

**Structural studies of *Campylobacter jejuni*:  
Virulence proteins (Cj0977 and JlpA) and novel lipoproteins**

---

**A Dissertation**

**Presented to**

**the Faculty of the Department of Biology and Biochemistry**

**University of Houston**

---

**In Partial Fulfillment**

**of the Requirements for the Degree**

**Doctor of Philosophy**

---

**By**

**Seonghee Paek**

**May 2013**

**Structural studies of *Campylobacter jejuni*:  
Virulence proteins (Cj0977 and JlpA) and novel lipoproteins**

APPROVED:

---

Seonghee Paek

---

Dr. Hye-Jeong Yeo, Chairperson

---

Dr. Anne Delcour

---

Dr. Chengzhi Cai

---

Dr. Joseph Eichberg

---

Dr. Dan E. Wells, Dean,  
College of Natural Sciences and Mathematics

## **Acknowledgements**

First and foremost, I would like to express my profound gratitude to my advisor, Dr. Hye-Jeong Yeo. I am truly blessed and honored to have been able to work in her laboratory over the past several years. Her full support and guidance has led me to continue and complete this study. I am truly grateful to my current and past committee members, Dr. Anne Delcour, Dr. Chengzhi Cai, Dr. Joseph Eichberg, and Dr. Rigoberto Advincula for their invaluable advice.

I am very thankful to all the my past and current lab members Dr. Kyoung-Jae Choi, Dr. Fumihiro Kawai, Dr. Takeshi Yokoyama, Philip Ho, Michal Szymanski, and Andrew Kirkpatrick. Their collaboration and support helped my research progress. My thanks also go to my dear friends, Kristina Henkel, Lucy Vela, Heather Brasher, Victor Stepanov, Danny N Yerly, Anna Dey, Prasenjit Dey, Xin Wang, Mediha Esra Altuntop, Pailinrut Chinwangso, Ashwini Shanbhogue, Deepika Kumar, and Hae Won Chung for their friendship. I really thank Missionary Bokson Kim and my church members for their intercessory prayer for my study.

I would like to thank my husband, Yo-il Yun who has supported me with love and my cute son, Paul Yun who gives me joy. I would like to thank my parents, Kyung-ho Paek and Mi-ja Seong, and my sisters, Seong-sil Paek, Seung-yeon Paek, and Seung-ju Paek, and my father-in-law, Suk-gu Yun, and all of my family-in-law for their encouragement. I also would like to thank my mother Dong-suk Jang in peace. I believe she is so happy for my accomplishment. Most of all, I would like to thank God who has led me here. Thank you all.

**Structural studies of *Campylobacter jejuni*:  
Virulence proteins (Cj0977 and JlpA) and novel lipoproteins**

---

**An Abstract of a Dissertation**

**Presented to**

**the Faculty of the Department of Biology and Biochemistry**

**University of Houston**

---

**In Partial Fulfillment**

**of the Requirements for the Degree**

**Doctor of Philosophy**

---

**By**

**Seonghee Paek**

**May 2013**

## Abstract

*Campylobacter jejuni* is a major cause of human gastroenteritis worldwide; however, its pathogenesis is not understood well. Cj0977, a cytoplasmic protein regulated by the flagellar promoter  $\sigma^{28}$ , and JlpA, a cell-surface exposed lipoprotein adhesin, are virulence factors, but their contributions in pathogenesis of *C. jejuni* remains elusive. Since there is no clue for these two protein structures, we aimed to discover structure-function relationship of these proteins. My work focused on establishing expression system, purification method, and crystallization method of these proteins, leading up to obtaining high quality crystals for successful X-ray diffraction.

Several techniques were employed to yield diffracting crystals: limited proteolysis and stability test for Cj0977 and introduction of two additional Met sites within the JlpA sequence. Here we present the first view of the crystal structures of Cj0977 and JlpA. Cj0977 adopts a ‘hot dog’ fold, a famous protein folds found in numerous CoA derivative binding enzymes. The Cj0977 structure suggests its possible function as an acyl-CoA binding regulatory protein. The JlpA structure reveals a novel fold of unclosed half  $\beta$ -barrel and is reminiscent of other bacterial lipoproteins. The JlpA structure suggests that JlpA may accommodate multiple ligands and that a similar role for JlpA as a carrier of as yet unidentified *Campylobacter*-specific lipids. The unique fold of JlpA led us to initiate other *C. jejuni* lipoprotein study.

Most of *C. jejuni* lipoproteins are unknown due to no apparent sequence homology in the sequence databank; however, they are considered to contribute to the pathogenesis. We aimed to uncover the crystal structures of *C. jejuni* lipoproteins to

provide framework for investigating their functions. My work focused on four lipoproteins: Cj0090, Cj1026c, Cj1090c, and Cj1649. Among these, the Cj0090 structure reveals a novel variant of the immunoglobulin fold with  $\beta$ -sandwich architecture, suggesting that Cj0090 may be involved in protein-protein interactions, consistent with a possible role for bacterial lipoproteins. In addition, a phage display technique was performed to identify ligands for our target lipoproteins.

The crystal structures revealed in this study contribute to provide clues for possible roles of our target proteins and enable to perform further structure-based functional analyses.

## Table of Contents

<b>Acknowledgements.....</b>	<b>iii</b>
<b>Abstract.....</b>	<b>v</b>
<b>Table of Contents.....</b>	<b>vii</b>
<b>List of Figures.....</b>	<b>xi</b>
<b>List of Tables.....</b>	<b>xiii</b>
<b>List of Abbreviations.....</b>	<b>xiv</b>
<b>Chapter I. Introduction</b>	<b>1</b>
1.1. <i>Campylobacter jejuni</i>	2
1.2. Virulence factors and pathogenesis of <i>C. jejuni</i>	5
1.2.1. Flagella and flagella-related virulence factors	6
1.2.2. Adhesins	11
1.2.3. Cytolethal distending toxin (CDT)	13
1.2.4. Capsular polysaccharide (CPS)	15
1.2.5. Lipooligosaccharide (LOS)	17
1.3. Bacterial lipoproteins	20
1.3.1. Lipoprotein biosynthesis	20
1.3.2. Lipoprotein localization	23
1.4. Methods in structural biology	25
1.5. Objectives of this dissertation	28

<b>Chapter II. Materials and methods</b>	<b>31</b>
2.1. Materials	32
2.2. Protein production	32
2.2.1. Cloning and expression of Cj0977	32
2.2.2. Cloning and expression of JlpA	34
2.2.3. Cloning and expression of novel lipoproteins	35
2.3. Protein purification	37
2.3.1. Bacterial cell lysis	38
2.3.2. Ammonium sulfate precipitation	38
2.3.3. His-tag affinity chromatography	38
2.3.4. GST affinity chromatography	39
2.3.5. Anion exchange chromatography	40
2.3.6. Gel filtration chromatography	40
2.4. Protein crystallization	41
2.4.1. Screening	41
2.4.2. Optimization and crystallization growth	43
2.5. Limited proteolysis assay	44
2.6. Phage display	44
2.6.1. Phage titering	46
2.6.2. Panning procedure	46
2.6.3. Binding phage M13 clone amplification	47
2.6.4. Rapid purification of sequencing templates	47
2.6.5. DNA sequencing and peptide analysis	48



2.7. Phage Enzyme-linked ImmunoSorbent Assay (ELISA)	48
<b>Chapter III. Structural Studies of the Virulence Protein Cj0977</b>	<b>50</b>
3.1. Introduction	51
3.2. Results	53
3.2.1. Purification and crystallization of His-Cj0977	53
3.2.2. Probing domain structures of Cj0977	55
3.2.3. Construct design of Cj0977 variants	57
3.2.4. Purification and crystallization of Cj0977 variants	61
3.3. Discussion	67
3.3.1. Cj0977 is a hot dog fold protein	67
3.3.2. Structural comparison and putative binding site for acyl-CoA	68
3.3.3. Proposed model of coupled folding and binding of Cj0977	72
<b>Chapter IV. Structural Studies of the Lipoprotein Adhesin JlpA</b>	<b>75</b>
4.1. Introduction	76
4.2. Results	78
4.2.1. Purification and crystallization of JlpA	78
4.2.2. Electron density map of JlpA crystals	81
4.2.3. Strategy for phase improvement: Introduction of methionine	81
4.2.4. Purification and crystallization of JlpA variants	86
4.3. Discussion	89

4.3.1. Molecular architecture of JlpA	89
4.3.2. Structure-function relationship of JlpA	92
<b>Chapter V. Searching for Novel Lipoproteins of <i>Campylobacter jejuni</i></b>	<b>96</b>
5.1. Introduction	97
5.2. Results	100
5.2.1. Identification of putative lipoproteins from <i>C. jejuni</i> genome	100
5.2.2. Production of soluble proteins: Cj0090, Cj1026c, Cj1090c, and Cj1649	103
5.2.3. Crystallization of Cj0090, Cj1090c, and Cj1649	109
5.2.4. Search for potential ligands for lipoproteins	111
5.2.4.1. Screening of small peptides by phage display experiment	111
5.2.4.2. Specificity of the potential ligands	113
5.3. Discussion	115
5.3.1. Crystal structure of Cj0090, a novel immunoglobulin fold	115
5.3.2. Biochemical assays of lipoproteins of <i>C. jejuni</i>	119
<b>Chapter VI. Concluding Remarks</b>	<b>122</b>
<b>References</b>	<b>128</b>

## List of Figures

<b>Figure 1-1.</b> <i>Campylobacter jejuni</i> and a model of <i>C. jejuni</i> flagellum structure.	3
<b>Figure 1-2.</b> Internalization and activity of cytolethal distending toxin (CDT).	16
<b>Figure 1-3.</b> <i>C. jejuni</i> lipooligosaccharides (LOSs) that mimic human gangliosides.	18
<b>Figure 1-4.</b> The biosynthetic pathway of bacterial lipoproteins.	22
<b>Figure 1-5.</b> Sorting and outer-membrane localization of lipoproteins by the Lol system.	24
<b>Figure 2-1.</b> Crystallization setup.	42
<b>Figure 2-2.</b> Schematic procedures of phage display.	45
<b>Figure 3-1.</b> Purification of His-Cj0977.	54
<b>Figure 3-2.</b> Crystals of N-terminal His-tagged full-length Cj0977 (His-Cj0977).	56
<b>Figure 3-3.</b> Limited proteolysis and stability test of His-Cj0977.	58
<b>Figure 3-4.</b> Sequence alignment and peptide mapping.	60
<b>Figure 3-5.</b> Purification of Cj0977 <sub>p19</sub> .	62
<b>Figure 3-6.</b> Purification of Cj0977 <sub>p17</sub> .	63
<b>Figure 3-7.</b> Analytical gel filtration chromatograph.	65
<b>Figure 3-8.</b> Crystals of smCj0977 <sub>p19</sub> and smCj0977 <sub>p17</sub> .	66
<b>Figure 3-9.</b> Crystal structure of Cj0977 <sub>p17</sub> .	69
<b>Figure 3-10.</b> Secondary structure of Cj0977 <sub>p17</sub> and structure-based sequence alignment.	71
<b>Figure 3-11.</b> Domain architecture and proposed model of coupled folding and binding of Cj0977.	74

<b>Figure 4-1.</b> Purification of His-JlpA.	80
<b>Figure 4-2.</b> JlpA crystals.	82
<b>Figure 4-3.</b> Initial electron density map of a JlpA crystal.	83
<b>Figure 4-4.</b> Sequence alignment of JlpA homologs.	85
<b>Figure 4-5.</b> smHis-JlpA <sub>L45M/I160M</sub> crystals.	88
<b>Figure 4-6.</b> Structure of JlpA.	90
<b>Figure 4-7.</b> JlpA and other bacterial lipoproteins.	93
<b>Figure 5-1.</b> Purification of His-Cj0090 <sub>16-122</sub> .	104
<b>Figure 5-2.</b> Purification of His-Cj1026c.	106
<b>Figure 5-3.</b> Purification of His-Cj1090c <sub>16-170</sub> .	107
<b>Figure 5-4.</b> Purification of His-Cj1649 <sub>18-199</sub> .	108
<b>Figure 5-5.</b> Crystal pictures of Cj0090, Cj1090c, and Cj1649.	110
<b>Figure 5-6.</b> Binding phage peptides.	114
<b>Figure 5-7.</b> Structure of Cj0090.	116
<b>Figure 5-8.</b> Topology diagrams of the Ig fold and comparison with the Cj0090 fold.	117
<b>Figure 5-9.</b> Topology diagrams of Cj0090 and close structural homologs with Ig fold variants.	120

## List of Tables

<b>Table 5-1.</b> Signal peptide sequence and lipobox motif of 20 putative lipoproteins from <i>C. jejuni</i> .	101
<b>Table 5-2.</b> List of 20 putative lipoproteins from <i>C. jejuni</i> NCTC 11168.	102
<b>Table 5-3.</b> Peptide identified by phage display.	112

## List of Abbreviations

ABC	ATP-binding cassette
AGG	Autoagglutination
Ala	Alanine
AMAN	Acute motor axonal neuropathy
ArHBT	<i>Arthrobacter</i> sp. 4-hydroxybenzoyl-CoA thioesterase
ASP	Ammonium sulfate precipitation
Asp	Aspartic acid
BME	$\beta$ -mercaptoethanol
BSA	Bovine serum albumin
BsFapR	<i>Bacillus subtilis</i> fatty acid & phospholipid biosynthesis regulator
CadF	<i>Campylobacter</i> adhesion to fibronectin
CapA	Campylobacter adhesion protein A
CDT (CLDT)	Cytotoxic distending toxin
CHO	Chinese hamster ovary
Cia	<i>Campylobacter</i> invasion antigen
CoA	Coenzyme A
CPS	Capsular polysaccharide
Cryo-EM	Cryo-electron microscopy
Cys	Cysteine
DIM	Phthiocerol dimycocerosates
DNase I	Deoxyribonuclease

DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FabA	$\beta$ -hydroxydecanoyl thiol ester dehydratase
FN3	Fibronectin type III
FspA	Flagella secreted protein A
FT	Flow through
GBS	Guillain-Barré syndrome
GEL	Gel filtration column
Glu	Glutamic acid
GST	Glutathione S-transferase
Hep	Heptosyltransferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HiTrapQ	HiTrapQ anionic exchange column
HMW	High-molecular-weight
HRP	Horseradish Peroxidase
Hsp	Heat shock protein
HTH	Helix-turn-helix
IEF	Isoelectric focusing
Ig	Immunoglobulin
JlpA	Jejuni lipoprotein A
kDa	Kilodalton

$K_d^{\text{app}}$	Apparent dissociation constants
KDO	2-keto-3-deoxy-octulosonic acid
LB media	Luria-Bertani media
Lgt	Pre-prolipoprotein diacylglycerol transferase
Lnt	Apolipoprotein N-acyl transferase
LolA	Localization of lipoprotein A
LolB	Localization of lipoprotein B
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
LspA	Prolipoprotein signal peptidase
LTB	The B subunit of heat-labile enterotoxin
Lys	Lysine
MAD	Multi-wavelength anomalous dispersion
MAP	Mitogen-activated protein
Met	Methionine
MFS	Miller Fisher syndrome
MIR	Multi-isomorphous replacement
MOMP	Major outer membrane protein
MPD	2-methyl-2,4-pentanediol
MR	Molecular replacement
Mr	Protein molecular weight marker
NaCl	Sodium chloride
NF- $\kappa$ B	Nuclear factor kappa-light-chain-enhancer of activated B cells



Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance spectroscopy
NspA	Neisserial surface protein A
NTA	Ni-NTA affinity column
OD	Optical density
Omp	Outer membrane protein
Opa	Opacity-associated
OPD	o-Phenylenediamine dihydrochloride
ORF	Open reading frames
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PEG	Polyethylene glycol
pfu	Plaque-forming unit
Phe	Phenylalanine
pI	Isoelectric point
PseAc	Pseudaminic acid
PseAm	An acetamidino form of PseAc
PseAmOGLnNAc	A form of PseAm with an N-acetyl glutamic acid attached
PseOAc	An acetylated form of PseAc
PVDF	Polyvinylidene difluoride
rpm	Revolutions per minute
SAD	Single wavelength anomalous dispersion
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

sm / SeMet	Selenomethionine
SPR	Surface plasmon resonance
TBS	Tris-buffered saline
TBST	Tris-Buffered Saline containg Tween 20
TLR	Toll-like receptor
TtPaal	<i>Thermus thermophilus</i> phenylacetate thioesterase
UV	Ultraviolet

## **Chapter I. Introduction**

### 1.1. *Campylobacter jejuni*

*Campylobacter jejuni*, a Gram-negative microaerophilic and spiral-shaped bacterium, is one of the most common bacterial causes of human gastroenteritis throughout the world (Poly and Guerry, 2008 and Zilbauer *et al.*, 2008) (Figure 1-1 A). Historically, awareness of the public health implications of *Campylobacter* infections has evolved over more than a century (Altekruse *et al.*, 1999). In 1886, Escherich observed organisms resembling campylobacters in stool samples of children with diarrhea, and in 1913, McFaydean and Stockman identified campylobacters (called related *Vibrio*) in fetal tissues of aborted sheep. While King described the isolation of related *Vibrio* from blood samples of children with diarrhea in 1957, clinical microbiologists in Belgium achieved the first isolation of campylobacters from stool samples of patients with diarrhea. In the 1970s, the development of selective growth media permitted more laboratories to test stool specimens for *Campylobacter*. Soon *Campylobacter* spp. were established as common human pathogens. This organism is now the leading cause of bacterial food poisoning (campylobacteriosis) worldwide, and is more prevalent than *Salmonella enteritis* (salmonellosis). *C. jejuni* is a member of the epsilon group of proteobacteria, along with the gastric pathogen *Helicobacter pylori* and a non-pathogenic bacterium *Wolinella succinogenes*. Since the first genome sequence of *C. jejuni* NCTC 11168 was reported, genome sequences of several other *C. jejuni* strains have been completed or are near completion (NCBI genome database) (Parkhill *et al.*, 2000). A noticeable characteristic in the genome was the presence of hypervariable sequence regions. These short sequences are mostly found in genes encoding the biosynthesis or modification of surface structures (e.g. lipooligosaccharides and flagella) or in closely linked genes of



unknown function. The apparent hypervariability of these regions may be important in the survival strategy of this organism (Parkhill *et al.*, 2000).

Clinical symptoms of human *C. jejuni* infection range from mild non-inflammatory diarrhea to severe inflammatory bloody diarrhea with abdominal pain, fever, nausea, and vomiting. *C. jejuni* inhabits the gastrointestinal tract of animal hosts such as poultry, other birds, cattle, and swine, and is considered to be a commensal organism. Consuming contaminated poultry is the major source of infection in human. In addition, unpasteurized milk and untreated drinking water are uncommon sources. The infection occurs more often in infants, young children, the elderly, and patients with weak immune systems than in the general population. When humans ingest *C. jejuni*-contaminated food, this bacterium survives the stress in the stomach and the small intestine, colonizes the small intestine in the early stage of infection, and then moves to the large intestine that is the target organ of the disease (Poly and Guerry, 2008).

In most immunocompetent individuals, campylobacteriosis is a self-limited illness and generally does not require antibiotic treatment. However, antimicrobial therapy can be introduced on very sick patients with *C. jejuni* infection to decrease the duration and the severity of illness. Macrolides and fluoroquinolones are commonly used for treatment of *Campylobacter* enteritis. However, antimicrobial resistance has become a major public health concern worldwide in recent years and more importantly, so has the rapid emergence of antibiotic-resistant *Campylobacter* strains (Moore *et al.*, 2006). In industrialized countries, the emergence of fluoroquinolone-resistant *C. jejuni* strains illustrates the need for prudent antibiotic use in food-animal production, mainly in poultry. Furthermore the increased use of fluoroquinolones in human and veterinary

medicine coincides with the increase of quinolone-resistant *C. jejuni* strains (Endtz *et al.*, 1991).

*C. jejuni* infections have been linked to two important sequelae: reactive arthritis and Guillain-Barré syndrome (GBS). GBS is an autoimmune-mediated demyelinating disorder of the peripheral nervous system causing paralysis and even death (Nachamkin *et al.*, 2002). The development of GBS is associated with lipooligosaccharides (LOSs) that mimic human gangliosides. Thus, antibodies to the LOS core structures result in an autoimmune response affecting the peripheral nervous system (Poly and Guerry, 2008). More recently, *C. jejuni* has also been associated with a rare form of mucosa-associated lymphoid tissue lymphoma called immunoproliferative small intestinal disease (also known as alpha chain disease) (Lecuit *et al.*, 2004). Although molecular and immunobiochemical studies have shown a link between these diseases and *C. jejuni* infections, further studies are necessary to determine the molecular mechanisms.

## **1.2. Virulence factors and pathogenesis of *C. jejuni***

Although *C. jejuni* is a major bacterial cause of human gastroenteritis and complete multiple genome sequences of *C. jejuni* are available, the pathogenic mechanism of *C. jejuni* is still poorly understood at the molecular level. It is partly due to the lack of a suitable animal model that reproduces the human disease and the lack of homologs of *C. jejuni* virulence factors from other pathogens. Here, recent advances in our understanding of the virulence and pathogenesis of *C. jejuni* are presented.

### 1.2.1. Flagella and flagella-related virulence factors

*C. jejuni* utilizes polar flagella for its rapid, darting motility. The *C. jejuni* flagella structures as well as its motility have long been recognized as crucial to pathogenesis. The *C. jejuni* motility is necessary for intestinal colonization in animal disease models and for the invasion of human intestinal cells *in vitro* (Morooka *et al.* 1985; Wassenaar *et al.*, 1991; and Nachamkin *et al.*, 1993).

