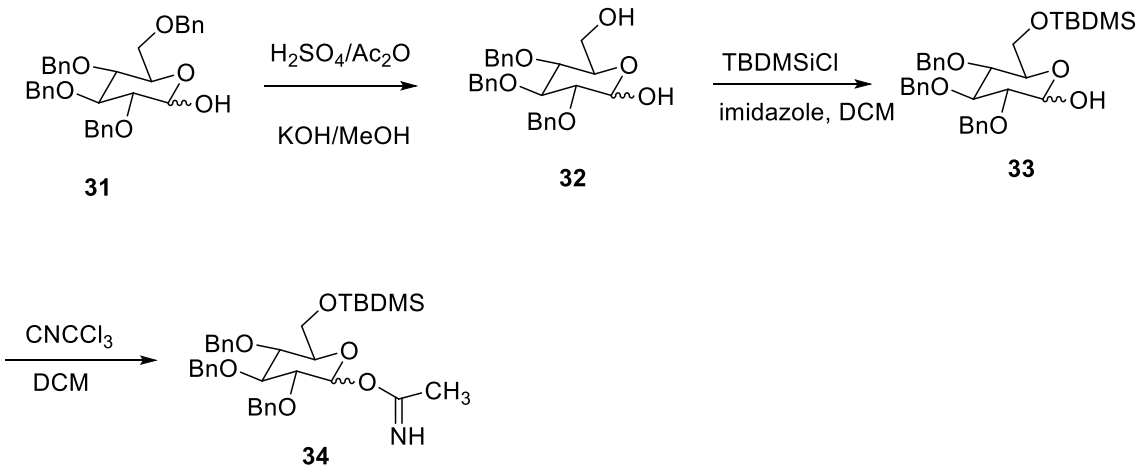
(ESI$^-$) m/z [M - H]$^-$ calcd. for $C_4H_9NO_5P$: 182.0218, found: 182.0215.

Synthesis of Compound 13 from Scheme S6, Scheme S7, and Scheme S8.

Glycosyl donor **34** and acceptor **38** were synthesized and combined by glycosylation into compound **39**, further transformations allowed disaccharide **13**. Synthesis is outlined on **Schemes S4**, **S5**, and **S6**.

Synthesis of Compound 34 from Scheme S6.

Scheme S6

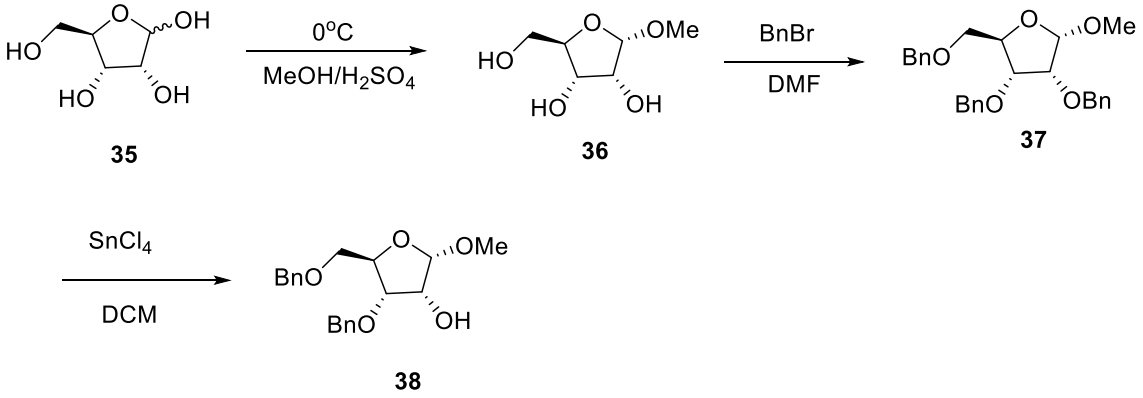


Glycosyl donor **34** was synthesized from commercially available D-glucose derivative **31**. Compound **33** was synthesized according to literature procedures (12). Reaction of **33** with trichloro acetonitrile in the presence of base afforded trichloroacetimidate **34**, as a mixture of α- and β- enantiomers. (13)