The bacterial flagellum that is responsible for motility and chemotaxis is a lash-like appendage and protrudes from the cell surface. The flagellum is composed of three parts: the filament, the hook, and the basal body (Figure 1-1 B). The filament is composed of a single protein subunit component, flagellin. The hook is composed of FlgE and located at the base of the flagellum *via* the hook-filament junction formed from FlgK and FlgL. The basal body consists of a series of rings connected by a rod. The distal rod is made up of FlgF and FlgG, and the proximal rod is composed of FliE, FlgB, and FlgC. There are five rings in the flagellum structure: the L-, P-, S-, M-, and C-ring. The L-ring is bound to the lipopolysaccharide layer in the outer membrane and consists of FlgH. The P-ring is associated with the peptidoglycan layer and is composed of FlgI. The two rings, the S- and M-rings, are in close contact with the inner membrane. Both the S- and M-ring are made up of subunits of a single protein, FliF. The M-ring is embedded in the membrane, and the S-ring is above the membrane. The C-ring located in the inner membrane consists of three proteins: FliG, FliM, and FliN. The FliM and FliN proteins are components of the motor switch and enable rotation and determining its direction (Yamaguchi *et al.*, 1986). The flagellum is driven by motor complex proteins located in the basal body. It is powered by proton motive force.



The *Campylobacter* flagella show two unusual structural features. In general, the N- and C-terminal amino acids of bacterial flagellins are highly conserved and serve as recognition sites for export and for subunit interactions that are important for flagellar filament formation. These conserved regions of bacterial flagellins are recognized by Toll-like receptor (TLR) 5. The TLRs are a class of proteins that play an important role in pathogen recognition and initiation of inflammatory and immune responses (Brightbill *et al.*, 1999). One distinct structural feature is that the flagellins from *Campylobacter* spp. and the related bacterium *H. pylori* lack these TLR5-recognition sites. This structural feature of the flagellin appears to contribute to avoidance of the innate immune system. The other structural feature is unusually high glycosylation in the *Campylobacter* flagellin protein. Flagellin from *C. jejuni* strain 81-176 is glycosylated at 19 sites (Ser or Thr residues) with pseudaminic acid (PseAc) and its derivatives (Thibault *et al.*, 2001). This unusual 9-carbon sugar is similar structurally to sialic acid. In *C. jejuni* 81-176, the other major modification is an acetamidino form of PseAc, called PseAm. An acetylated form of PseAc (PseOAc) and a form of PseAm with an N-acetyl glutamic acid attached (PseAmOGlnNAc) are also observed as minor modifications (Thibault *et al.*, 2001 and Schirm *et al.*, 2005). The genes required for biosynthesis and/or transfer of PseAc and PseAm to flagellin have been identified (Thibault *et al.*, 2001; Goon *et al.*, 2003; and Guerry *et al.*, 2006). Interestingly, these genes are located adjacent to the flagellin structural genes and are among the most variable loci observed in the *C. jejuni* genomes. The function of the glycosyl modifications in the flagellar structure is not understood well. However, glycosylation is particularly important to the epsilon group of proteobacteria including *C. jejuni* and *H. pylori*, since neither organism can assemble a

filament in the absence of glycosylation. Also some *C. jejuni* mutants in flagellin glycosylation were unable to autoagglutinate (Golden and Acheson, 2002 and Guerry *et al.*, 2006). Autoagglutination (AGG) is often associated with virulence in other bacterial pathogens. A mutant of *C. jejuni* 81-176 that cannot synthesize PseAm was still able to assemble a flagella filament, but the glycosylation sites were glycosylated with PseAc instead of PseAm. This mutant was defective in AAG and showed a reduction in adherence and invasion of INT407 cells as well as attenuation in a ferret diarrheal disease model (Guerry *et al.*, 2006). These observations suggest that flagellin glycosylation has an important role in *C. jejuni* virulence.

In *C. jejuni*, flagella play complex roles in the pathogenesis other than just motility or chemotaxis, and are involved in secretion of non-flagella virulence proteins and the co-regulation of secreted and non-secreted virulence factors with the flagella regulon. In the absence of a specialized type III secretion system, which is utilized to secrete and inject virulence proteins into the cytosol of eukaryotic host cells in many other Gram-negative enteropathogens, the *C. jejuni* flagella function as specialized secretion systems to present several non-flagellar proteins (often virulence factors) to the extracellular space or host cells. Such proteins include Cia (Campylobacter invasion antigen) proteins (CiaB, CiaC, and CiaI), FlaC, and FspA (Konkel *et al.*, 1999; Song *et al.*, 2004; Poly *et al.*, 2007; Christensen *et al.*, 2009; and Buelow *et al.*, 2010). These secreted proteins are required for the *C. jejuni* invasion of human epithelial cells (Konkel *et al.*, 1999).

CiaB was the first protein reported as a non-flagella protein secreted through the *C. jejuni* flagellum (Konkel *et al.*, 1999) (Figure 1-1 B). *C. jejuni* *ciaB* null mutants were

fully motile but showed reduced internalization into host cells and were unable to secrete other Cia proteins *in vitro* suggesting an important role of CiaB in the secretion process. CiaC is required for maximal *C. jejuni* invasion of host cells *in vitro* (Christensen *et al.*, 2009). CiaC mutant showed reduced invasion *in vitro* when compared with the wild type. CiaI appears to play a role in enhancing the *C. jejuni* survival within human epithelial cells (Buelow *et al.*, 2010). Secretion of the Cia proteins depends on a minimum flagella structure. The *flhB* mutants, which are defective in export, or mutants in *flgB*, *flgC*, and *flgE*, which are defective in components of the flagellar basal body and hook, were unable to secrete Cia proteins.

FlaC, another non-flagella-secreted protein, is a flagellin-like protein but not a part of the flagellar structure. The secreted FlaC protein binds to human epithelial cells *in vitro*, and invasion of human epithelial cells by a *flaC* null mutant is reduced compared to wild type, suggesting an important role of FlaC in host cell invasion (Song *et al.*, 2004). Secretion of the FlaC also depends on a minimum flagella structure. The mutants in the flagellar hook protein FlgE, the flagellar basal-body rod protein FlgF, flagellar P-ring protein FlgI, or flagellar hook-associated protein FlgK were unable to secrete FlaC (Song *et al.*, 2004 and Poly *et al.*, 2007).

FspA (Flagella-secreted protein A), is a small, acidic protein that is expressed by a  $\sigma^{28}$  promoter (see below) and appear to be a *C. jejuni* specific protein. Analysis of the *fspA* gene in 41 isolates of *C. jejuni* revealed two overall variants of the predicted protein, FspA1 and FspA2. The incubation of INT407 cells with recombinant FspA2, but not FspA1, *in vitro* resulted in a rapid induction of apoptosis of INT407 cells (Poly *et al.*, 2007). Apparently, mutation in *fspA* had no effect in invasion or adhesion. The observed

heterogeneity among *fspA* alleles suggests alternate virulence potential among different strains. Secretion of FspA also requires a minimum flagellar structure, as mutants in *flgF*, *flgE*, *flgI*, or *flgK* failed to secrete FspA *in vitro* (Song *et al.*, 2004 and Poly *et al.*, 2007).

A large amount of energy is required to produce flagella in bacteria, and thus regulation of flagellar structural gene expression is important. Bacterial gene expression is regulated by numerous sigma factors. The number of sigma factors utilized by a given species of bacteria is highly variable. In contrast to *Bacillus subtilis* with 14 sigma factors or *Escherichia coli* with 7 sigma factors, the *C. jejuni* genome contains only three:  $\sigma^{28}$ ,  $\sigma^{54}$ , and  $\sigma^{70}$  (Parkhill *et al.*, 2000 and Carrillo *et al.*, 2004), suggesting that certain pathways in this organism may be coordinately regulated. Indeed, *Campylobacter* flagellar genes can be classified into three groups based on their promoters. The transcription factor  $\sigma^{70}$  is responsible for regulation of housekeeping genes including several early genes (class I) required for the flagellar basal body assembly. Most structural genes (class II) coding for the components of the basal body, hook, and minor flagellin (FlaB) are controlled by  $\sigma^{54}$  and the major flagellin (FlaA) is regulated by  $\sigma^{28}$  (Guerry *et al.*, 1991; Hendrixson *et al.*, 2001; Hendrixson and DiRita, 2003; Carrillo *et al.*, 2004; Wösten *et al.*, 2004; and Goon *et al.*, 2006). Interestingly, recent microarray-based studies also identified numerous non-flagellar genes that are regulated by  $\sigma^{54}$  and  $\sigma^{28}$ . For example, the *cj1026c* (one of our target proteins; see Chapter V) and *cj0428* genes were expressed by  $\sigma^{54}$  promoter, and the *cj0977* (see Chapter III) and *fspA1* genes were expressed by  $\sigma^{28}$  promoter (Goon *et al.*, 2006 and Poly *et al.*, 2007). The predicted proteins encoded by these genes are usually small (<29 kDa) and the functions of most these proteins are unknown. Given the strong connection of flagella and virulence, some

of these unknown genes were hypothesized as novel virulence factors that were co-regulated with flagella. More recently, studies on two of these proteins have confirmed their association with virulence. One is FspA that is involved in induction of apoptosis of intestinal epithelial cells, as described above. The other is the Cj0977 protein. A Cj0977 mutant of *C. jejuni* resulted in >3 logs reduced invasion of intestinal epithelial cells *in vitro* compared with the wild type strain. This mutant was also attenuated in an animal disease model (Goon *et al.*, 2006). The expression of Cj0977 is dependent on a minimal flagella structure, consistent with regulation by  $\sigma^{28}$ . Unlike FspA, however, Cj0977 was not secreted through the flagella and its function remains unknown (Figure 1-1 B).

### **1.2.2. Adhesins**

Adherence of *C. jejuni* to host epithelial cells is a critical step in *Campylobacter* infection. This process appears to require multiple binding factors to achieve an efficient interaction with host cells. A number of proteins, such as CadF, CapA, JlpA, and PEB1, have been described to bind cultured epithelial cells, but the relative significance and contribution of each to disease remains unknown.

CadF (*Campylobacter* adhesion to fibronectin), a 37 kDa protein, specifically binds to fibronectin, a component of the extracellular matrix of mammalian cells (Konkel *et al.*, 1997). Fibronectin has been reported as a common host cell target of a number of other pathogens, including *Staphylococcus aureus*, *Streptococcus pyogenes*, *Salmonella enteritidis*, *Escherichia coli*, *Neisseria gonorrhoeae*, *Mycobacterium avium*, and *Treponema* spp. (Kuusela, 1978; Myhre and Kuusela, 1983; Rydén *et al.*, 1983; Fröman *et al.*, 1984; Baloda *et al.*, 1985; Thomas *et al.*, 1985; Dawson and Ellen, 1990 and 1994; Visai *et al.*, 1991; Jaffe *et al.*, 1996; Schorey *et al.*, 1996; and van Putten *et al.*,

1998). CadF is required for maximal adherence and invasion by *C. jejuni* *in vitro* as a *cadF* mutant showed a reduced adhesion to human INT407 cells compared to a wild-type (Monteville *et al.*, 2003). CadF also appears to be required for the chick colonization as *cadF* mutants showed reduced chick colonization compared with the wild-type (Ziprin *et al.*, 1999).

CapA, a putative autotransporter, was recently identified *in silico* from the genome sequence of *C. jejuni*. Subsequently, biochemical characterization of this large protein (~116 kDa) demonstrated that CapA is an autotransporter protein of *C. jejuni* and localized on the cell surface. A *capA* insertion mutant had a significantly reduced adherence to Caco-2 cells and invasion *in vitro* and failed to colonize and persist in a chick model system *in vivo* (Ashgar *et al.*, 2007). Thus, CapA appears to play a role in host association and colonization by *Campylobacter*.

The PEB1 protein (28kDa) was identified as an antigenic factor and adhesin exposed on the cell surface *C. jejuni* (Kervella *et al.*, 1993 and Pei and Blaser, 1993). Disruption of the *peb1A* gene encoding PEB1 showed reduced *C. jejuni* adherence to human HeLa cells (Pei *et al.*, 1998). Mutation in the *peb1A* locus of *C. jejuni* reduced interactions with epithelial cells and intestinal colonization in a mouse model. PEB1 is also the periplasmic-binding protein component of an aspartate/glutamate ATP binding cassette (ABC) transporter essential for optimal microaerobic growth on these dicarboxylic amino acids. Recent structure-function studies provided a molecular basis for both the solute transport and adhesin/virulence functions of PEB1 (Müller *et al.*, 2007).

JlpA is a bacterial cell surface-exposed lipoprotein and plays a role in HEp-2 cell binding (Jin *et al.*, 2001). Disruption of *jlpA* gene showed reduced adherence of *C. jejuni* to human HEp-2 epithelial cells by about 19% compared with the wild-type strain. Moreover, JlpA was reported to interact with HEp-2 cell surface heat shock protein (Hsp) 90 $\alpha$  and initiate signaling pathways leading to the activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF- $\kappa$ B) and p38 mitogen-activated protein (MAP) kinase, contributing to proinflammatory responses in host cells (Jin *et al.*, 2003). A recent study also showed that JlpA is a glycoprotein and is immunogenic (Scott *et al.*, 2009). The JlpA protein is the main topic of this thesis, and we will discuss it fully in Chapter IV.

### **1.2.3. Cytolethal distending toxin (CDT)**

*C. jejuni* produces cytolethal distending toxin (CDT) that represents the best characterized virulence factor of the organism. CDTs are a family of heterotrimeric toxins produced by many Gram-negative bacteria including *C. lariidis*, *C. coli*, *C. fetus*, as well as many other bacterial species, including *E. coli*, *Actinobacillus actinomycetemcomitans*, *Shigella dysenteriae*, *Haemophilus ducreyi*, and *Helicobacter hepaticus* (Johnson and Lior, 1988). As mentioned earlier, except for CDT, *C. jejuni* lacks homologous of virulence factors that are common to other pathogens, which in part hinders progress in understanding the pathogenesis. The completed genome sequence of *C. jejuni* NCTC11168 in 2000 has revealed the three closely linked genes encoding the CDT (*cdtA*, *cdtB*, and *cdtC*), which are the only toxin genes present in *C. jejuni* (Parkhill *et al.*, 2000).

CDT is a tripartite toxin and CDT activity requires the expression of all three genes. CdtA, CdtB, and CdtC interact with one another to form an active tripartite holotoxin that exhibit full cellular toxicity (Lara-Tejero and Galán, 2001). CdtB is the enzymatically active subunit and CdtA and CdtC form the heterodimeric subunit required for the delivery of CdtB (Lara-Tejero and Galán, 2001). CdtB exhibits activity similar to the enzyme deoxyribonuclease (DNase I) as shown by microinjection of CdtB alone into host cells, which induces cytoplasmic distention and cell cycle arrest (Lara-Tejero and Galán, 2000). CdtB mutants in DNase-specific active site showed reduced DNase activity and failed to induce cellular distension or arrest division of HeLa cells (Elwell and Dreyfus, 2000).

*C. jejuni* CDT induces cell distension in various mammalian cell lines such as Chinese hamster ovary (CHO) cells, Vero, HeLa, and HEp-2, and Caco-2 cells and causes eventual cell death (Johnson and Lior, 1988 and Whitehouse *et al.*, 1998). However, not all types are responsive to CDT, since mouse Y-1 adrenal cells show no effect for CDT treatment (Johnson and Lior, 1988).

CDT blocks cell division by arresting at the G2/M transition of the eukaryotic cell cycle (Whitehouse *et al.*, 1998 and Cortes-Bratti *et al.*, 2001). The role of CDT in diarrheal disease caused by *C. jejuni* remains unclear, but it has been proposed that cell cycle arrest of intestinal epithelial cells would increase the contact time between *C. jejuni* and host cells. The function of CdtA and CdtC appears to mediate binding to host cells. Indeed, CdtA has a domain that shares similarity to the B chain of ricin toxin that is responsible for receptor-mediated endocytosis of ricin (Lara-Tejero and Galán, 2001). Moreover, both CdtA and CdtC bind to the surface of HeLa cells with specificity, linking

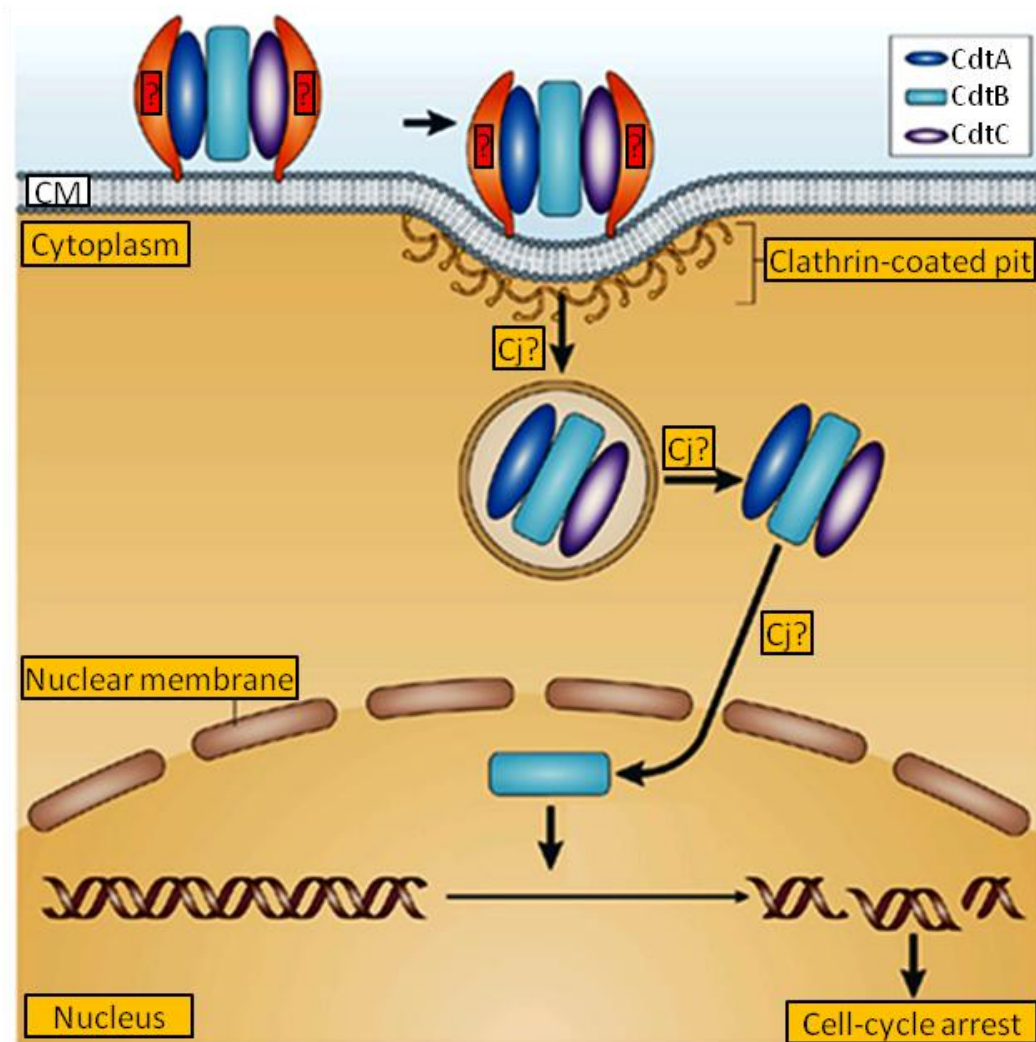


using the same receptor (Lee *et al.*, 2003). As it has been demonstrated for *H. ducreyi* that CDT enters the host cells via clathrin-coated pits, it may be possible that *C. jejuni* CdtA and CdtC mediate binding and translocation of CdtB into the host cells through a similar pathway (Cortes-Bratti *et al.*, 2000) (Figure 1-2).

#### **1.2.4. Capsular polysaccharide (CPS)**

*C. jejuni* is unusual for an intestinal pathogen in that its surface is coated with a polysaccharide capsule. Capsular polysaccharide (CPS), an important component of the outer membrane, mimics host cell antigens and undergoes structural variation. The phase-variable structural modifications include the incorporation of methyl, ethanolamine, and aminoglycerol groups on CPS sugars (Szymanski *et al.*, 2003; McNally *et al.*, 2005; and McNally *et al.*, 2007). CPS is considered as a virulence factor, since CPS-conjugated vaccines have protected against diarrheal disease in primate models. Interestingly, the capsule locus is one of the most variable loci among *C. jejuni* strains.

Many strains of *C. jejuni* have long been thought to produce a high-molecular-weight lipopolysaccharide (HMW LPS, also referred to as O-antigen), but it is now recognized as CPS (Karlyshev *et al.*, 2000). The presence of CPS on the membrane surface of *C. jejuni* was demonstrated using electron microscopy (Karlyshev *et al.*, 2001). The structures of several CPS types of *C. jejuni* have been reported to date (Aspinall *et al.*, 1992; Aspinall *et al.*, 1995; Hanniffy *et al.*, 1999; Muldoon *et al.*, 2002; Karlyshev *et al.*, 2005; McNally *et al.*, 2005; Gilbert *et al.*, 2007; McNally *et al.*, 2007; and Chen *et al.*, 2008). These types vary in sugar composition and linkage. The presence of heptosyl residues and phosphoramidate side chains is the structural features of the capsules of *C. jejuni*.



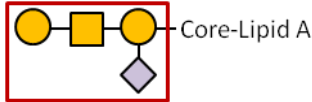
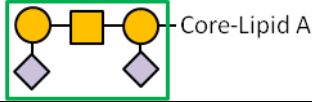
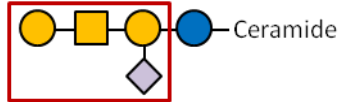
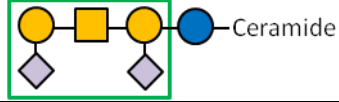
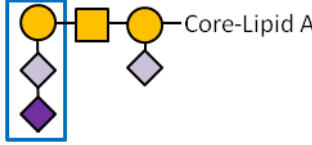
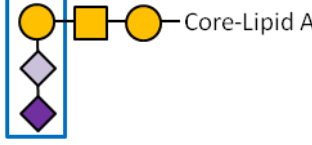
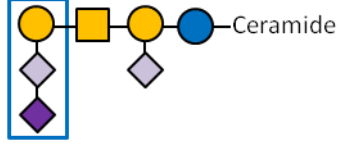
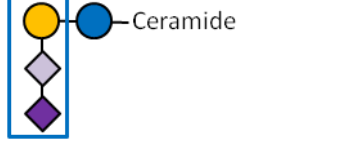
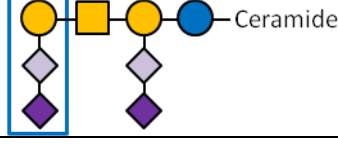
**Figure 1-2. Internalization and activity of cytolethal distending toxin (CDT).** The active cytolethal distending toxin (CDT) is composed of three subunits, CdtA, CdtB, and CdtC. CDT attaches to the host cell membrane via an unknown receptor. CDT may enter the host cell by way of clathrin-coated pits. Following internalization, CdtB enters the nucleus via the typical nuclear import cycle and acts as a DNase causing double-strand DNA breaks and leads to cell-cycle arrest. The arrows marked 'Cj?' indicate the aspects of CDT internalization or activity that have been studied for CDT of bacteria other than *C. jejuni* (adapted from Young *et al.*, 2007).

Generally, CPS is thought to be important for bacterial survival and persistence in the environment and to play a critical role in pathogenesis for *C. jejuni* (Roberts *et al.*, 1996). CPS mutants in *C. jejuni* showed the important roles of CPS in human serum resistance, adherence and invasion of epithelial cells *in vitro*, and virulence in a ferret diarrhoeal model (Bacon *et al.*, 2001 and Bachtiar *et al.*, 2007).

#### **1.2.5. Lipooligosaccharide (LOS)**

Lipooligosaccharide (LOS) is one of the major cell-surface structures of *C. jejuni* and considered as an important virulence factor. LOS is also observed in other Gram-negative pathogens such as *Neisseria spp.*, *Haemophilus spp.* and *Bordetella spp.* (Jennings *et al.*, 1980; Inzana *et al.*, 1985; and Ray *et al.*, 1991). While LPS (lipopolysaccharide) comprises three major parts including lipid A, core oligosaccharide, and O-antigen, LOS lacks O-antigen. The core oligosaccharide consists of a conserved inner core and a highly variable outer core (Dorrell *et al.*, 2001 and Parker *et al.*, 2008). The core is attached via 2-keto-3-deoxy-octulosonic acid (KDO) molecule to the lipid A, which is anchored to the outer membrane (Naess and Hofstad, 1984 and Beer *et al.*, 1986).

LOS is important for adherence and invasion of host cells and undergoes phase variation (Moran *et al.*, 1996 and Guerry *et al.*, 2002). The outer LOS cores of *C. jejuni* strains are hypervariable due to variations in sugar linkage and composition (Figure 1-3). The outer core structures of LOS from certain strains of *C. jejuni* resemble human gangliosides, a group of glycosphingolipids predominantly distributed in the surface membrane of nerve cells and anchored to the external leaflet of the lipid bilayer by a

<i>C. jejuni</i> <i>cst-II</i>	LOS structures produced by <i>C. jejuni</i>	Human gangliosides
<p>Sialyltransferase Cst-II (Thr51) activity</p> <p>monofunctional (<math>\alpha</math>-2,3-sialyltransferase)</p>	<p>GM1-like</p>  <p>Core-Lipid A</p> <p>GD1a-like</p>  <p>Core-Lipid A</p>	<p>GM1</p>  <p>Ceramide</p> <p>GD1a</p>  <p>Ceramide</p>
<p>Sialyltransferase Cst-II (Asn51) activity</p> <p>bifunctional (<math>\alpha</math>-2,3-sialyltransferase and <math>\alpha</math>-2,8-sialyltransferase)</p>	<p>GT1a-like</p>  <p>Core-Lipid A</p> <p>GD1c-like</p>  <p>Core-Lipid A</p>	<p>GT1a</p>  <p>Ceramide</p> <p>GD3</p>  <p>Ceramide</p> <p>GQ1b</p>  <p>Ceramide</p>

**Figure 1-3. *C. jejuni* lipooligosaccharides (LOSs) that mimic human gangliosides.** *C. jejuni* expresses oligosaccharide structures, which protrude from the LOS core. The terminal structure of GM1-like LOS is identical to that of GM1 (shown in red box) and GD1a-like LOS to that of GD1a (shown in green box). The terminal disialosyl group in GT1a is identical to that of GD3, GQ1b, GT1a-like LOS, and GD1c-like LOS. Variation in the nucleotide sequence at 51 of *cst-II* (Thr or Asn) seems to affect enzymatic activity resulting in four different human ganglioside mimics ●, Glucose; ●, Galactose; ■, *N*-acetyl-galactosamine; ◆,  $\alpha$  2,3 *N*-acetyl-neuraminic acid; and ◆,  $\alpha$  2,8 *N*-acetyl-neuraminic acid (adapted from Yuki, 2005 and Rinaldi and Willison, 2008).

ceramide moiety, thus resulting in ganglioside mimicry that can lead to the development of neurological complications including Guillain-Barré syndrome (GBS), Miller Fisher syndrome (MFS), and acute motor axonal neuropathy (AMAN) (Figure 1-3) (Ang *et al.*, 2001; Misawa *et al.*, 2001; Nachamkin *et al.*, 2002; and Yuki, 2005). In addition, the presence of sialylated LOS cores may be a mechanism of immune avoidance, since loss of sialic acid from the core resulted in increased immunogenicity and increased sensitivity to killing by normal human serum (Poly and Guerry, 2008). Human ganglioside-like LOS is synthesized by sialyltransferase Cst-II. Variation in the nucleotide sequence of *cst-II* may affect enzymatic activity; Cst-II (Thr51) has only  $\alpha$ -2,3-sialyltransferase activity (monofunctional) and can produce GM1-like and GD1a-like LOSs, whereas Cst-II (Asn51) has both  $\alpha$ -2,3- and  $\alpha$ -2,8-sialyltransferase activities (bifunctional) and can produce GT1a-like and GD1c-like LOSs that mimic terminal disialosyl groups (Figure 1-3).

The genes involved in LOS biosynthesis of *C. jejuni* have been identified. Inactivation of *galE* (encoding UDP-glucose 4-epimerase) or *waaF* (encoding heptosyltransferase II) resulted in organisms expressing truncated LOS and with reduced invasion of epithelial cells *in vitro* (Fry *et al.*, 2000). Both *waaF* and *lgtF* (encoding a two-domain glucosyltransferase responsible for the addition of  $\beta$ -1,4 glucose to heptosyltransferase I (HepI) and  $\beta$ -1,2 glucose to HepII) mutants of *C. jejuni* 81-176 expressed truncated LOS and resulted in impaired colonization in a animal disease model (Kanipes *et al.*, 2008 and Naito *et al.*, 2010). In addition, complete deletion of the LOS biosynthesis locus in *C. jejuni* NCTC11168 resulted in attenuated growth and increased sensitivity to antibiotics and detergents (Marsden *et al.*, 2009).

### 1.3. Bacterial lipoproteins

Bacterial lipoproteins are universal components of the bacterial membranes and play various roles in the cell. One study estimated that typically up to 3% of bacterial genomes encode lipoproteins (Babu *et al.*, 2006). There is also a renewed interest in lipoproteins due to their roles in bacterial pathogenesis. Although the majority of lipoproteins remain putative, characterized lipoproteins have been involved in formation and stabilization of the cell surface structure, nutrient uptake, transmembrane signaling, adhesion, cell growth, conjugation, sporulation, antibiotic resistance, substrate transport, and extracytoplasmic folding of proteins (Lampen and Nielsen, 1984; Mathiopoulou *et al.*, 1991; Perego *et al.*, 1991; Alloing *et al.*, 1994; and Sutcliffe and Russell, 1995). In pathogenic bacteria, lipoproteins play a role in virulence-associated functions such as colonization, invasion, and evasion of host defense (Jenkinson *et al.*, 1994; Khandavilli *et al.*, 2008; and Hutchings *et al.*, 2009).

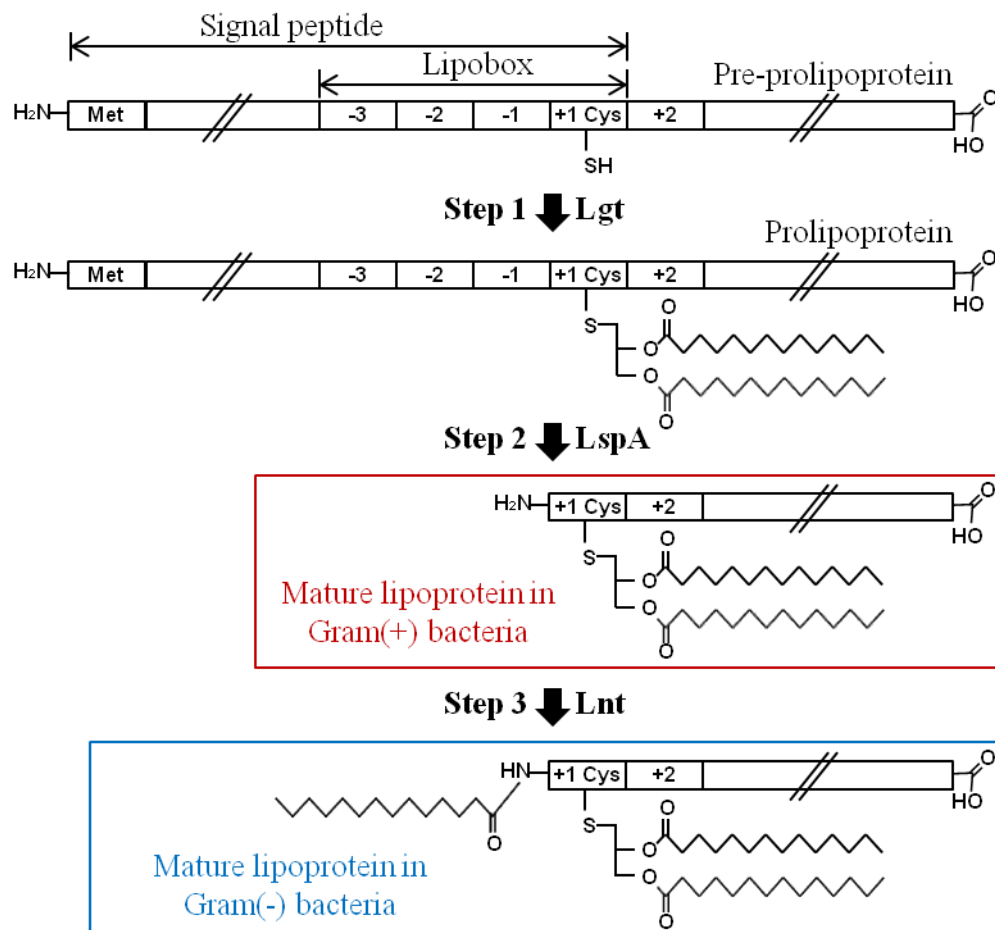
#### 1.3.1. Lipoprotein biosynthesis

Lipoproteins are characterized by the presence of an N-terminal signal peptide and a conserved consensus sequence called a lipobox, located in the C-terminal end of the signal peptide sequence. The lipobox sequence motif is essential for correct lipoprotein processing and is typically four amino acids, [LIVMFWSTAG]<sub>-3</sub>-[LIVMFYSTAGCQ]<sub>-2</sub>-[AGS]<sub>-1</sub>-C<sub>+1</sub>, in which the +1 Cys residue is universally conserved and is targeted for lipid modification (Taylor *et al.*, 2006). The lipobox motif is recognized by the cytoplasmic membrane prolipoprotein modification and processing enzymes that lead to the formation of N-acyl-diacylglycerylcysteine, in which a diacylglycerol moiety is covalently attached to the thiol group on the side chain of the N-

terminal Cys residue. The lipid-modified Cys residue serves to anchor the lipoprotein to the membrane and is the first amino acid of the mature lipoprotein. The mature lipoprotein is then translocated to its final destination according to its sorting signal.

Lipoproteins are initially translated in the cytoplasm as pre-prolipoproteins that possess an N-terminal signal peptide containing a conserved lipobox sequence. Pre-prolipoproteins are translocated across the inner membrane via a SecA-dependent pathway. Then the lipid modification of the pre-prolipoprotein to the mature lipoproteins is carried out by three enzymatic steps on the periplasmic side of the inner membrane (Kovacs-Simon *et al.*, 2011) (Figure 1-4). The N-terminal Cys residue lipid-modification is processed by pre-prolipoprotein diacylglycerol transferase (Lgt) resulting in a prolipoprotein (Step 1). In this process, a diacylglyceryl moiety is attached by a thioether linkage to the lipobox Cys residue. Then prolipoprotein signal peptidase (LspA) cleaves the signal sequence of the prolipoprotein (Step 2). The genes of *lgt* and *lspA* are present in a wide range of both Gram-positive and Gram-negative bacteria. In Gram-negative, the cleaved prolipoprotein is further processed by aminoacylation of the N-terminal Cys by apolipoprotein N-acyl transferase (Lnt) (Step 3). Thus, the mature lipoprotein possesses three acyl chains linked to its N-terminal Cys. Both the diacylglyceryl group and the amino terminal acyl group are derived from membrane phospholipids and provide anchoring of the lipoprotein to the membrane (Hantke and Braun, 1973).

In many Gram-negative bacteria, the three lipoprotein biosynthetic enzymes (Lgt, LspA, and Lnt) are conserved and essential for viability. Mutations in these enzymes are lethal in Gram-negative bacteria. On the other hand, in many Gram-positive bacteria, Lnt homologs have not been found and mutations in enzymes (Lgt and LspA) cause only



**Figure 1-4. The biosynthetic pathway of bacterial lipoproteins.** Two-step biosynthetic pathway is observed in Gram-positive bacteria (Step 1 to Step 2). Three-step biosynthetic pathway is shown in Gram-negative bacteria (Step 1 to Step 3). The first enzyme, pre-prolipoprotein diacylglyceryl transferase (Lgt), transfers a diacylglycerol moiety from a membrane phospholipid to the sulfhydryl group of +1 Cys of the conserved lipobox motif generating a thioether linkage. The second enzyme, prolipoprotein signal peptidase (LspA), cleaves the signal peptide at the N-terminus of the +1 S-diacylglyceryl Cys of the prolipoprotein. In Gram-positive bacteria, LspA cleaves the signal peptide, leaving Cys as the new N-terminus of the mature lipoprotein. In Gram-negative and some Gram-positive bacteria, apolipoprotein N-acyltransferase (Lnt), transfers an acyl group from another phospholipid to the newly-generated  $\alpha$ -amino group of the S-diacylglyceryl Cys, forming a mature triacylated lipoprotein. (adapted from Kovacs-Simon *et al.*, 2011).

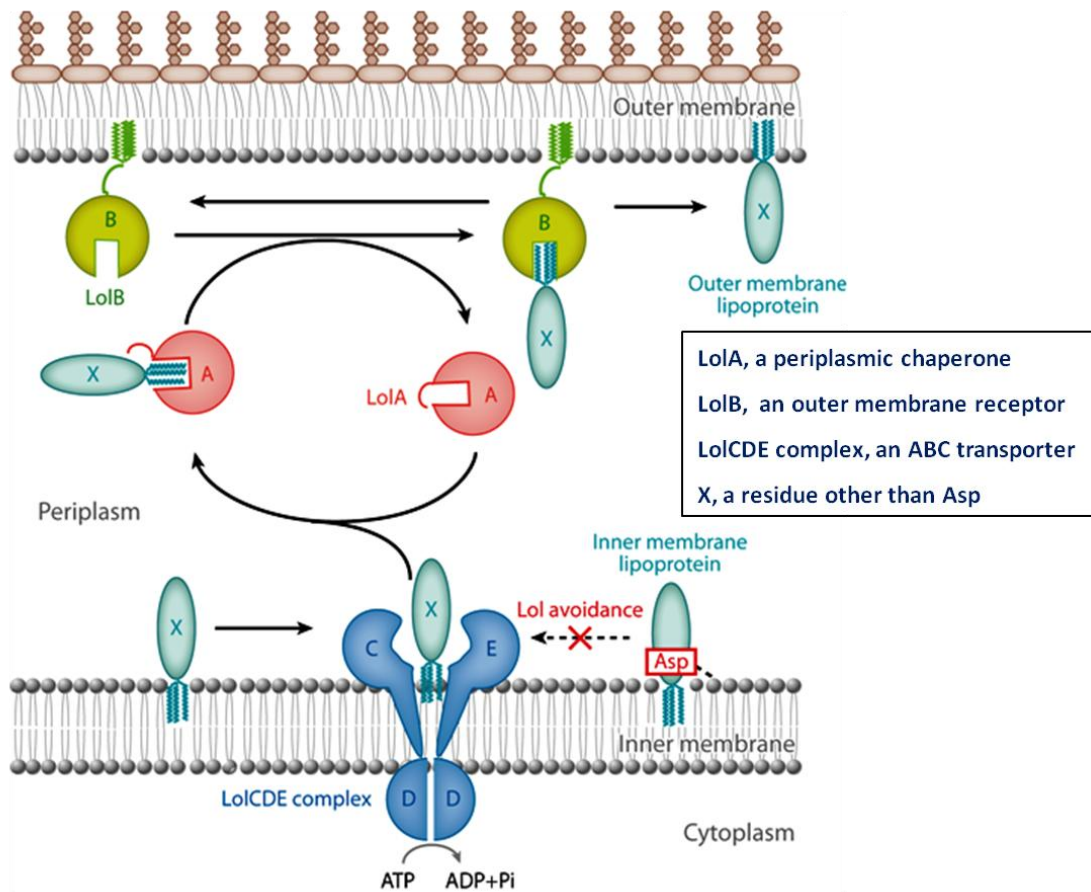


slight growth defects (Petit *et al.*, 2001; Hamilton *et al.*, 2006; and Baumgartner *et al.*, 2007), but the presence of lipoproteins is necessary for virulence (Tidhar *et al.*, 2009). Amino acid residues at position +2, +3, and +4 are responsible for designating the membrane localization of Gram-negative bacterial lipoproteins.

### **1.3.2. Lipoprotein localization**

After the processing of lipoproteins, the lipoproteins are translocated to their final destinations. In Gram-negative bacteria, mature lipoproteins are localized to the periplasmic side of the inner membrane or the outer membrane or they are localized to the outer surface of the outer membrane. Lipoproteins in Gram-positive bacteria are localized on the outer surface of the cytoplasmic membrane (Tokuda *et al.*, 2009 and Okuda and Tokuda, 2011). Lipoprotein sorting has been studied in detail in the Gram-negative *E. coli* system. *E. coli* has over 90 lipoproteins and the majority of these are located at the inner surface of the outer membrane, while others are present at the periplasmic side of the inner membrane (Narita *et al.*, 2004). In *E. coli*, the mature lipoproteins destined for the outer membrane are translocated by the lipoprotein localization (Lol) transport system, which is composed of five Lol proteins, Lol ABCDE. The Lol machinery consists of a transmembrane protein complex (LolCDE), a periplasmic chaperone (LolA), and an outer-membrane receptor (LolB) (Tokuda *et al.*, 2009 and Okuda and Tokuda, 2011) (Figure 1-5). Since Lol proteins are widely present in various Gram-negative bacteria, the Lol pathway found in *E. coli* has been served as the prototype of the lipoprotein sorting system.

The N-terminal second residue of lipoproteins is important for localization of lipoproteins. Asp at position 2 acts as the LolCDE avoidance signal. It prevents the



**Figure 1-5. Sorting and outer-membrane localization of lipoproteins by the Lol system.** The outer membrane-specific lipoproteins are transported from the inner membrane to the outer membrane by the Localization of lipoproteins (Lol) system, which is composed of LolA, LolB, and LolCDE complex. The outer membrane-specific lipoproteins are released from the inner membrane by LolCDE complex leading to the formation of a hydrophilic complex between the released lipoprotein and LolA. The lipoprotein is transferred from LolA to LolB and then incorporated into the outer membrane. The inner membrane retention signal Asp at position 2 causes the lipoprotein to retain on the inner membrane (adapted from Okuda and Tokuda, 2011).

recognition of lipoproteins by LolCDE and causes retention of lipoproteins in the inner membrane (Figure 1-5). Substitution with a different amino acid at this position 2 results in translocation by the Lol machinery to the outer membrane. It was reported that substitution of Ser at position 2 of an outer membrane-specific lipoprotein by Asp caused it to remain in the inner membrane. In addition, an inner membrane-specific lipoprotein was mislocalized to the outer membrane when another residue was substituted for Asp at position 2. These results indicate that Asp at position 2 is the inner membrane retention signal (Yamaguchi *et al.*, 1988).