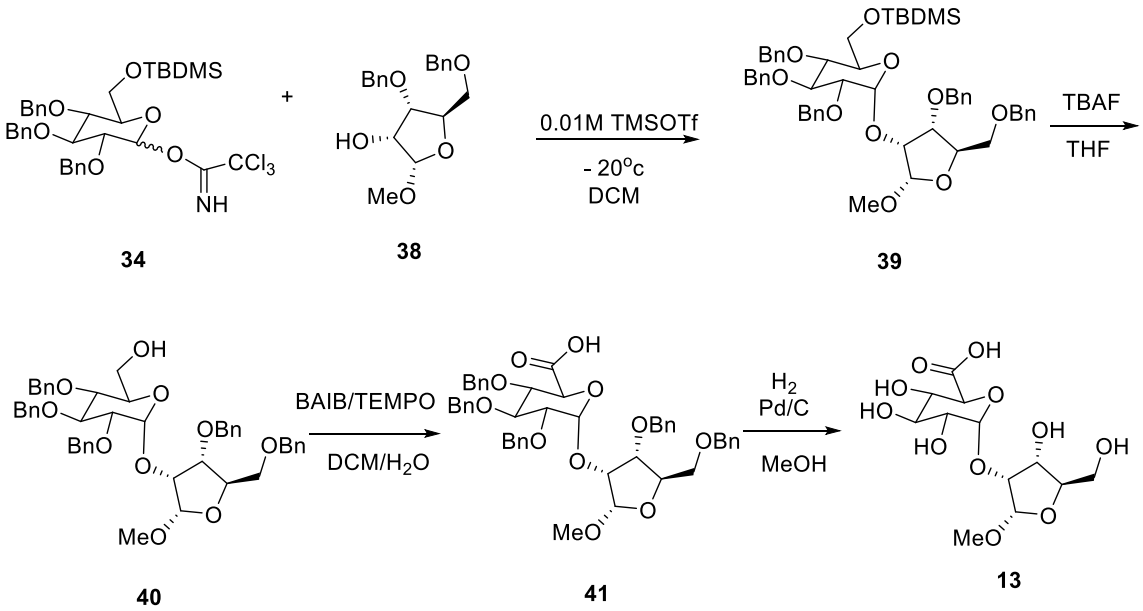
Synthesis of compound 38 from Scheme S7.

Glycosyl acceptor **38**, β-OMe D-ribose derivative, was synthesized from D-ribose **35** according literature procedures (**Scheme S5**) (14).

Scheme S7



Scheme S8



Synthesis of Compound 39 from Scheme S8.

Compounds **34** (0.12 g, 0.17 mmol, 1 equiv.) and **38** (0.11 g, 0.31 mmol, 1.8 equiv.) were dissolved in 3.0 mL anh. dichloromethane and 4A^{o} (0.24 g) was added. The mixture was stirred for 20 min at room temperature and then cooled to - 20^{o}C; trimethylsilyltrifluoro methanesulfonate (TMSOTf) (1.0 mL, 0.01 M in DCM) was added and stirred for 15 min at the same temperature. Triethylamine was used to neutralize the mixture and after concentrating, the residue was subjected to silica gel column chromatography (Hexanes:EtOAc, 4:1) to afford alpha enantiomer **39** (0.11 g, 0.12 mmol) as a colorless oil (73 % yield) (alpha enantiomer was the major product of the glycosylation reaction).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 - 7.12 (m, 25H), 5.06 (d, J = 3.4 Hz, 1H), 5.04 (d, J = 4.2 Hz, 1H), 4.91(d, J = 7.2 Hz, 1H), 4.44(d, J = 7.6 Hz, 1H), 4.72-4.58 (m, 4H), 4.57-4.52 (m, 2H), 4.51(d, J = 8.2 Hz, 2H), 4.34-4.25 (m, 1H), 4.13-4.07 (m, 1H), 4.05 (d, J = 9.5 Hz, 1H), 3.91-3.87 (m, 1H), 3.71-3.65 (m, 3H), 3.52-3.51(m, 2H), 3.48 (s, 3H), 0.85 (s, 9H), 0.01 (s, 3H), 0.00 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 138.8, 138.7, 138.6, 138.3, 138.0, 128.5, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.7, 127.6, 127.5, 127.5, 1275.4, 101.7, 95.5, 81.9, 81.6, 79.9, 76.2, 75.6, 74.9, 74.4, 73.5, 72.9, 72.4, 71.9, 70.1, 62.3, 55.3, 25.9, 18.3, 5.34, 5.15.