#### **1.4. Methods in structural biology**

As essential components of organisms, proteins carry out numerous functions in the cells. Three-dimensional structures of proteins can give insight in the molecular mechanisms underlying important processes in life. Over many decades of research, effective methods have been developed to visualize protein structures on an atomic scale. These methods include X-ray crystallography, atomic force microscopy, nuclear magnetic resonance spectroscopy (NMR), and cryo-electron microscopy (cryo-EM). In recent years, purely computational methods have been advanced to predict protein structures. Especially, X-ray crystallography and NMR represent the most common methods to provide structure information at the atomic level of resolution (Guerry and Herrmann, 2011 and Yonath, 2011). While X-ray crystallography can be applied to any sizes of proteins including large molecules (routinely, >100 kDa) as long as suitable crystals are obtained, NMR spectroscopy is limited to relatively small and soluble molecules (typically, <40 kDa). However, NMR can provide information about the

dynamics of proteins and flexible regions or unstructured domain of proteins. Thus, these two methods are often complementary in structural studies of macromolecules (Snyder *et al.*, 2005; Yee *et al.*, 2005; and Feng, 2011).

The method used in our laboratory for protein structure determination is protein X-ray crystallography. Although the first protein structures solved by this method date from the 1960s (Kendrew *et al.*, 1960 and Perutz *et al.*, 1960), advances in the last two decades have revolutionized structure determination of proteins and protein complexes in ways that were not possible before. Major improvements have been achieved in the expression and purification of proteins and crystallographic methods. Automation and high throughput, as found in structural genomics, may allow rapid structure determination of all possible proteins folds (Chandonia and Brenner, 2006).

The process of structure determination begins with obtaining large amounts of pure, monodispered, and highly concentrated protein samples. Recombinant expression in *E. coli* can yield high amounts of proteins, as long as any post-translational modifications (e.g. glycosylation) do not interfere with the protein production. Purification is generally done by chromatographic methods, typically after producing recombinant proteins with a high affinity tag like a His-tag. The highly concentrated and highly pure (>95%) protein samples are used for crystallization experiments.

Crystallization of a protein can be viewed as a controlled aggregation process. Among many different protocols, the most common is the hanging drop vapor-diffusion method, in which a drop of protein sample is mixed with an equal volume of precipitant solution and placed above a reservoir precipitant solution. With time, by vapor diffusion, the concentration of the protein as well as the precipitant in the drop will slowly increase,

causing the protein to precipitate in some cases. Precipitation in an ordered manner can yield protein crystals. Yet, it is impossible to predict the crystallization condition(s) for a protein. Therefore, many different conditions should be screened and examined by trial-and-error, varying parameters including protein concentration (as well as different protein variants), pH, temperature, precipitant, additives, *etc.*

The protein crystal consists of many protein molecules forming a periodic structure of unit cells (crystal lattice). In X-ray diffraction experiments, the electrons of the crystal interact with the X-rays, causing them to scatter. The periodicity of the crystal lattice causes X-rays only in defined directions producing a diffraction pattern of regularly spaced spots. To reconstruct the electron cloud in the unit cell, the structure factor (amplitude) and the phase are necessary. While the amplitude can be derived from the intensities measured in the X-ray experiment, the phase cannot be measured. The phase is, however, essential for the calculation of the electron density. This is the so-called the crystallographic phase problem. The phase problem, loss of phase information, can be solved by several ways including multi-wavelength anomalous diffraction (MAD), single-wavelength anomalous diffraction (SAD), multi-isomorphous replacement (MIR), and molecular replacement (MR). Especially, the ability to incorporate the heavy atom selenium (as selenomethionine) in the methionine sites of proteins during the bacterial cell culture has greatly facilitated solution of the phase problem. During this project, we have produced and explored selenomethionine-labelled proteins whenever possible (as long as proteins contained Met residues).

The bottleneck step of protein X-ray crystallography is to grow sizable diffraction quality protein crystals, from which high-resolution (at least 3Å) diffraction data can be

obtained. To obtain high quality crystals (well diffracting crystals), many different approaches can be employed. Certain chemical compounds or small molecules may have drastic effects on crystallization of a protein. Hence an effective additive identified from additive screenings can be added in the crystallization solution to improve crystallization. Proteins in complex with their binding partners (or ligands), if known, appear to crystallize better than the proteins alone. (Ostetmeier *et al.*, 1995; Hunte and Michel, 2002; Shore *et al.*, 2006; Shore *et al.*, 2008; and Solmaz and Hunte, 2008). When a target protein is not crystallizing, one can sometimes succeed in crystallization using its homolog proteins (from different organisms) (Kawai *et al.*, 2011). By analogy, usage of protein variants and protein engineering can result in great outcome (**Chapters III and IV**). The disordered or flexible regions of protein often hinder crystallization. These regions can be identified and removed using limited proteolysis, resulting in a compact domain structure of protein. This technique was successfully implemented for crystallization in **Chapter III**.

### **1.5. Objectives of this dissertation**

*C. jejuni* has proven extremely resistant to molecular pathogenesis analysis. One of the difficulties encountered in the field would be the lack of virulence factors or effectors in the *C. jejuni* genome sequence that are similar to those of better-understood pathogens. Recent studies on novel virulence factors Cj0977 and JlpA appeared to open up a hopeful avenue towards the molecular basis for *C. jejuni* virulence. As noticed before, the sequences of these proteins did not show any significant similarity to other protein sequences or structures with a known protein function. How is Cj0977, a

cytoplasmic protein, linked to flagellar gene expression? What is the role of Cj0977 in invasion of *C. jejuni* into host cells? What is the molecular basis for host cell recognition by JlpA, a surface-exposed lipoprotein adhesin? Does JlpA adopt a protein fold? Clearly, three-dimensional structures of these virulence proteins would give insights into the functions and mechanisms underlying *C. jejuni* virulence processes. For the purpose of structure determination, my work focused on growing high quality crystals of Cj0977 (**Chapter III**) and JlpA (**Chapter IV**). By exploiting protein expression systems, purification methods, and crystallization methods, we aimed to surmount the difficulties associated with initial non-productive crystallization of these proteins. These studies led up to visualize the first structural view of each protein, providing some insights in their functions.

In search of novel virulence factors of *C. jejuni*, we initiated structural studies of putative lipoproteins of this organism (**Chapter V**). Bacterial lipoproteins are known to play various important roles in bacterial pathogenesis. Since our target lipoproteins were putative proteins and did not reveal any significant sequence similarities to lipoproteins from other bacteria, several fundamental questions were raised. Can these putative lipoproteins be expressed in a heterologous system (*E. coli*)? Do these proteins with seemingly novel sequences adopt new protein folds? Do *C. jejuni* lipoproteins share common properties with other bacterial lipoproteins? During this project, much effort was made to establish a platform for structure-function relationships of *C. jejuni* novel lipoproteins, the long-term objective of this project. For this purpose, we employed a systematic approach to rapidly clone and produce target proteins. One major accomplishment of this study is the crystal structure of Cj0090 that reveals a novel

variant of the immunoglobulin fold among bacteria. Progress on other selected proteins Cj1026c, Cj1090c, and Cj1649 paved the way for structure-function studies and application to potential therapeutics.



## **Chapter II. Materials and methods**

## **2.1. Materials**

Restriction enzymes and the Ph.D.-12 phage display peptide library kit were obtained from New England Biolabs. *Pfu* DNA polymerase and QuikChange® II XL Site-Directed Mutagenesis Kit were purchased from Stratagene. T4 DNA ligase and IPTG were purchased from Promega. Primers used for PCR were synthesized by IDT. pGEX-6P-1 vector and pET vectors were purchased from GE Healthcare and Novagen, respectively. Ni-NTA agarose, QIAquick PCR purification kit, QIAquick Gel extraction kit, and QIAprep® Spin Miniprep kit were purchased from Qiagen. *Escherichia coli* strains BL21(DE3), BL21(B834), and DL41 were obtained from Stratagene. SelenoMet Medium was purchased from Molecular Dimensions. EDTA-free protease inhibitor cocktail and trypsin (sequencing grade) were purchased from Roche. Glutathione Sepharose™ 4 fast flow, Hitrap™ Q HP, HiPrep 16/60 Sephacryl S-100 HR, and HiPrep 16/60 Sephacryl S-200 HR were purchased from GE Healthcare. The Crystal Screen and Crystal Screen 2 reagent kits, crystallization plates, and siliconized cover slides were purchased from Hampton Research. The pHClear Suite and JCSG+ Suite crystallization reagent kits were obtained from Qiagen. Unless stated otherwise, chemicals were purchased from Sigma-Aldrich.

## **2.2. Protein production**

### **2.2.1. Cloning and expression of Cj0977**

The initial expression strain harboring the plasmid (pET15b::*cj0977*) was constructed in the Dr. Guerry's laboratory (Naval Medical Center, Silver Spring, MD). The *cj0977* gene was cloned from *C. jejuni* strain 81-176 (CJJ81176\_0996) and

expressed as N-terminal His-tagged full-length Cj0977 (His-Cj0977) in the BL21(DE3) strain.

For construction of three Cj0977 variants, Cj0977<sub>p21</sub>, Cj0977<sub>p19</sub>, and Cj0977<sub>p17</sub>, as GST fusion proteins, the following primers were designed: pFbam (5'-GCA GGA TCC GAT AAT TTT GAA GAA TAT GCA C-3'), pR1xho (5'-GGT CTC GAG TTA TTT TCC ACC CAC AGA GGC C-3'), pR2xho (5'-GGT CTC GAG TTA GGT ACC TTG TTC TTG ATT TTC G-3'), and pR3xho (5'-CGG CTC GAG TTA GAG TTT AAA TAT ATG CTC ATC G-3'). PCR was carried out using primer pairs of pFbam/pR1xho, pFbam/pR2xho, and pFbam/pR3xho for Cj0977<sub>p21</sub>, Cj0977<sub>p19</sub>, and Cj0977<sub>p17</sub>, respectively, and the plasmid pET15b::cj0977 as template. Each DNA fragment was cloned into a pGEX-6P-1 expression vector, and each construct was transformed into BL21(DE3) strain and a methionine auxotroph strain (DL41) for protein expression.

For production of native N-terminal His-tagged Cj0977 (native His-Cj0977), the *E. coli* strain BL21(DE3) harboring pET15b::cj0977 was grown in 1L of liquid Luria Bertani (LB) medium supplemented with 100 µg/mL ampicillin at 37°C in a shaker (New Brunswick Scientific). When the culture reached an OD<sub>600</sub> of 0.7-0.8, IPTG was added to a final concentration of 0.2 mM and the culture was grown for 3 hr more at room temperature. Subsequently, bacterial cells were harvested by centrifugation at 5,200 rpm for 20 min at 4°C (Sorvall), and the cell pellet was resuspended in 40 mL of lysis buffer containing 20 mM Tris (pH 8.0), 250 mM NaCl, and 2 mM EDTA-free protease inhibitor cocktail, and stored at -20°C until purification.

For production of selenomethionine-substituted N-terminal His-tagged Cj0977 (smHis-Cj0977), the methionine auxotroph strain B834(DE3) harboring pET15b::cj0977

was grown in 1L of SelenoMet medium supplemented with 100 µg/ mL ampicillin. When using SelenoMet medium, the bacteria were grown at 37°C for 8-11 hr without light prior to IPTG induction. The bacterial culture and cell preparation were carried out as described for the nature His-Cj0977.

For production of each of selenomethionine-substituted GST-Cj0977 variants, GST-Cj0977<sub>p21</sub>, GST-Cj0977<sub>p19</sub> and GST-Cj0977<sub>p17</sub> (smGST-Cj0977<sub>p19</sub>, smGST-Cj0977<sub>p19</sub>, and smGST-Cj0977<sub>p17</sub>), the methionine auxotroph strain DL41 harboring the plasmid encoding the target gene was used. Typically, bacteria were grown in 1L of SelenoMet medium as described for smHis-Cj0977. Bacterial cells were collected by centrifugation at 5,200 rpm for 20 min at 4°C and resuspended in 40 mL of lysis buffer containing PBS, 2 mM EDTA-free protease inhibitor cocktail, and 1 mM DTT, and stored at -20°C for later use in purification.

### **2.2.2. Cloning and expression of JlpA**

The initial expression strain harboring the plasmid (pET19b::*jlpA*<sub>18-372</sub>) was obtained from the Dr. Guerry's laboratory. The DNA fragment encoding the mature form of JlpA lacking the signal peptide (JlpA<sub>18-372</sub>) was amplified by PCR using *C. jejuni* strain 81-176 genomic DNA and was inserted into the NdeI/BamHI sites of pET19b to express an N-terminal His-tagged protein in the BL21(DE3) strain.

For the phasing purpose, additional methionine sites were introduced in the JlpA sequence. Four JlpA<sub>18-372</sub> variants, JlpA<sub>L45M/I160M</sub>, JlpA<sub>L45M/L284M</sub>, JlpA<sub>I160M/L284M</sub>, and JlpA<sub>L45M/I160M/L284M</sub> were produced by using the plasmid pET19b::*jlpA*<sub>18-372</sub> as template and QuikChange® II XL Site-Directed Mutagenesis Kit according to the manufacturer's instructions. To confirm the correct sequences in four clones, DNA sequencing was

performed by SeqWright (Houston, TX). Primers used in site-directed mutagenesis to create JlpA<sub>18-372</sub> variants are as follows: I45M\_forward (5'- GTT AAA CAA GAA ATT GCA AGC ATG TCT CAG GAT TCT GGA ATA AAG - 3'), I45M\_reverse (5'- CTT TAT TCC AGA ATC CTG AGA CAT GCT TGC AAT TTC TTG TTT AAC - 3'), I160M\_forward (5'- GAT CCA AAA ATC AGC TCT TTT ATG AAT AAA TTA AGC TCG GAT TCT - 3'), I160M\_reverse (5'- AGA ATC CGA GCT TAA TTT ATT CAT AAA AGA GCT GAT TTT TGG ATC - 3'), L284M\_forward (5'-ATA GCA ACT GCT AAG GAA AAT ATG CAA ACC TTA AAA GCT CAA AGT - 3'), and L284M\_reverse (5'-ACT TTG AGC TTT TAA GGT TTG CAT ATT TTC CTT AGC AGT TGC TAT - 3').

The plasmids containing the genes encoding JlpA<sub>18-372</sub> and its variants were expressed in *E. coli* BL21(DE3) or B834(DE3) using liquid LB or SelenoMet media supplemented with 100 µg/mL ampicillin. Bacterial cells were harvested by centrifugation at 5,200 rpm for 20 min at 4°C, and the cell pellet was resuspended in 40 mL lysis buffer containing 20 mM Tris (pH8.0), 250 mM NaCl, 0.1% Triton-X, 5 mM β-mercaptoethanol (BME), and 2 mM EDTA-free protease inhibitor cocktail, and stored at -20°C until purification.

### **2.2.3. Cloning and expression of novel lipoproteins**

The DNA fragment corresponding to Cj0090 lacking the N-terminal signal sequence (Cj0090<sub>16-122</sub>) was PCR amplified from genomic DNA of *C. jejuni* NCTC 11168 (ATCC 700819) using primers Cj0090f (5'-CCG GGT ACC TGT GCA CCA AGT TAT CAA ATA AAT TC-3') and Cj0090r (5'-CCG GAG CTC TTA ATT TTT TGC TTT AAT TTC TAA TC-3'). The PCR product containing KpnI and SacI cloning

sites (underlined) was purified, digested with KpnI and SacI, and ligated with KpnI/SacI-digested pET45b(+). The plasmid encoding the target protein was transformed into BL21(DE3) to express an N-terminal His-tagged Cj0090<sub>16-122</sub>.

The DNA fragment encoding Cj1026c was PCR amplified from genomic DNA of *C. jejuni* NCTC 11168 using primers Cj1026cf (5'-CCG CAT ATG AAA AAA ATT TAT TTT ATG CTA GCA-3') and Cj1026cr (5'-CGC CTC GAG ATA AGC AAA CAA TTC TTT CCA CTT G-3'). For the cloning purpose, the NdeI and XhoI cloning sites (underlined) were introduced to the forward and reverse primers, respectively. The PCR product was ligated with NdeI/XhoI-digested pET20b(+). The plasmid encoding the Cj1026c protein was transformed into BL21(DE3) to express a C-terminal His-tagged protein (His-Cj1026c).

The DNA fragment encoding Cj1090c without the N-terminal signal sequence (His-Cj1090c<sub>16-170</sub>) was PCR amplified from genomic DNA of *C. jejuni* NCTC 11168 using primers Cj1090cf (5'-CCG GGT ACC TGT GGA TAT ATT CCT ACA TC -3') and Cj1090cr (5'-CCG GAG CTC TTA GTA TTT GGA ATC ACG TTT TTG-3'). The PCR product contained the KpnI/SacI sites (underlined), and thus molecular cloning and protein expression were carried out as described for Cj0090 above.

The DNA fragment corresponding to Cj1649 without the N-terminal signal sequence (Cj1649<sub>18-199</sub>) was amplified by PCR using genomic DNA from *C. jejuni* NCTC 11168 using primers Cj1649f (5'-CCG GGT ACC TGT TCT TTA AGA CAC GAA ACC-3') and Cj1649r (5'-CCG GAG CTC TTA GGA TAA GTT AGA ATT TAC CC-3'). Like the Cj0090 and Cj1090c systems, the PCR product was cloned into

pET45b(+)) to express Cj1649<sub>18-199</sub> as an N-terminal His-tagged protein. Molecular cloning and protein expression followed as mentioned in Cj0090 above.

For production of Cj0090, Cj1090c, and Cj1649, each BL21(DE3) expression strain was grown in 1L of liquid LB medium supplemented with 100 µg/mL of ampicillin at 37°C. Bacterial cells were collected by centrifugation and resuspended in lysis buffer as described for native His-Cj0977 above (See 2.2.1.).

For production of Cj1026c, the BL21(DE3) expression strain harboring the plasmid pET20b(+):*cj1026c* was grown in 1 L of liquid LB medium supplemented with 100 µg/mL of ampicillin at 37°C. During the culture, the His-Cj1026c protein was moderately expressed without IPTG induction and secreted to the media. When the culture reached an OD<sub>600</sub> of ~0.9, the culture was centrifuged at 5,200 rpm for 20 min at 4°C, and the supernatant was collected for the protein purification.

### 2.3. Protein purification

To obtain highly pure proteins, multiple purification steps were applied for each protein:

Protein	Purification procedure
His-Cj0977	3 steps (NTA – HiTrapQ – GEL)
Cj0977 <sub>p19</sub> and Cj0977 <sub>p17</sub>	3 steps (GST – HiTrapQ – GEL)
His-JlpA <sub>18-372</sub> and its four variants	3 steps (NTA – HiTrapQ – GEL)
His-Cj0090 <sub>16-122</sub>	2 steps (NTA – GEL)
His-Cj1026c	3 steps (ASP – NTA – GEL)
His-Cj1090c <sub>16-170</sub>	3 steps (NTA – HiTrapQ – GEL)
His-Cj1649 <sub>18-199</sub>	3 steps (NTA – HiTrapQ – GEL)

NTA, Ni-NTA affinity column; HiTrapQ, HiTrapQ anionic exchange column; GEL, gel filtration column; GST, GST affinity column; and ASP, ammonium sulfate precipitation.

### **2.3.1. Bacterial cell lysis**

To purify target proteins overexpressed in the cytosol, bacterial cells were first disrupted. The frozen cell stock in 40 mL lysis buffer was thawed and sonicated at 30% duty cycles and 5 output control for 3 min (3 times) on ice using a Branson 450 sonifier. The cell lysate was centrifuged at 15,000 rpm for 20 min at 4°C, and the supernatant containing soluble proteins was collected. Typically, subsequent purification steps were carried out without delay.

### **2.3.2. Ammonium sulfate precipitation**

To purify Cj1026c secreted in the culture medium, the supernatant was collected from the culture and subjected to ammonium sulfate precipitation at 60% saturation by adding 390 g of ammonium sulfate to the 1000 mL sample. Following the salting-out process on ice for about 1hr, the precipitated proteins were collected by centrifugation at 15,000 rpm for 30 min at 4°C. The pellet containing His-Cj1026c was resuspended in 20 mL of 20 mM Tris-Cl, pH 8.0 and 500 mM NaCl and dialyzed against buffer containing 20 mM Tris-Cl, pH 8.5 and 250 mM NaCl overnight at 4 °C. The protein sample was collected from the dialysis bag and used in further purification steps.

### **2.3.3. His-tag affinity chromatography**

Typically, we used batch/gravity-flow purification techniques for His-tagged proteins. The soluble fraction containing the protein of interest (from 2.3.1.) was



combined with 3 mL of Ni-NTA agarose beads and subjected to a binding reaction for 30 min at 4°C on a rotator. Subsequently, protein-bead complexes were loaded into a gravity flow manual column, and the flow-through (FT) fraction was removed. The column was washed with 20 mL of buffer-A (20 mM Tris-HCl pH 8.5, 250 mM NaCl). The bound proteins were eluted by a step gradient method using imidazole up to 250 mM in buffer-A. The presence and purity of the target protein in the elution fractions were checked by SDS-PAGE. The highly pure elution fractions were combined and subjected to a further purification step.

#### **2.3.4. GST affinity chromatography**

For GST-fusion proteins, we used batch/centrifugation purification techniques. The soluble fraction of cell lysate was combined with 4 mL of Glutathione Sepharose™ 4 fast flow beads and subjected to a binding reaction on a rotator at 4°C for 3hr. Bead-protein complexes were collected by centrifugation at 2,000 rpm for 4 min and the supernatant was discarded. To remove non-specifically bound proteins, 10 mL of PBS containing 1 mM DTT and 0.005% Triton-X was added to tubes and gently mixed. The beads were collected by centrifugation and the supernatant (wash 1) was decanted. This washing step was repeated three times (wash 2-4).

To remove the GST moiety on column, bead-protein complexes were further washed twice with 10 mL of GST cleavage buffer containing 50 mM Tris-Cl (pH 7.0), 150 mM NaCl, 1 mM EDTA, and 1 mM DTT. Subsequently, 4 mL of GST cleavage buffer and 50 µL (100 units) of PreScission protease were added to the tube containing bound GST-proteins. The cleavage reaction took place at 4°C with gentle agitation on the rocker for 8 hr. After centrifugation, the supernatant (elution 1) containing target

proteins (now GST-free) was collected. This step was typically repeated three times (elution 2-4) by eluting with 3 mL of cleavage buffer for each elution sequence. The elution fractions were checked using SDS-PAGE for the presence and purity of the target proteins. The highly pure elution fractions were combined and subjected to a further purification step.

### **2.3.5. Anion exchange chromatography**

To prepare the injection sample suitable for anion exchanger binding, the salt concentration and buffer pH was first adjusted. This was done by diluting (typically two to four-fold) the protein sample (from the previous step) with buffer-B (20 mM Tris-HCl, pH 8.0). After correctly connecting a 5 mL HiTrap Q column to an AKTA Purifier system, the sample was injected to the column equilibrated with 50 mM NaCl in buffer-B. Following a washing step, the bound proteins were eluted with a linear gradient of 50 mM to 1M NaCl in buffer-B (total elution volume 50 mL). The peak fractions were evaluated for purity and presence of the target protein using SDS-PAGE. The selected elution fractions were combined and applied to the next purification step.

### **2.3.6. Gel filtration chromatography**

A gel filtration column was equilibrated with buffer-C (20 mM Tris-HCl, pH 8.0 and 150 mM NaCl) or buffer-D (50 mM HEPES, pH 7.0, 200 mM NaCl, 0.1 mM EDTA, and 5% glycerol) on an AKTA Purifier system. We used a HiPrep™16/60 SephacrylS-100 (or S-200) HR column for a large sample volume (up to 3 mL) or a high-resolution column Superdex75 10/300 GL for a small sample volume (0.5 mL). Thus, before injection into the column, the sample from the previous step was concentrated to ~3 mL or 0.5 mL using a Vivaspin concentrator filter at 3,600 rpm. The proteins were eluted by

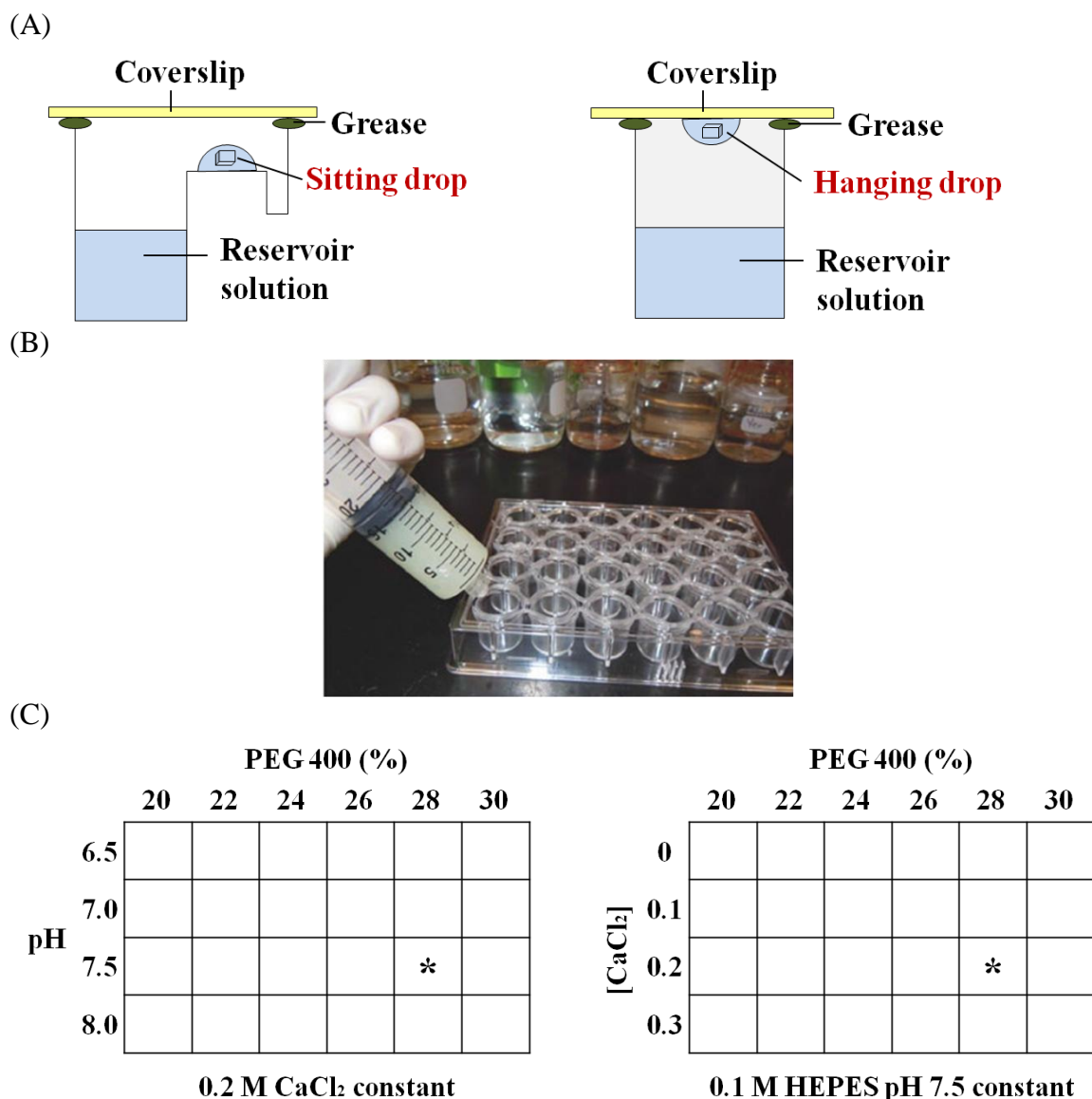
one column volume of buffer-C or buffer-D. The peak fractions were analyzed for purity and presence of target protein by SDS-PAGE. The highly pure proteins fractions were selected for crystallization.

## **2.4. Protein crystallization**

### **2.4.1. Screening**

As a high concentration (2-50 mg/mL) of the protein sample is a prerequisite for protein crystallization, the highly pure fractions from the final step of purification (gel filtration) were concentrated by using a Vivaspin concentrator filter, and the protein concentration was determined using a UV spectrometer (Cary 50 Varian). The solubility among different proteins widely varies: His-Cj0977 and its variants were concentrated ~35 mg/mL, while His-JlpA<sub>18-372</sub> and its variants and His-Cj0090<sub>16-122</sub>, His-Cj1026c, His-Cj1090c<sub>16-170</sub>, and His-Cj1649<sub>18-199</sub> were concentrated ~10 mg/mL. The concentrated protein sample was kept at 4°C during crystallization trial. Since crystallization conditions of a target protein are unknown, a screening process was initiated with crystallization reagent kits including pH Clear, Crystal Screen, Crystal Screen 2, and JCSG+ Suite.

The Oryx robot system (Douglas instrument) was used for initial high-throughput screening. A 96 well plate was prepared manually by placing 100 µL of unique solution in each reservoir well. Subsequently, 0.2 µL of protein solution and 0.2 µL of crystallization solution were mixed together by the Oryx robot system using the sitting drop vapor diffusion method (Figure 2-1 A left). The 96 well plate was then tightly sealed, incubated at 17°C, and observed for crystal growth under a microscope over time.



\* = CS #14. 0.2 M  $\text{CaCl}_2$ , 0.1 M HEPES pH 7.5, 28% v/v PEG 400

**Figure 2-1. Crystallization setup.** (A) Schematics of crystallization methods: sitting drop (left) and hanging drop (right). (B) A  $4 \times 6$  well plate and crystallization plate preparation. Shown is an example of greasing the crystallization plate with Petroleum jelly. (C) Schematics of the optimization strategy. For example, if the condition of Crystal Screen #14 (0.2 M  $\text{CaCl}_2$ , 0.1 M HEPES pH 7.5, 28% v/v PEG 400) is found as a promising condition of a protein, the condition is optimized to produce larger crystals. During optimization, crystallization parameters are varied. Along a given axis, the concentration of one component of the well solution is altered while the others are kept constant. Shown are two plates that could be set up in the first round of optimization. The initial condition #14 is also incorporated in the plate (\*) (adapted from Yeo, 2013).

Systematic trials of manual screening were pursued using the hanging drop vapor diffusion method (Figure 2-1 A right). Before the hanging drop method, the crystallization plate was prepared by placing sealant to seal the cover slides on the plate (Petroleum jelly is cost effective and easy to handle) (Figure 2-1 B). The 24 well plate was prepared by placing 1 mL of unique crystallization reagent in each reservoir well. To set up hanging drops, 1  $\mu$ L of protein solution and 1  $\mu$ L of reservoir solution were mixed on a circular siliconized cover slide, which was then placed upside down and covered the reservoir well airtightly.

#### **2.4.2. Optimization and crystallization growth**

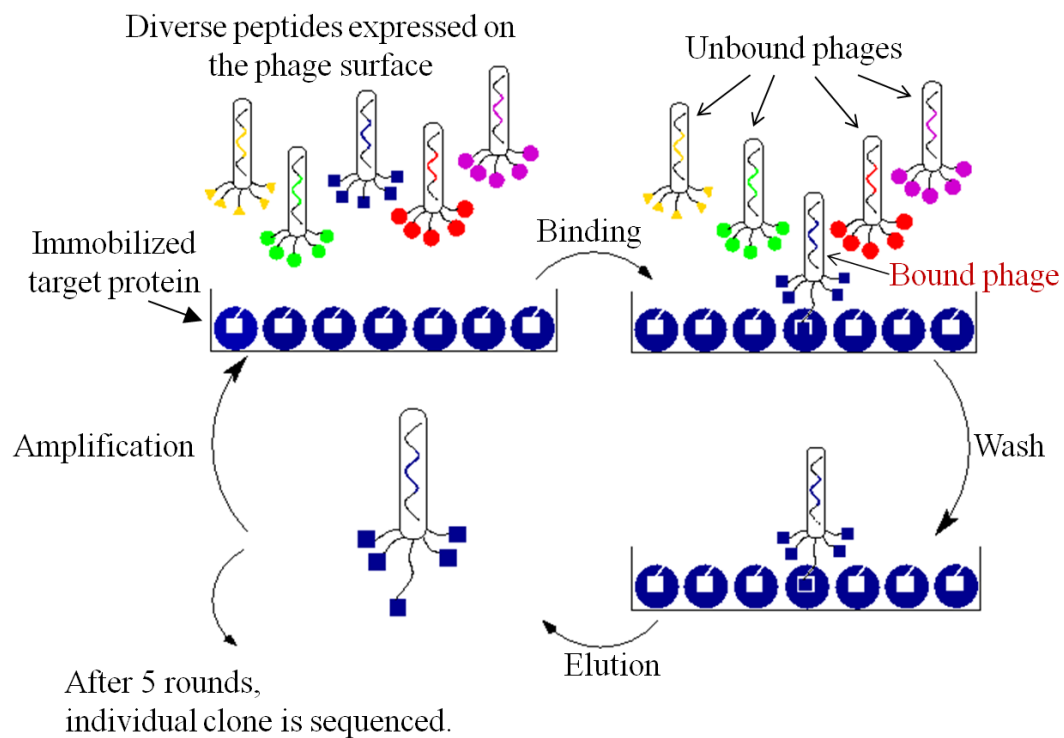
To obtain crystals suitable for data collection, optimization procedures were almost always necessary. Optimization is commonly done manually using the hanging drop method. Optimization of initial crystallization conditions was performed by varying one parameter (buffer, salt, or precipitant) along an axis of a 4x6 well plate while holding the others constant. On the other 4x6 well plate axis, a second parameter was varied. This way produced a 24-well plate with each well containing a unique mixture of components (Figure 2-1 C). Each of the 24 wells contained a solution slightly different than the initial condition used as a starting point for the optimization, and therefore usually drops yielded improved crystals. In addition, a kit composed of 96 different additives (Hampton Research) was used to screen improved crystallization conditions. Conditions yielding promising behavior were further optimized in an effort to grow diffraction quality crystals.

## **2.5. Limited proteolysis assay**

When a preprotein resisted crystallization, limited proteolysis was extensively explored to identify stable domain(s) of the protein. This technique was particularly successful for Cj0977 in this study. To find an optimal protease to protein ratio for limited proteolysis, reactions with different trypsin to protein ratios (w/w) (typically 1:50, 1:100 and 1:250) were performed. For each 100  $\mu$ L reaction, 100  $\mu$ g of a pure protein was incubated with trypsin in proteolysis buffer (1 mM EDTA, 50 mM Tris-HCl pH 7, 100 mM NaCl) at 37°C. The reaction was monitored at 0, 10, 20, 30, and 60 min by taking a 15  $\mu$ L aliquot of the reaction at each time point. The aliquots of the reactions were stopped by adding 15  $\mu$ L of 2X SDS-PAGE loading buffer followed by boiling for 4 min, and were analyzed with SDS-PAGE. The optimal reaction ratio that yields a stable domain was determined by the SDS-PAGE analysis. To identify stable domain(s) of the target protein, SDS-PAGE gels was transferred to a PVDF membrane using the Trans-Blot SD Semi-Dry Transfer Cell™ (Biorad) and submitted for N-terminal sequencing.

## **2.6. Phage display**

Using the Ph.D.-12 phage display peptide library kit, specific peptides that interact with Cj0090, Cj1026c, Cj1090c, or Cj1649 were screened. The general procedures (Figure 2-2) used for phage production or analysis, and media and solution preparations were carried out according to the manufacturer's instructions (New England Biolabs).



**Figure 2-2. Schematic procedures of phage display.** A library of phage, each displaying a different peptide sequence, is exposed to a plate coated with the target protein. Unbound phage is washed away and specifically bound phage is eluted. The eluted phage is amplified, and the process is repeated for a total 5 rounds followed by isolation and sequencing of individual clones (adapted from instruction manual ‘New England Biolabs’).

### **2.6.1. Phage titering**

A single colony of *E. coli* ER2738 was inoculated to 100 mL of LB medium supplemented with 20 µg/mL tetracycline and incubated overnight at 37°C. The overnight culture was diluted 1:100 in LB medium and incubated at 37°C until mid-log phase ( $OD_{600} \sim 0.5$ ). Agarose top (1 g Bacto-Tryptone, 0.5 g yeast extract, 0.5 g NaCl, 0.1 g  $MgCl_2 \cdot 6H_2O$ , 0.7 g agarose / 100 mL) was melted, dispensed in 3 mL aliquots, and kept at 45°C until use. 10-fold serial dilutions of phage solutions were prepared in LB medium. 200 µL of *E. coli* ER2738 culture that had reached to mid-log phase was mixed with 10 µL of each phage dilution and incubated at room temperature for 2 min. Each of the infected cells was transferred to a 3 mL agarose top aliquot and poured onto a pre-warmed LB/ IPTG/Xgal plate. After incubation at 37°C overnight, the blue plaques on plates were counted to determine phage titres.

### **2.6.2. Panning procedure**

His-Cj0090<sub>16-122</sub>, His-Cj1026c, His-Cj1090c<sub>16-170</sub>, and His-Cj1649<sub>18-199</sub> were prepared as 20 µg/mL, 50 µg/mL, 100 µg/mL, and 40 µg/mL solutions, respectively. 150 µL of each protein was immobilized to five wells of the Costar 96-well polystyrene microplate at 4°C overnight. Subsequently, the wells were completely filled up with blocking buffer (5 mg/mL BSA, 0.1 M  $NaHCO_3$ , pH 8.6) and incubated for 1h at 4°C. Wells were washed six times with TBST (TBS containing 0.1% Tween-20). 10 µL of phage library Ph.D.-12 ( $4 \times 10^{10}$  phage) was diluted to 500 µL in TBST and equally dispensed in the five wells (100 µL each). The plate was incubated for 1hr at room temperature with gentle agitation. Unbound phages were removed from the wells by washing ten times with TBST. Bound phages were eluted by adding 100 µL of the target



protein in TBS buffer to compete the bound phage away from the immobilized target on the plate. The plate was gently agitated for 1 hr at room temperature. The eluate was collected in a microcentrifuge tube. A small amount (~1  $\mu$ L) of the eluate was titered as described above (See 2.6.1.). The rest of the eluate was amplified in *E. coli* ER2738 and isolated by PEG precipitation. This panning procedure was repeated for a total of five rounds. The stringency of selection was increased by using 0.5% Tween-20 in TBS from the second round to reduce the frequency of non-specific phage binding.

### **2.6.3. Binding phage M13 clone amplification**

To amplify phage, an overnight culture of *E. coli* ER2738 was diluted 1:100 in LB medium. 1 mL of diluted culture was dispensed into culture tubes. Phage plaques from titration of the fifth round were stabbed using a sterile pipette tip and transferred to a tube containing 1 mL of diluted culture. The cultures were incubated at 37°C for 4.5 hr, transferred to microcentrifuge tubes, and centrifuged at 10,000 rpm for 1 min to pellet bacterial cells. Supernatants were transferred to a fresh tube and recentrifuged. The upper 80% of supernatant was transferred to a fresh tube and used as the amplified phage stock.

### **2.6.4. Rapid purification of sequencing templates**

To purify M13 single-stranded DNA of each individual clone, 500  $\mu$ L of the amplified phage stock described above was transferred to a fresh microcentrifuge tube. 200  $\mu$ L of PEG/NaCl (20% (w/v) PEG 8000, 2.5 M NaCl) was added, mixed, incubated at room temperature for 10 min, and then centrifuged for 10 min. After discarding the supernatant, the pellet was suspended in 100  $\mu$ L Iodine Buffer (10 mM Tris-HCl (pH8.0), 1 mM EDTA, 4 M NaI), and 250  $\mu$ L of ethanol was added. After 10 min incubation at room temperature followed by centrifuging for 10 min, the supernatant was discarded.

The pellet was washed with 70% ethanol and dried. The pellet was suspended in 30  $\mu$ L TE buffer (10 mM Tris-HCl pH8.0, 1 mM EDTA).

#### **2.6.5. DNA sequencing and peptide analysis**

DNA sequencing analysis was carried out by SeqWright DNA Technology Services using -96 gIII sequencing primer (5'-CCC TCA TAG TTA GCG TAA CG-3'). The DNA sequences were translated into amino acids by using the “translate” program on the proteomics server of the Swiss Institute of Bioinformatics Expert Protein Analysis System (ExPASy, <http://www.expasy.ch/>).

#### **2.7. Phage Enzyme-linked ImmunoSorbent Assay (Phage ELISA)**

Binding specificities of selected phage clones were examined by Phage ELISA. 100  $\mu$ L of the target protein was immobilized on each well (total 36 wells) of the Costar 96-well polystyrene microplate at 4°C overnight with gentle agitation. The protein coating solution was discarded, and the wells were filled up with blocking buffer (5 mg/mL BSA, 0.1 M NaHCO<sub>3</sub>, pH 8.6) and incubated for 2 hr at room temperature. To measure the nonspecific binding of phages, the target protein-uncoated wells were also incubated with blocking buffer. After six washes with TBST buffer (TBS containing 0.5% Tween-20), 100  $\mu$ L of four-fold serial dilutions of the phage solutions in TBST buffer were added to each well, and the plate was incubated for 1 hr at room temperature with gentle agitation. After ten washes with TBST buffer, 100  $\mu$ L of HRP conjugated anti-M13 monoclonal antibody (GE Healthcare, diluted 1:5,000 in TBST) was added to each well, and the plate was incubated for 1 hr at room temperature. After ten washes with TBST, 200  $\mu$ L of Sigma FAST OPD substrate was added to each well, and the plate

was incubated in the dark for 30 min at room temperature. The reactions were terminated by the addition of 50  $\mu$ L of 4*N*-HCl. The absorbance at 492 nm was measured using Multiskan EX microplate photometer (Thermo scientific). Absorbance data were calculated by subtracting the absorbance of uncoated wells from the absorbance of target protein-coated wells. The binding curves were plotted by fitting data to the one site saturation equation:  $y = (B_{\max} \times x) / (K_d + x)$  using Sigma Plot 10.0 program.

The numbers of amplified phage were determined according to the phage titting as described in 2.6.1.. The concentrations of phage were calculated as follows: molar concentration = (the total number of phage in solution /  $6.02 \times 10^{23}$ ) / solution volume.

### **Chapter III. Structural Studies of the Virulence Protein Cj0977**

Contributed to:

Yokoyama, T., Paek, S., Ewing, C.P., Guerry, P., and Yeo, H.J. (2008) Structure of a sigma28-regulated nonflagellar virulence protein from *Campylobacter jejuni*. *J. Mol. Biol.* **384**, 364-376.