(ESI^{+}) m/z $[M + H]^{+}$ calcd. for $C_{53}H_{69}O_{10}Si$: 891.4503, found: 891.4512.

Synthesis of Compound 40 from Scheme S8.

A solution of **39** (0.27 g, 0.30 mmol, 1 equiv.) in anhydrous THF (10 mL) was chilled to 0^{o}C and TBAF (0.9 mL, 0.9 mmol, 3 equiv., 1M THF) was added. Solution was stirred for 5 hours at ambient temperature (23 ^{o}C), then concentrated and the residue subjected to silica gel column chromatography (DCM:MeOH, 20:1) to obtain **40** (0.075 g, 0.1 mmol) as a colorless oil (32 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 - 7.07 (m, 25H), 4.99 (d, J = 3.4 Hz, 1H), 4.97 (d, J = 4.2 Hz, 1H), 4.86 (d, J = 10.9 Hz, 1H), 4.82 (d, J = 11.0 Hz, 1H), 4.75-4.37 (m, 8H), 4.24(q, J = 3.6 Hz, 1H), 4.07-3.98 (m, 2H), 3.85 (dd, J_1 = 6.6 Hz, J_2 = 3.3 Hz, 1H), 3.68-3.61 (m, 1H), 3.61-3.51 (m, 2H), 3.50-3.43 (m, 3H), 3.42 (s, 3H), 3.42-3.38 (m, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 138.8, 138.5, 138.3, 138.2, 137.9, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 101.7, 95.8, 82.1, 81.4, 79.8, 77.2, 77.1, 76.2, 75.6, 74.9, 74.9, 73.5, 72.9, 71.9, 71.5, 70.1, 55.4.

(ESI^+) m/z $[M + H]^+$ calcd. for $C_{47}H_{53}O_{10}$: 777.3639, found: 777.3630.

Synthesis of Compound 41 from Scheme S8.

Compound **40** (0.07 g, 0.1 mmol, 1 equiv.) was dissolved in dichloromethane-water (3.3 mL-0.6 mL) and cooled to 0^{o}C. TEMPO (0.005 g, 0.025 mmol, 0.24 equiv) and BAIB (0.081 g, 0.25 mmol, 2.5 equiv) were added and stirred for 18 h at room temperature. Sodium sulfite (0.016 g, 0.12 mmol, 1.2 equiv) was added to the reaction, stirred for 10 min, organic solvent was removed and aqueous solution was acidified with1 drop of 4 M HCl. Acidified aqueous solution was extracted with EtOAc (10 mL x3), washed with 1 M HCl (5 mL x1) and after condensation, the residue was subjected to silica gel column chromatography (Hexanes:EtOAc, 1:1 and EtOAc) to obtain **41** (0.040 g, 0.05 mmol) as a colorless oil (51 % yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 - 7.03 (m, 25H), 4.99 (d, J = 3.4 Hz, 1H), 4.97 (d, J = 4.2 Hz, 1H), 4.86 (d, J = 10.9 Hz, 1H), 4.82 (d, J = 11.0 Hz, 1H), 4.75-4.37 (m, 8H), 4.24(q, J = 3.6 Hz, 1H), 4.07-3.98 (m, 2H), 3.85 (dd, J_1 = 6.6 Hz, J_2 = 3.3 Hz, 1H), 3.68-3.61 (m, 1H), 3.61-3.51 (m, 2H), 3.50-3.43 (m, 3H), 3.42 (s, 3H), 3.42-3.38 (m, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 138.8, 138.5, 138.3, 138.2, 137.9, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 101.7, 95.8, 82.1, 81.4, 79.8, 77.2, 77.1, 76.2, 75.6, 74.9, 74.9, 73.5, 72.9, 71.9, 71.5, 70.1, 55.4.

(ESI$^-$) m/z [M - H]$^-$ calcd. for $C_{47}H_{49}O_{11}$: 789.3275, found: 789.3219.

Synthesis of Compound 13 from Scheme S8.