### 3.1. Introduction

*Campylobacter jejuni*, a Gram-negative motile bacterium, is a leading cause of human gastrointestinal infections worldwide. Besides the canonical role in locomotion, the flagella of *C. jejuni* play multiple and complex roles in virulence including chemotaxis, autoagglutination, invasion and colonization (Black *et al.*, 1988; Nachamkin *et al.*, 1993; Wassenaar *et al.*, 1993; and Hendrixson *et al.*, 2004). Moreover, in the absence of a specialized type III secretion system in the genome, *C. jejuni* utilizes the flagellum to secrete several nonflagellar virulence proteins to the extracellular space or host cell cytosol (Konkel *et al.*, 1999; Song *et al.*, 2004; Poly *et al.*, 2007; and Christensen *et al.*, 2009).

Flagella biosynthesis is energetically costly and is regulated in an ordered manner. Two sigma factors  $\sigma^{54}$  and  $\sigma^{28}$  are important in flagellar gene regulation. Most genes encoding protein components of the basal body, hook, and minor flagellin are controlled by  $\sigma^{54}$ , and the major flagellin and other late genes in the flagellar regulon are controlled by  $\sigma^{28}$ . Interestingly, several nonflagellar genes also appear to be regulated by  $\sigma^{54}$  or  $\sigma^{28}$  promoters, as demonstrated by a microarray-based study (Carrillo *et al.*, 2004). The *cj0977* gene was first identified as a  $\sigma^{28}$ -regulated nonflagella gene.

The *cj0977* gene encodes an acidic protein with 192 amino acids (Mr = 21.2 kDa; pI = 4.8). Cj0977 lacks a leader sequence or transmembrane domains, suggesting that this protein is cytoplasmic. Cj0977 is highly conserved within *Campylobacter* spp. and also conserved within the proteobacteria. Another study demonstrated that Cj0977 is a virulence factor regulated by a  $\sigma^{28}$  promoter in *C. jejuni* and its expression is dependent on a minimal flagellar structure (Goon *et al.*, 2006). A Cj0977 mutant in *C. jejuni* 81-176

was fully motile and produced a normal flagella filament, but was significantly reduced in invasion of intestinal epithelial cells *in vitro* (3-logs lower than the parent strain) (Goon *et al.*, 2006). Although the invasion defect in the Cj0977 mutant of 81-176 was more significant than the mutants of CiaB or FlaC (secreted virulence proteins through the *C. jejuni* flagella) (see Chapter I 1.2.1.), Cj0977 was not secreted into the supernatant, suggesting a virulence function of Cj0977 within the bacterium. The Cj0977 mutant was also attenuated in the ferret diarrhea model (Goon *et al.*, 2006). The isoelectric focusing (IEF) analysis revealed that an IEF pattern of purified flagellin from this Cj0977 mutant was identical to that of *C. jejuni* 81-176, suggesting that Cj0977 is not related with flagellin glycosylation (Goon *et al.*, 2006). The expression level of the Cj0977 protein was compared among mutants in the flagella regulon of *C. jejuni* 81-176. The mutants were either nonmotile with no flagellin expression or were reduced in motility with a truncated flagella filament. Reduced expression of Cj0977 was observed in these mutants suggesting coregulation of Cj0977 with the flagella regulon.

Based on these observations, Cj0977 appeared to be a virulence factor involved in invasion of host with a novel function. However, how Cj0977 contributes to pathogenesis of *C. jejuni* remained elusive. When we started this study, there was no crystal structure of Cj0977. In addition, Cj0977 does not share any significant sequence homology with other known proteins. Therefore, we aimed to determine the first crystal structure of this *C. jejuni* virulence protein with a novel function. It is well appreciated that protein tertiary structure is better conserved than protein primary sequence. Thus, we hypothesized that by solving the structure of Cj0977, we can uncover structural homolog proteins with a known function, which may reveal a clue of the function of Cj0977 within

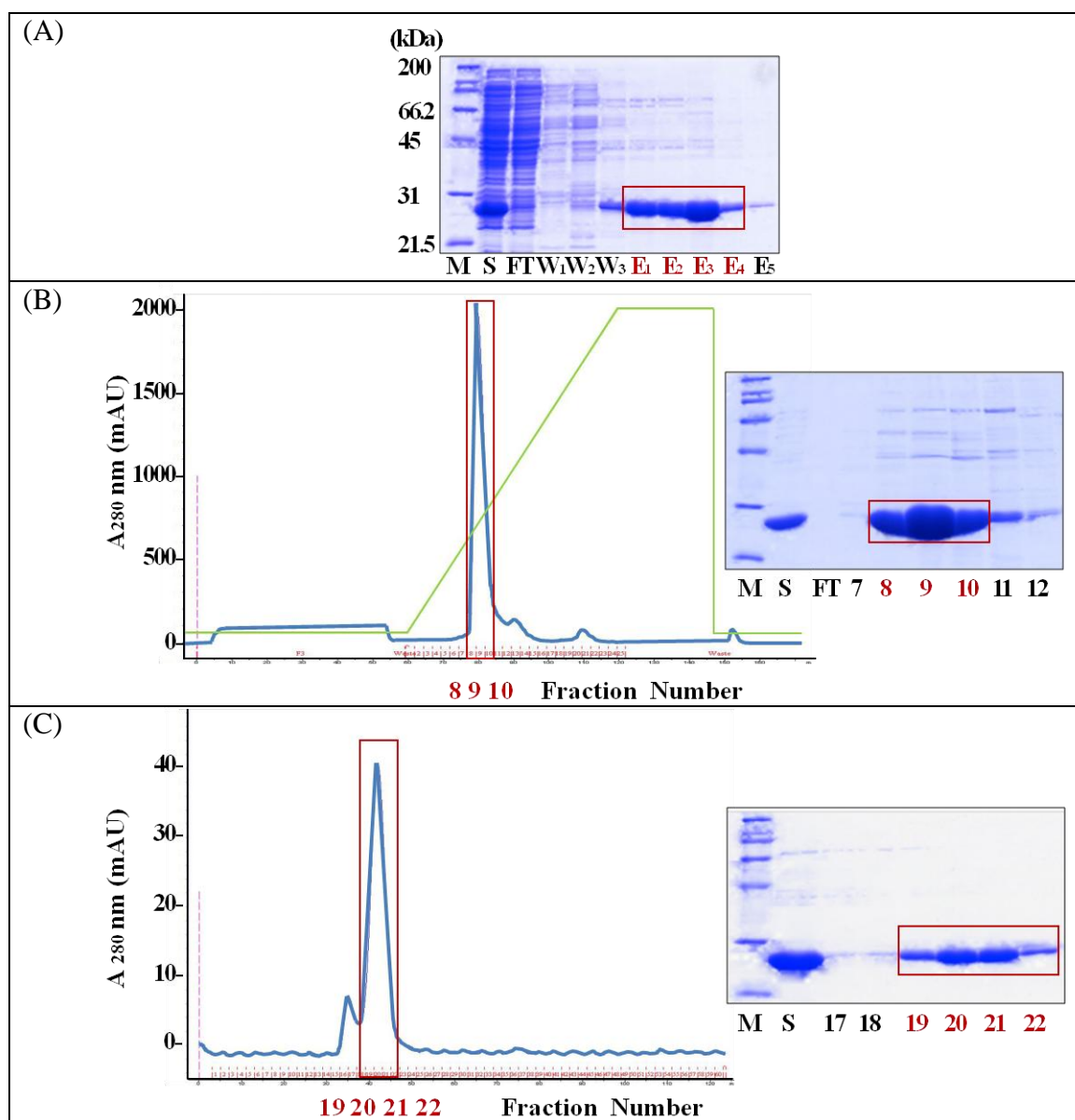
the organism and its roles in pathogenesis associated with the flagellar regulon. As Cj0977 is highly conserved within *Campylobacter* spp. as well as other epsilon-proteobacteria, the structure may further help to understand pathogenesis of other pathogens. In this chapter, we present crucial experimental steps for the structure determination. In the process of obtaining high quality crystals, one key result was defining a stable domain of Cj0977 by limited proteolysis and protein stability test. We also discuss the first view of the crystal structure and possible functions of Cj0977.

## **3.2. Results**

### **3.2.1. Purification and crystallization of His-Cj0977**

The objective of this study was to obtain highly pure protein solution and high quality diffracting crystals of Cj0977. Initially, N-terminal His-tagged full-length Cj0977 (His-Cj0977) was pursued for crystallization.

The His-Cj0977 protein was well expressed and highly soluble in the BL21(DE3) strain. The first purification process of nickel affinity chromatography (Ni-NTA agarose) resulted in a nice protein elution profile when using a step gradient of imidazole concentration of 50 mM (elution E1 and E2), 125 mM (E3), and 150 mM (E4 and E5) (Figure 3-1 A). Most impurities were removed after this purification and the isolated Cj0977 protein remained soluble, which facilitated protein purification. Considering the characteristic high net negative charge of Cj0977 (pI 4.8) at pH 8, His-Cj0977 was further purified using an anion exchange column (HiTrap Q) on AKTA Purifier. As expected, the Cj0977 protein bound tightly to the column and a single elution peak was



**Figure 3-1. Purification of His-Cj0977.** (A) SDS-PAGE analysis of nickel affinity chromatography. Elution fractions E1-E4 were combined for the next purification step. (B) Anion exchange chromatogram and SDS-PAGE analysis. Fractions 8-10 corresponding to the peak were collected and concentrated for the gel filtration column. (C) Gel filtration chromatogram and SDS-PAGE analysis. This size exclusion chromatography indicates dimeric His-Cj0977 in solution. M, Marker; S, Sample; FT, Flow-through; W, Wash; and E, Elution.

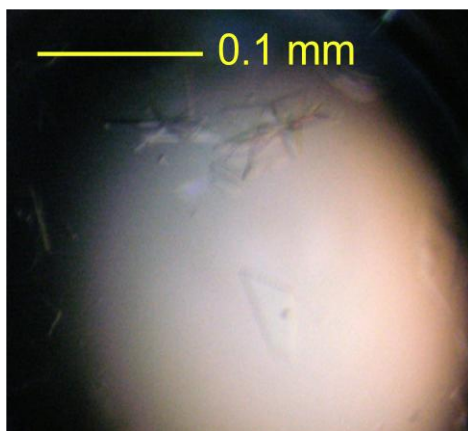


observed between 350 mM and 420 mM of NaCl (with the highest peak at 390 mM) (Figure 3-1 B). The final purification step included gel filtration column (size exclusion chromatography). The protein was eluted between 40 and 50 mL. This peak corresponds to an apparent molecular weight of 46 kDa, compared with the calibration curve, suggesting that Cj0977 adopts a dimer structure in solution (Figure 3-1 C). Subsequently, the quite pure elution fractions were concentrated for crystallization purpose. We found this protein was unusually soluble, and could routinely concentrate to ~35-45 mg/mL.

This three-step purification approach was suitable for generating a highly pure protein sample (ideally >95% purity), a prerequisite for crystallization. Indeed, several promising conditions were identified during initial crystal screening trials. After optimization process, sizable crystals (0.2 x 0.2 x 0.05 mm) were produced by vapor diffusion in hanging drops against a reservoir solution containing ammonium sulfate (1.8 M-2.2 M) and 0.1 M HEPES (pH 6.8-7.4) (Figure 3-2). However, these crystals were recalcitrant to diffraction. Thus, we next turned to other strategies to obtain diffracting crystals.

### **3.2.2. Probing domain structures of Cj0977**

There would be reasons for why those well-formed Cj0977 crystals were not diffracting. Among others, floppy or disordered regions of a protein could result in poorly ordered crystals and thus very poor diffraction. We hypothesized that crystals of His-Cj0977 were not well ordered, because the Cj0977 contains inherently disordered regions that preclude forming well ordered protein crystals. To examine this hypothesis, we used limited proteolysis to probe stable domains of Cj0977. With this approach, we can remove disordered regions of the protein, which likely disturb ordered crystal packing.



**Figure 3-2. Crystals of N-terminal His-tagged full-length Cj0977 (His-Cj0977).** Thin crystals of His-Cj0977 with various shapes were nicely grown at 17°C by vapor diffusion in hanging drops against a reservoir solution containing 2.1M ammonium sulfate precipitant within one week.

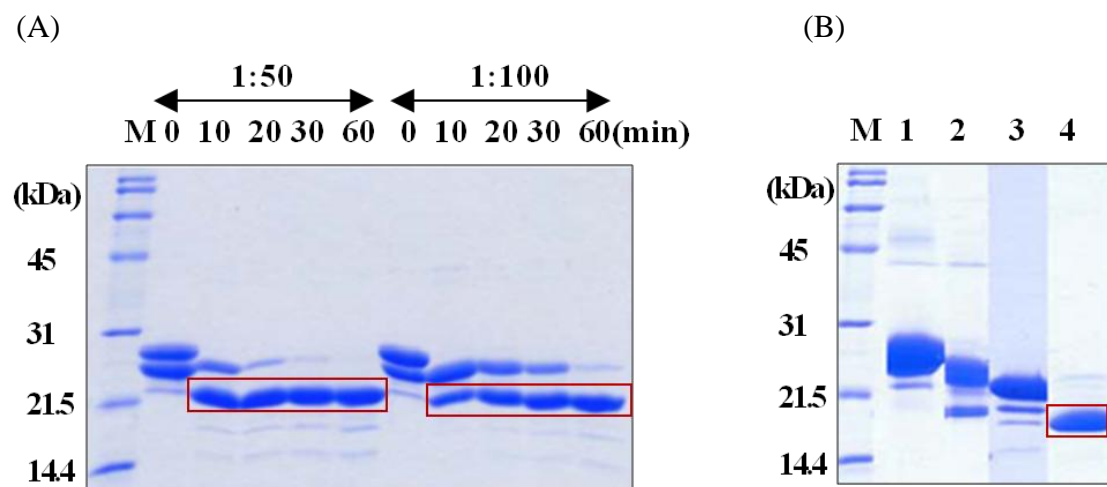
Among several available proteases, we first tested trypsin for limited proteolysis of His-Cj0977. To identify an optimal ratio of protease:protein for the efficient limited proteolysis experiment, two reactions with trypsin: His-Cj0977 ratios of 1:50 and 1:100 were carried out (Figure 3-3 A). The SDS-PAGE analysis showed that the initial full length His-Cj0977 became a smaller protein (~21 kDa) in the 1:50 reaction at 30 min and in the 1:100 reaction at 60 min, and the smaller protein remained stable for about 2hr (Figure 3-3 A). This result indicated that Cj0977 consists of a large single domain structure and small flexible or disordered segments.

Next, the stability of purified His-Cj0977 in solution was monitored to assess a half-life of the protein. After final purification, the protein was concentrated, kept at 4°C, and monitored by SDS-PAGE at each week point. This experiment nicely demonstrated the evolution and fate of purified His-Cj0977 (Figure 3-3 B). At the 4-week point, the full-length protein became a fragment of ~21 kDa similar to the domain observed after limited proteolysis. Interestingly, at the 6 week point, the initial protein resulted in an apparent single species of ~17 kDa that remained very stable for months.

Based on these results, we concluded that the 17 kDa fragment corresponded to the most compact core structure of Cj0977 and the 21 kDa fragment represented a transiently stable domain structure of Cj0977. We next hypothesized that these compact domains would yield high quality diffracting crystals, if we could produce and crystallize these smaller variants of Cj0977.

### **3.2.3. Construct design of Cj0977 variants**

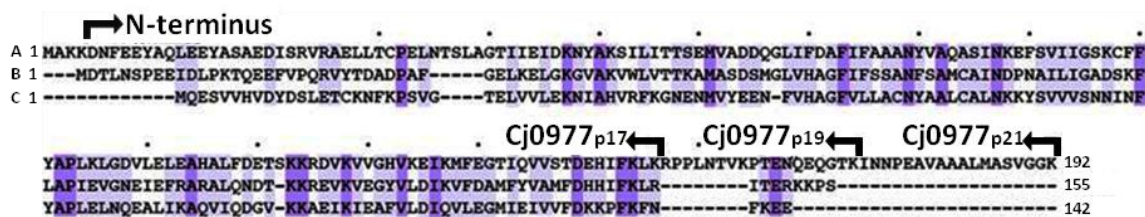
To identify the amino acid residues of the 21 kDa and 17 kDa fragments, the protein bands were transferred from the SDS-PAGE gel to a PVDF membrane, and



**Figure 3-3. Limited proteolysis and stability test of His-Cj0977.** (A) Limited proteolysis of His-Cj0977 using trypsin. Two reactions with trypsin:Cj0977 ratios of 1:50 and 1:100 were performed and the reaction progress was monitored by the SDS-PAGE gel (15%) at 0, 10, 20, 30, and 60 min time points. (B) His-Cj0977 stability test. The purified His-Cj0977 was kept in 4°C and was monitored on the SDS-PAGE gel to check stability. Lanes 1 to 4 correspond to 1 day-old sample, 2 weeks-old sample, 4 weeks-old sample, and 6 weeks-old sample, respectively (adapted from Yokoyama *et al.*, 2008). M, Marker.

submitted to the N-terminal sequencing. The peptide mapping result showed that they shared the same N-terminus (Asp-4) (Figure 3-4), meaning that the N-terminal sequence (including His-tag) of His-Cj0977 was cleaved by trypsin and the two fragments differ in the C-terminal regions. The C-terminal cleavage site of each fragment was determined by calculating the molecular mass of the amino acid sequence (Figure 3-4). Since trypsin cuts amino acid residues after Lys (K), the expected cleavage site of Cj0977<sub>p21</sub> is after K192, meaning that trypsin only cleaved N-terminal His-tag and the first four residues (MAKK; Met-Ala-Lys-Lys). The expected cleavage site of Cj0977<sub>p17</sub> is after K146, meaning that N-terminal His-tag and the first four residues (MAKK) as well as C-terminal 36 residues were removed. Based on the peptide mapping of 17 kDa fragment, both the N-terminal His-tag and the C-terminal 36 amino acid tail might hamper well-ordered crystals in His-Cj0977.

On the other hand, when we performed multiple sequence alignment, it was clear that the 17 kDa domain represented the highest conserved core domain among Cj0977 homologs (Figure 3-4). Based on the peptide mapping, three Cj0977 variants, referred to as Cj0977<sub>p21</sub>, Cj0977<sub>p19</sub>, and Cj0977<sub>p17</sub>, were constructed as GST-fusion proteins (Figure 3-4). The Cj0977<sub>p21</sub> (21 kDa) and Cj0977<sub>p17</sub> (17 kDa) domains were selected based on the experimental results, as described above. We also selected an additional construct Cj0977<sub>p19</sub> (19 kDa), since the sequence alignment suggested this fragment as a possible stable structure. The small N-terminal His-tag appeared to have a negative effect on crystal packing in Cj0977 protein. Thus, by constructing GST-fusion proteins, we planned to purify Cj0977 variants by affinity chromatography and subsequently remove the GST moiety to obtain Cj0977 variants without any affinity tag. We anticipated that



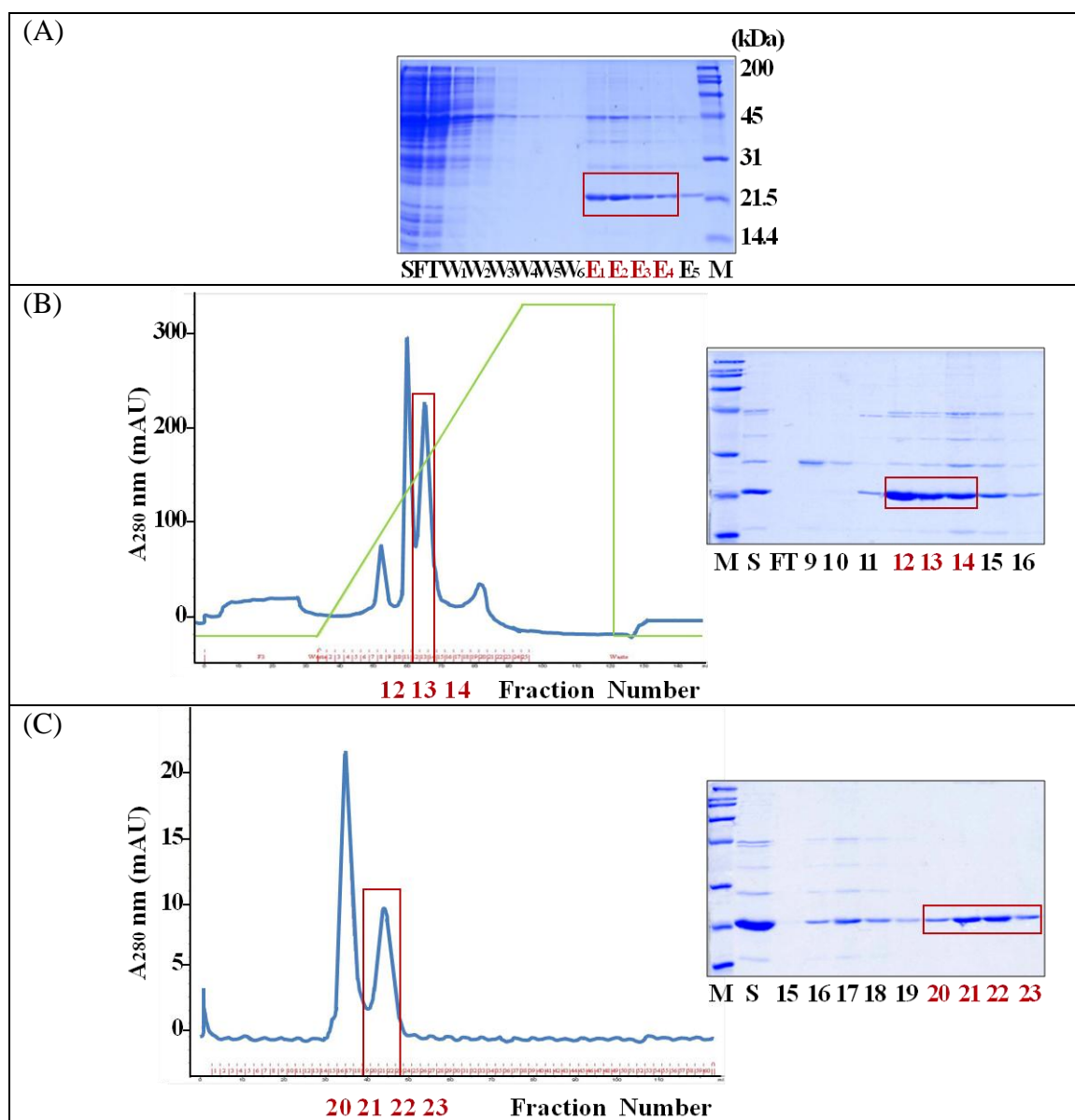
**Figure 3-4. Sequence alignment and peptide mapping.** A, B, and C represent the sequence of Cj0977, Ws0669, and Hp0420, respectively. The most significant homologs of Cj0977 included the hypothetical proteins WS0669 and HP0420 from *Wolinella succinogenes* and *Helicobacter pylori*, respectively. Identical residues and conserved residues are highlighted as purple and light purple, respectively. Based on N-terminal sequencing and molecular weight calculation, amino acid residues included in Cj0977<sub>p21</sub>, Cj0977<sub>p19</sub>, and Cj0977<sub>p17</sub> were mapped (in arrows) (adapted from Yokoyama *et al.*, 2008).

these new constructs would result in improved crystal packing.