To the solution of **41** (0.040 g, 0.05 mmol) in 10 mL of methanol, Pd/C (10%) (0.05 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtering and concentrating, 1 mL of water was added to the product and lyophilized. **13** (0.015 g, 0.044 mmol) was obtained as white solid (88 % yield).

^{1}H NMR (400 MHz, D_2O) δ 5.16 (d, J = 3.4 Hz, 1H), 5.14 (d, J = 3.9 Hz, 1H), 4.51 (d, J = 9.8 Hz, 0.2H), 4.43 (d, J = 9.8 Hz, 0.8H), 4.18-4.11 (m, 3H), 3.87-3.81 (m, 1H), 3.75-3.63 (m, 2H), 3.62-3.56 (m, 2H), 3.44 (s, 3H). ^{13}C NMR (125 MHz, D_2O) δ 173.2, 101.8, 99.2, 85.6, 77.0, 72.5, 71.4, 71.3, 71.2, 69.7, 61.5, 55.1.

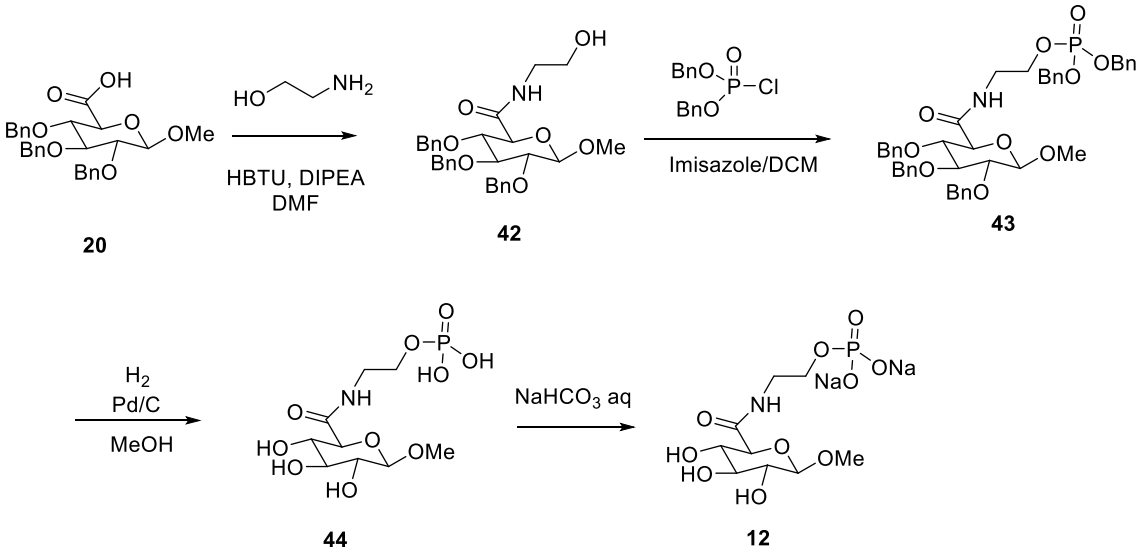
(ESI$^-$) m/z [M - H]$^-$ calcd. for $C_{12}H_{19}O_{11}$: 339.0927, found: 339.0936.

Synthesis of methyl-β-D-glucuronic acid ethanolamide phosphate 12 from Scheme S9.

Compound **20** was coupled to ethanolamine to obtain **42** in identical conditions described for the synthesis of compound **21.** Phosphorylation of **42** to obtain **43** was conducted as described for the compound **22**.

Catalytic hydrogenation of **43** allowed compound **44**, which was neutralized with $NaHCO_3$ sat. solution and lyophilized to afford final salt **12**.

Scheme S9



Synthesis of Compound 44 and 12 from Scheme S9.

To the solution of **43** (0.054 g, 0.07 mmol) in 10 mL of methanol, Pd/C (10%) (0.08 g) was added and the reaction mixture was subjected to H_2 (1 atm) for 16 h. After filtering and concentrating, product **44** was isolated (0.020 mg, 88%).

44 was dissolved in 1 mL of water, neutralized with $NaHCO_3$ sat. water solution and lyophilized to obtain **12** as a white solid.