#### **3.2.4. Purification and crystallization of Cj0977 variants**

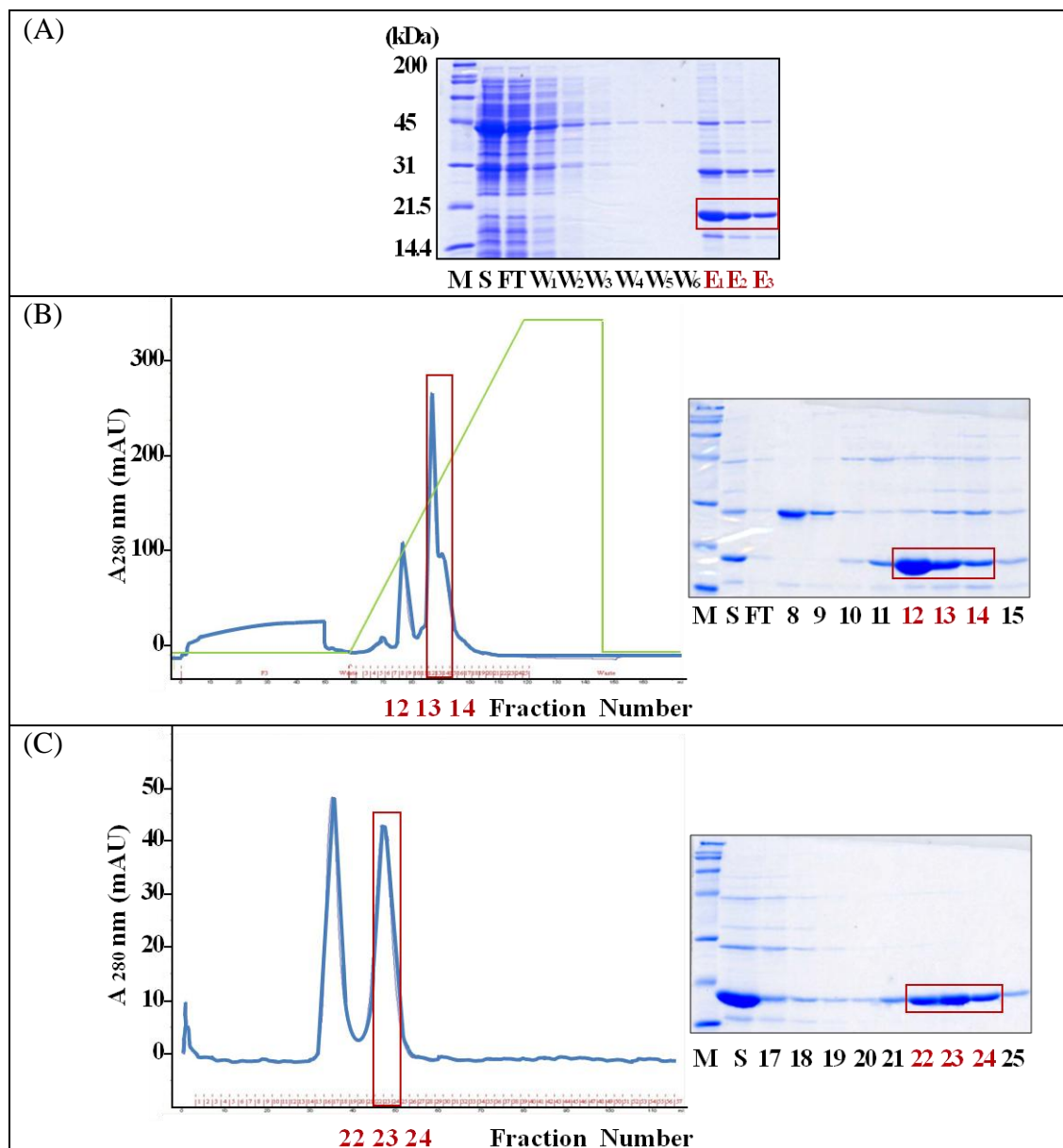
Selenomethionine incorporation into recombinant proteins facilitates protein phasing by the single- or multiwavelength anomalous dispersion (SAD or MAD) method. For the MAD phasing purpose, selenomethionine-substituted GST-Cj0977 variants, smGST-Cj0977<sub>p21</sub>, smGST-Cj0977<sub>p19</sub>, and smGST-Cj0977<sub>p17</sub> were produced. Although expressions of three variants were good, smCj0977<sub>p21</sub> was degraded during the first purification process. The other two variants, smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub> could be purified as soluble and stable proteins. Hence, we focused on these two variants for purification and crystallization.

The first purification step included affinity chromatography using glutathione-sepharose beads. The theoretical molecular weights of GST fusion moiety, Cj0977<sub>p19</sub> and Cj0977<sub>p17</sub> are 26 kDa, 19 kDa and 17 kDa, respectively. Both smGST-Cj0977<sub>p19</sub> and smGST-Cj0977<sub>p17</sub> were correctly expressed with an expected molecular weight of ~19 kDa and ~17 kDa, respectively, as shown in Figures 3-5 A and 3-6 A. After cleavage by the Prescission protease, moderate amounts of GST-free smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub> proteins were obtained. However, unlike His-tagged protein purification (see 3.2.1.), we observed that the yield and enrichment of the target proteins were relatively low (Figures 3-5 A and 3-6 A). The second purification step using an anion exchange column resulted in several peaks in the chromatograms, showing further enrichment of the target proteins (Figures 3-5 B and 3-6 B). Both smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub> bound to the anion exchanger and showed very similar elution profiles with the peak around 500 mM NaCl. In the final gel filtration column, we could further remove impurities, resulting in highly



**Figure 3-5. Purification of Cj0977<sub>p19</sub>.** (A) SDS-PAGE analysis of affinity chromatography using glutathione-sepharose beads. Cj0977<sub>p19</sub> moiety was eluted after PreScission protease cleavage. (B) Anion exchange chromatogram and SDS-PAGE analysis. The first peak and second peak correspond to GST and Cj0977<sub>p19</sub>, respectively. The peak fractions of Cj0977<sub>p19</sub> were collected and concentrated for the gel filtration chromatography. (C) Gel filtration chromatogram and SDS-PAGE analysis. The pure elution fractions 20-23 were collected for crystallization. His-Cj0977<sub>p19</sub> also adopts a dimer in solution. M, Marker; S, Sample; FT, Flow-through; W, Wash; and E, Elution.



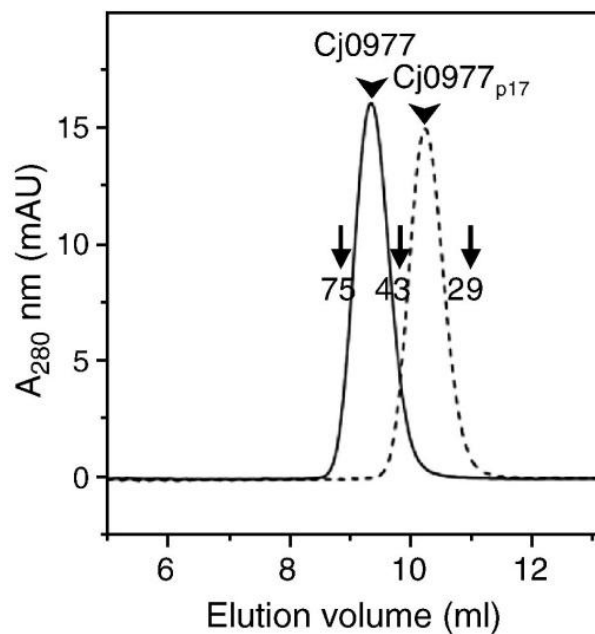


**Figure 3-6. Purification of Cj0977<sub>p17</sub>.** (A) SDS-PAGE analysis of affinity chromatography. The elution fractions E1-E3 were combined for the next chromatography. (B) Anion exchange chromatogram and SDS-PAGE analysis. The first peak and the second peak correspond to GST and Cj0977<sub>p17</sub>, respectively. The combined elution fractions 12-14 were concentrated for the gel filtration. (C) Gel filtration chromatogram and SDS-PAGE analysis. Like His-Cj0977 and His-Cj0977<sub>p19</sub>, His-Cj0977<sub>p17</sub> also is a dimer in solution. M, Marker; S, Sample; FT, Flow-through; W, Wash; and E, Elution.

pure protein samples (Figures 3-5 C and 3-6 C). Size exclusion chromatography indicated that both smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub> adopted dimer structures in solution like His-Cj0977.

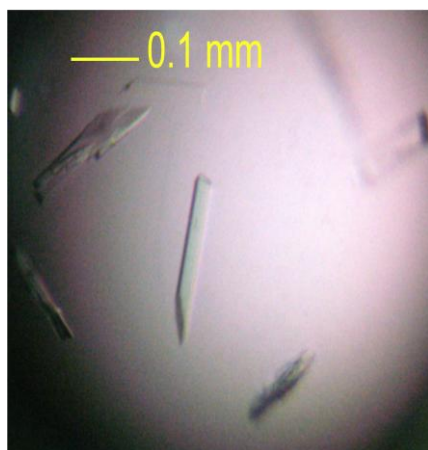
To confirm and compare the quaternary structure of Cj0977 and its variants, high-resolution gel filtration was performed using purified His-Cj0977 and smCj0977<sub>p17</sub> samples (Figure 3-7). Clearly, the elution volume of each protein indicated a dimer structure (His-Cj0977, ~46 kDa and smCj0977<sub>p17</sub>, ~35 kDa). Therefore, we confirmed that the truncation in smCj0977<sub>p17</sub> (and smCj0977<sub>p19</sub>) did not alter the quaternary structure and validated Cj0977 variants as good materials for crystallization.

Next, we pursued crystallization with smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub>. Interestingly, crystals of both smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub> grew in the same reservoir solution condition containing 24%-27% polyethylene glycol (PEG) 4000, 0.25 M ammonium acetate, and 0.1 M Tris-HCl, pH 7.0-7.4 (Figure 3-8). These crystals (for both proteins) in various shapes grew within three days, and needle clusters appeared more slowly in some crystallization drops. Some of these needle crystals grew very nicely to dimensions of 400  $\mu\text{m}$   $\times$  50  $\mu\text{m}$   $\times$  10  $\mu\text{m}$  within one week. Finally, we found one crystal of Cj0977<sub>p17</sub> (out of ~100 crystals tested for diffraction) that diffracted to a resolution of 2.6 Å.

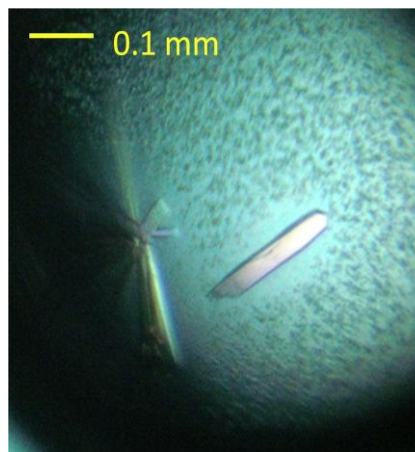


**Figure 3-7. Analytical gel filtration chromatograph.** The full length purified His-Cj0977 eluted with an estimated molecular weight of 46 kDa, a dimer structure. For comparison, the elution profile of Cj0977<sub>p17</sub> is overlaid. The purified Cj0977<sub>p17</sub> also eluted as a dimer with an estimated molecular mass of 35 kDa. The peak positions of molecular standard markers are shown as arrows (conalbumin = 75 kDa, ovalbumin = 43 kDa, and carbonic anhydrase = 29 kDa) (adapted from Yokoyama *et al.*, 2008).

(A)



(B)



**Figure 3-8. Crystals of smCj0977<sub>p19</sub> and smCj0977<sub>p17</sub>.** (A) Crystals of smCj0977<sub>p19</sub>. (B) Crystals of smCj0977<sub>p17</sub>. Both crystals were grown at 17°C by vapor diffusion in hanging drops against a reservoir solution containing 24%-27% PEG 4000 solution, 0.25 M ammonium acetate, and 0.1 M Tris-HCl, pH 7.0-7.4. Needle shaped crystals grew within one week from both proteins. One Cj0977<sub>p17</sub> crystal diffracted to a resolution of 2.6 Å.

### 3.3. Discussion

Intriguingly, the *cj0977* gene, which is not a flagellar gene, was controlled by  $\sigma^{28}$  promoter, like many other flagellar genes, suggesting its regulatory role linked to flagella and/or virulence (Yokoyama *et al.*, 2008). Subsequently, the corresponding protein Cj0977 was demonstrated as a virulence factor associated with invasion of intestinal epithelial cells *in vitro* (Goon *et al.*, 2006). In an effort to understand the mechanism, by which Cj0977 contributes to virulence of *C. jejuni* at the molecular level, we set out to solve the structure of Cj0977. My work focused on obtaining high quality crystals, a bottleneck step in protein crystallography. First, we used the full-length Cj0977 protein for crystallography work. Although the full-length protein yielded crystals, they did not diffract beyond 6 Å. Next, based on limited proteolysis as well as survey of protein stability, we designed several truncated versions of Cj0977. Among the different constructs, Cj0977<sub>p17</sub>, which includes residues 5 to 156 and corresponds to a full-length form of Ws0669 or Hp0420 (Figure 3-4), diffracted to 2.6 Å (at the peak wavelength) allowing MAD phasing. The structure determination and refinement steps were done in collaboration with Dr. Yokoyama in our laboratory.

#### 3.3.1. Cj0977 is a hot dog fold protein

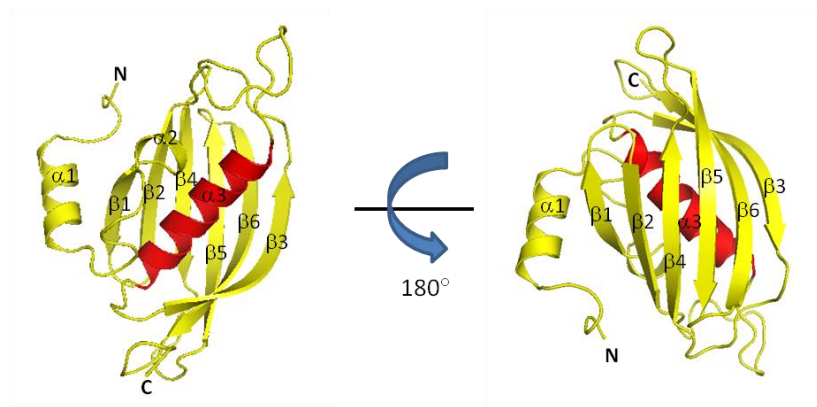
Despite no obvious clue from the primary sequence, the 3D structure of Cj0977<sub>p17</sub> revealed the ‘hot dog’ fold, one of the most ubiquitous protein folds discovered to date. The hot dog fold was first reported in the structure of *E. coli*  $\beta$ -hydroxydecanoyl thiol ester dehydratase (FabA) (Leesong *et al.*, 1996). This protein fold is associated with numerous coenzyme A derivative binding oligomeric enzymes (Dillon and Bateman, 2004). Consistent with the size exclusion chromatography result (Figure 3-6), the crystal

structure demonstrated that the biological assembly of Cj0977 is a dimer (containing two same subunits). The monomer structure of Cj0977<sub>p17</sub> is a typical mixed  $\alpha$  and  $\beta$  ‘hot dog’ fold and consists of six-stranded curved antiparallel  $\beta$  sheet  $\beta 1/\beta 2/\beta 4/\beta 5/\beta 6/\beta 3$  (bun) wrapping around the central helix,  $\alpha 3$  (sausage) (Figure 3-9). The monomer is a compact structure with dimensions of  $\sim 45 \text{ \AA} \times 35 \text{ \AA} \times 25 \text{ \AA}$ . Upon dimerization, two 6-stranded  $\beta$  sheets of each subunit are joined to form a large anti-parallel 12-stranded  $\beta$  sheet for the functional quaternary structure (Figure 3-9). Since Cj0977 is a structural homolog of proteins that bind coenzyme A derivatives, this study have revealed a clue of the function, as we initially hypothesized.

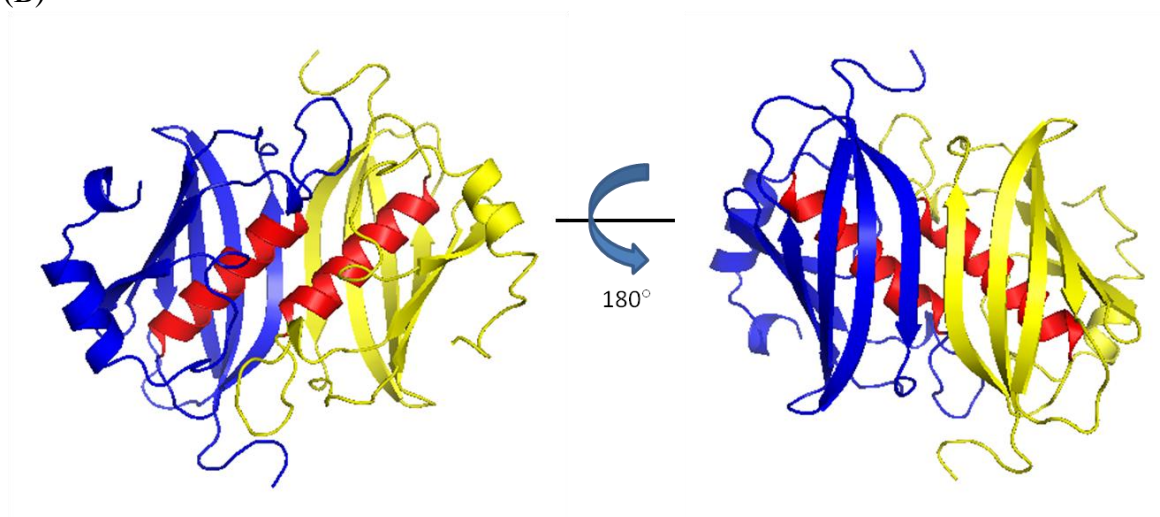
### **3.3.2. Structural comparison and putative binding site for acyl-CoA**

Once a new protein structure is determined, one can search its close structural homologs using the DALI server, a network tool for protein structure comparison (Holm and Sander, 1995). Our DALI search for structural homologs to Cj0977 identified over 20 proteins with a Z-score of  $\geq 8$ . The DALI server is a network service that compares protein structures in 3D; a query protein structure is compared against those in the Protein Data Bank (PDB) and the structural similarity is measured by Z-score. Structures that have a Z-score above 2 commonly have similar folds. The majority of these proteins are acyl-CoA thioesterases with differing substrate specificities from various organisms. Unexpectedly, however, the structure of Cj0977<sub>p17</sub> was the closest to that of *Bacillus subtilis* FapR with a Z-score of 17. FapR, not a thioesterase, is a global transcriptional repressor that controls the expression of many genes involved in fatty acids and phospholipids biosynthesis in *B. subtilis*. Binding of malonyl-CoA, an essential

(A)



(B)

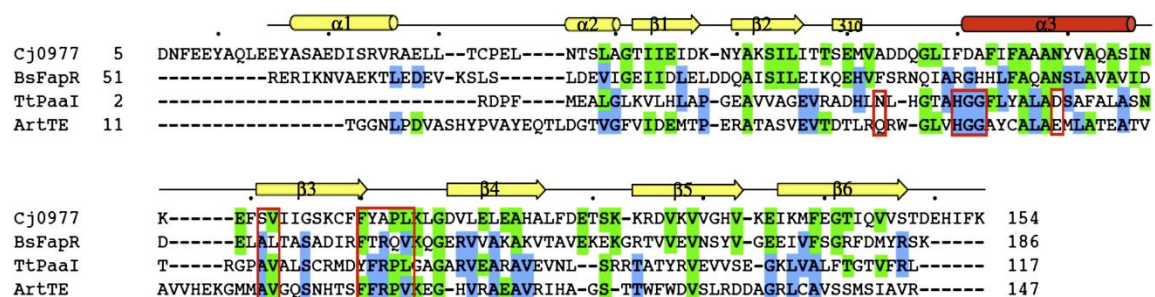


**Figure 3-9. Crystal structure of Cj0977<sub>p17</sub>.** (A) Ribbon diagrams of the Cj0977<sub>p17</sub> monomer structure. The secondary structures are labeled. “N” and “C” represent the N-terminus and C-terminus of Cj0977<sub>p17</sub>, respectively. The central helix (α3) is highlighted in red. (B) Ribbon diagrams of the Cj0977<sub>p17</sub> dimer structure (PDB code: 3BNV). Two subunits are shown in blue and yellow (adapted from Yokoyama *et al.*, 2008).

intermediate in fatty acid biosynthesis, to FapR induces conformational changes in the hot dog domain of FapR, which propagates to helix-turn-helix (HTH) motifs of FapR dimer, and therefore hinders productive association between FapR and its specific DNA promoter (Schujman *et al.*, 2006). Like Cj0977, FapR also adopts a dimer structure and these two proteins superimpose well throughout the hot dog fold domain, except that the N-terminal helices orient differently (Yokoyama *et al.*, 2008). Like crystallization of Cj0977, where the C-terminal 36 residues portion of Cj0977 was truncated in Cj0977<sub>p17</sub>, the crystal of FapR was obtained by deleting N-terminal 43 residues (named FapR<sub>D43</sub>). The N-terminal 43 residues domain of FapR represents HTH motif involved in DNA binding. By analogy, we speculate that this small domain (the C-terminal 36 residues portion of Cj0977) plays a role in Cj0977 virulence by DNA binding or interaction with other proteins (see 3.3.3.). In fact, the C-terminal 36 amino acid moiety was required for invasion of *C. jejuni* into INT407 cells (Yokoyama *et al.*, 2008). Other significant structural homologs to Cj0977 included *Thermus thermophilus* phenylacetate thioesterase (TtPaaI), and *Arthrobacter* sp. 4-hydroxybenzoyl-CoA thioesterase (ArHBT) (Thoden *et al.*, 2003 and Kunishima *et al.*, 2005). The PaaI thioesterase is involved in bacterial phenylacetic acid catabolic pathway, and HBT catalyzes the hydrolysis of the thioester moiety to yield 4-hydroxybenzoate and CoA.