Compound 44 : ^{1}H NMR (400 MHz, MeOH-d_4) δ 4.26 (d, J = 7.6 Hz, 1H), 4.09-4.02 (m, 2H), 3.75 (d, J = 9.4 Hz, 1H), 3.56 (s, 3H), 3.55 - 3.46 (m, 3H), 3.42 (t, J = 8.8 Hz, 1H), 3.27 - 3.21 (m, 1H).

^{31}P NMR (160 MHz, MeOH-d_4) δ + 0.12 (s).

(ESI$^-$) m/z $[M - H]^-$ calcd. for $C_9H_{17}NO_{10}P$: 330.0590, found: 330.0596.

References

1. Fawaz, M. V., Topper, M. E. and Firestine, S. M. (2011) The ATP-grasp enzymes. Bioorganic Chemistry 39, 185-191.
2. Parkhill, J., Wren, B. W., Mungall, K., Ketley, J. M., Churcher, C., Basham, D., Chillingworth, T., Davies, R. M., Feltwell, T., Holroyd, 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., Whitehead, S., and Barrell, B. G. (2000) The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences. Nature 403, 665– 668.
3. Parker, C. T., Huynh, S., & Heikema, A. P. (2016) Complete Genomic Sequence of Campylobacter jejuni subsp. jejuni HS:19 Strain RM1285 Isolated from Packaged Chicken. Genome announcements, 4(5), e01100-16.
4. Poly, F., Serichantalergs, O., Kuroiwa, J., Pootong, P., Mason, C., Guerry, P., & Parker, C. T. (2015) Updated Campylobacter jejuni Capsule PCR Multiplex Typing System and Its Application to Clinical Isolates from South and Southeast Asia. PloS one, 10(12), e0144349.
5. Harrison, K. J., Crécy-Lagard, V., & Zallot, R. (2018) Gene Graphics: a genomic neighborhood data visualization web application. Bioinformatics (Oxford, England), 34(8), 1406–1408.
6. Keller, S., Wetterhorn, K. M., Vecellio, A., Seeger, M., Rayment, I., & Schubert, T. (2019) Structural and functional analysis of an L-serine O-phosphate decarboxylase involved in norcobamide biosynthesis. FEBS letters, 593(21), 3040–3053.

7. Cheong, C. G., Escalante-Semerena, J. C., & Rayment, I. (2002) Structural studies of the L-threonine-O-3-phosphate decarboxylase (CobD) enzyme from Salmonella enterica: the apo, substrate, and product-aldimine complexes. Biochemistry, 41(29), 9079–9089.
8. J. Y. Choi and R. F. Borch (2007) Highly Efficient Synthesis of Enantiomerically Enriched 2-Hydroxymethylaziridines by Enzymatic Desymmetrization. Organic Letters 9, 215-218.
9. Ágoston, K., Ágoston, Á., Dorgan, C. R., & Fügedi, P. (2015). A new method testing the orthogonality of different protecting groups. Carbohydrate research, 418, 98–103.
10. Wan, I., Witte, M. D., & Minnaard, A. J. (2019). From d- to L-Monosaccharide Derivatives via Photodecarboxylation-Alkylation. Organic letters, 21(18), 7669–7673.
11. Doppalapudi, V. R., Cochran, M. C., Chu, D. S., Arias, J. D., Burke, R. (2020) Nucleic acid-polypeptide compositions and uses thereof. Patent number WO2020/247782. World Intellectual Property Organization (PCT).
12. Shiozaki, M., Tashiro, T., Koshino, H., Shigeura, T., Watarai, H., Taniguchi, M., & Mori, K. (2013). Synthesis and biological activity of hydroxylated analogues of KRN7000 (α-galactosylceramide). Carbohydrate research, 370, 46–66.
13. Aly, M. R. E., Castro-Palomino, J. C., El-Sayed, I. I., El-Sayed E. H., Schmidt, R. R. (1998). The Dimethylmaleoyl Group as Amino Protective Group – Application

to the Synthesis of Glucosamine-Containing Oligosaccharides. Euro. J. Org. Chem., 1998, 2305-2316.

14. Klasson, B., Eneroth, A., Nilsson, M., Pinho, P., Samuelsson, B., Sund, C. (2013). HCV polymerase inhibitors. Patent number WO2013/084165. World Intellectual Property Organization (PCT).