Our careful inspection of the superimposed structures indicated some clues as to the function of Cj0977. The liganded forms of all three proteins indicated residues involved in recognition of the common structural moiety of acyl-CoA derivatives, 4-phosphopantetheine. The binding pocket appears to be conserved in Cj0977 with consensus motifs S-V and F-Y-A-P-L (Figure 3-10). However, the prominently





**Figure 3-10. Secondary structure of Cj0977<sub>p17</sub> and structure-based sequence alignment.** Secondary structure elements of Cj0977<sub>p17</sub> are shown at the top of the sequences.  $\alpha$ -helices and  $\beta$ -strands are denoted as cylinders and arrows, respectively (central helix  $\alpha 3$  in red). BsFapR, TtPaaI, and ArtTE represent *B. subtilis* FapR, *T. thermophilus* PaaI, and *Arthrobacter* sp. 4-hydroxybenzoyl-CoA thioesterase, respectively. Green and blue shadings highlight residues identical with the Cj0977 sequence and identical residues among hot dog fold proteins except Cj0977, respectively. Key residues involved in ligand recognition and/or catalysis are highlighted by red boxes (adapted from Yokoyama *et al.*, 2008).

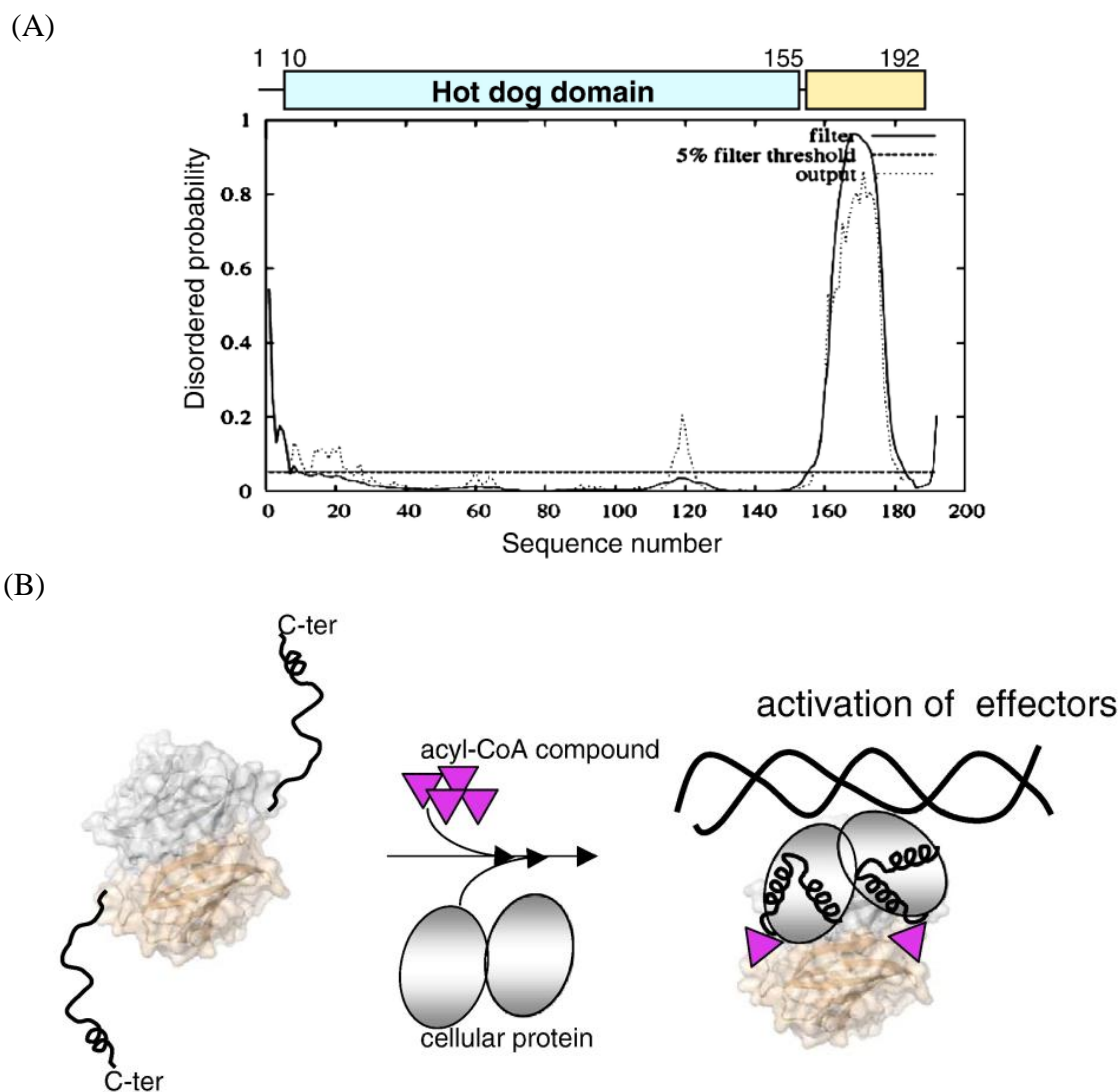
conserved residues in these subfamily thioesterases, HGG triad and catalytic base Asp or Glu, are completely lacking in Cj0977, indicating an unlikely function of Cj0977 as a thioesterase. As expected, we could not detect any significant enzyme activity with Cj0977 (data not shown). In considering moderate similarity to known binding pockets of the acyl chain moiety, we also initiated studies to examine binding properties of acyl-CoA compounds to Cj0977. Cj0977 contains Phe70 in the place of Arg106 of FapR, which is essential for interacting with malonyl carboxylate, yet Phe70 is indispensable for virulence (Yokoyama *et al.*, 2008). Considering a relatively small binding pocket in the Cj0977 structure, we speculate that the nature of the acyl moiety of the ligand would be a small hydrophobic group. It would be interesting to investigate the possibility that Cj0977 has a similar function as FapR, as a transcriptional regulator sensing an acyl-CoA intermediate of membrane lipid (or fatty acid) biosynthesis *via* the Cj0977 hot dog domain.

### **3.3.3. Proposed model of coupled folding and binding of Cj0977**

Many gene sequences, especially in eukaryotic genomes, encode entire proteins or large segments of proteins that lack a well-structured three-dimensional fold (Dyson and Wright, 2005). The existence of unstructured regions of significant size (>50 residues) is common in functional proteins. Functions of intrinsically disordered protein domains include the regulation of transcription and translation, cellular signal transduction, protein phosphorylation, and regulation of the self-assembly of large multiprotein complexes such as the bacterial flagellum. Many of these proteins fold on binding to their targets - that is, they undergo coupled folding and binding processes. Since the C-terminal 36 residue tail of Cj0977 was required for invasion of *C. jejuni* into INT407 cells, we

examined this tail sequence. Consistent with our limited protease experiment, this segment (rich in polar residues and very low composition of hydrophobic residues) appeared to be disordered. Indeed, the DISOPRED2 program (Ward *et al.*, 2004) predicted the Cj0977 sequence of disorder propensity as a completely disordered region at the C-terminal 36 residues of Cj0977 (Figure 3-11 A). From these results, we proposed that this segment of 36 amino acids is unstable and unstructured in solution and undergoes coupled folding and binding to an as yet unidentified cellular *C. jejuni* protein, which might be involved in transcriptional regulation (Figure 3-11 B) (Yokoyama *et al.*, 2008). Taken together, we proposed a model that this C-terminal region of Cj0977 might acquire structure only after binding to its partner, and an acyl-CoA intermediate of fatty acid biosynthesis might promote or hinder the favorable conformational change of this region (Yokoyama *et al.*, 2008).

In conclusion, our structural study of Cj0977 provided new clues in discovering a novel acyl-CoA effector as well as in studying the seemingly disordered C-terminal region. Since a similar acyl-CoA binding pocket was identified in Cj0977, exploring the binding properties of Cj0977 with various CoA compounds may give a clue to find its ligand(s). Furthermore, DNA binding properties of Cj0977 can be examined using *C. jejuni* genomic DNA by DNase footprinting assay. Together with the structure of Cj0977 bound to its CoA derivative ligand, the ordered structure promoted by the interaction with a cellular protein partner may aid in crystallization. These studies would indicate a novel linkage between flagella regulation, fatty acid biosynthesis and virulence in *C. jejuni*, a significant step toward the understanding of the structural basis for *C. jejuni* virulence.



**Figure 3-11. Domain architecture and proposed model of coupled folding and binding of Cj0977.** (A) Top, Domain organization of Cj0977 showing the main structural and functional domain (hot dog fold domain) and the C-terminal tail residues 157-192. Bottom, Disordered profile plot generated by the DISOPRED2 Disorder Prediction Server (Ward *et al.*, 2004). (B) The hot dog fold domain of Cj0977 binds as yet unidentified CoA derivative ligands (shown as triangles). Ligand binding may facilitate the initial transition from the disordered state to the ordered state of the C-terminal tail (shown as black strings) of Cj0977. By interacting with cellular protein(s), the C-terminal region is now folded and plays a role as a regulator, such as binding to DNA or other proteins (adapted from Yokoyama *et al.*, 2008). “C-ter” denotes the C-terminus of Cj0977.

## **Chapter IV. Structural Studies of the Lipoprotein Adhesin JlpA**

Contributed to:

Kawai, F., Paek, S., Choi, K.J., Prouty, M., Kanipes, M.I., Guerry, P., and Yeo, H.J.

(2012) Crystal structure of JlpA, a surface-exposed lipoprotein adhesin of *Campylobacter jejuni*. *J. Struct. Biol.* **177**, 583-588.

#### 4.1. Introduction

Adherence of bacteria to host epithelial cells is a critical step in a bacterial infection. Once ingested by human via contaminated food, *C. jejuni* survives the stresses in the stomach, colonizes the small intestine in early stages of infection, and then moves to the large intestine, the target organ of disease (Black *et al.*, 1988). After successful colonization of the mucus lining of the intestine, *C. jejuni* can adhere to epithelial cells (Poly and Guerry, 2008). Adherence of *C. jejuni* is a multifactorial process, in which multiple binding factors may be required to bind to their specific receptors to attain an efficient interaction with host cells. A number of *C. jejuni* proteins including PEB1, MOMP, CadF, CapA, and JlpA have been reported to bind to cultured epithelial cells, as detailed in Chapter I. Disruption of the *peb1A* gene encoding PEB1 showed reduced *C. jejuni* adherence to human HeLa cells by 50 to 100 fold (Pei *et al.*, 1998). The outer membrane protein MOMP encoded by the *porA* gene was observed to bind to INT407 cells (Moser *et al.*, 1997). CadF appears to bind to cell matrix protein fibronectin, as a *cadF* mutant showed a 50% reduction in adhesion to human INT 407 cells compared to a wild-type *C. jejuni* isolate (Monteville *et al.*, 2003). CapA, a putative autotransporter, has been reported to play a role in adherence and invasion of Caco-2 cells *in vitro* (Ashgar *et al.*, 2007). However, the relative contribution as well as significance of these factors to disease remains uncertain, given the absence of small animal models of diarrhea for *C. jejuni*.

Another putative adhesin is the JlpA protein, a surface-exposed lipoprotein (Jin *et al.*, 2001 and Jin *et al.*, 2003). The *jlpA* gene encodes JlpA of 372 amino acid residues with a molecular mass of 42.3 kDa. JlpA contains a typical signal peptide (18 amino acid

residues) at the N-terminus and a lipobox motif [L-F-S-A-C] at the distal end of the signal peptide. Earlier work demonstrated that the JlpA molecule contains fatty acids as confirmed by incorporation of [<sup>3</sup>H]-palmitic acid and binds to HEp-2 cells, establishing JlpA as a novel *C. jejuni* lipoprotein adhesin virulence factor (Jin *et al.*, 2001). A *jlpa* mutant showed about 19% adherence to human HEp-2 cells relative to the wild-type strain, although a more recent study failed to detect any JlpA-mediated binding to chicken cells *in vitro* (Jin *et al.*, 2001 and Flanagan *et al.*, 2009). Another study reported that JlpA interacts with HEp-2 cell surface heat shock protein 90α (Hsp90α) and initiates signaling pathways leading to activation of NF-κB and p38 MAP kinase, suggesting that JlpA contributes to the inflammatory responses associated with *C. jejuni* infection (Jin *et al.*, 2003). Together with several genes encoding known *C. jejuni* virulence factors, the *jlpa* gene was found to be upregulated in response to human mucin MUC2, the most abundant secreted mucin in the human intestine and a major chemoattractant for *C. jejuni*, suggesting an important role of JlpA in pathogenicity (Tu *et al.*, 2008). Moreover, recent work has shown that JlpA is a glycoprotein and is immunogenic during human infection (Scott *et al.*, 2009). Of note, impaired glycosylation in *C. jejuni* resulted in reduction in adherence to and invasion of INT407 cells *in vitro* and a reduced ability to colonize the intestinal tract of mice (Szymanski *et al.*, 2002).

Considered as a surface-exposed lipoprotein and immunogenic adhesin protein, JlpA is of significant therapeutic interest as a potential vaccine target. Like the Cj0977 protein (Chapter III), JlpA is highly conserved within *Campylobacter* spp.. Also, the crystal structure of JlpA is not available and JlpA does not show any significant sequence homology with other proteins with a known function. To understand the structural

determinants of JlpA mediated adherence and gain insights into the function of JlpA, we set out to solve the crystal structure of JlpA. In this chapter, we present essential experimental steps for the structure determination. One key experiment for solving the structure was improving MAD phasing using crystals of the JlpA protein that were modified to contain two additional Met sites. We also discuss the first view of the crystal structure and how JlpA serves as a lipoprotein adhesin.

## **4.2. Results**

### **4.2.1. Purification and crystallization of JlpA**

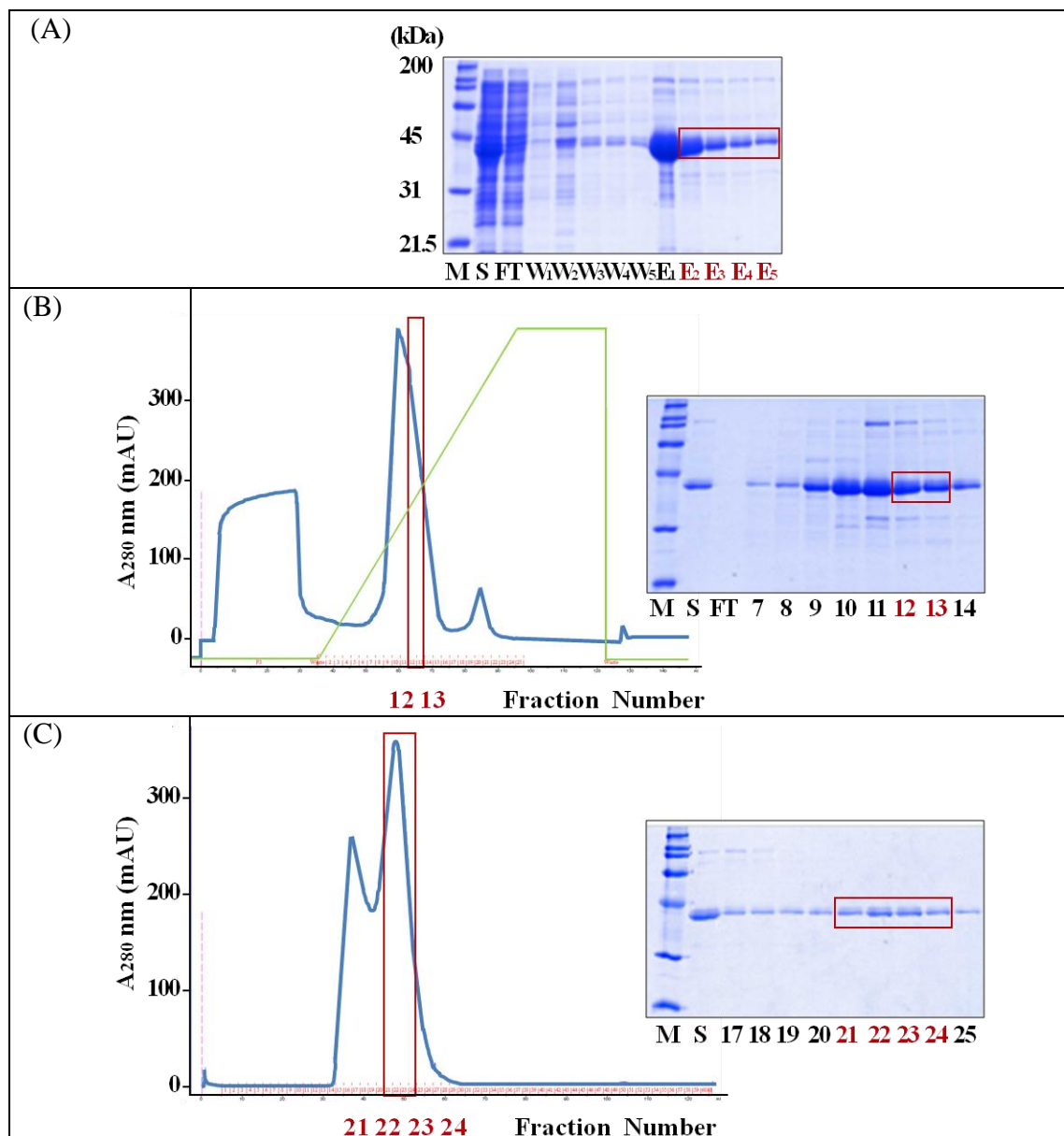
The objective of this study was to establish a purification method for JlpA and obtain high quality diffracting JlpA crystals. The mature form of JlpA (amino acid residues 18 to 372) lacking the N-terminal signal peptide was produced as an N-terminal His-tagged protein (His-JlpA) and pursued for crystallization.

The His-JlpA protein was well expressed and soluble in the BL21(DE3) strain. The first purification step of nickel affinity chromatography (Ni-NTA agarose) resulted in a broad elution profile with each fraction containing significant impurities, compared with the elution profile of Cj0977 that indicated over 90% purity after the Ni-NTA column step (Chapter III, Figure 3-1 A). This unfavorable behavior of JlpA in purification is probably due to the nature of the protein: JlpA is a lipoprotein adhesin (the molecule may be sticky by nature) anchored to the outer membrane, compared with the cytoplasmic Cj0977 protein. To improve purification, several parameters were adjusted in the cell lysis step. The addition of 0.1% triton-X 100 and 5 mM  $\beta$ -mercaptoethanol to lysis buffer resulted in an improved elution profile. The His-JlpA protein was eluted at an



imidazole concentration of 150 mM (elution E1-E5) when using a step gradient method (Figure 4-1 A). The theoretical molecular weight of JlpA (excluding signal peptide sequence and including the His-tag and additional residues integrated for cloning purpose) is 43 kDa. The target protein was observed with apparent molecular weight of 43 kDa on the gel suggesting a normal electrophoretic mobility of the protein. Among different purification trials, we found that collecting fractions E2-E5 only (excluding E1) for the next step was important for the final quality of the purified sample. Considering the characteristic high net negative charge of JlpA (pI 4.7) at pH 8.5, His-JlpA was further purified using an anion exchange column (HiTrap Q) on AKTA Purifier. The JlpA protein bound tightly to the column and a single broad elution peak was observed between 330 mM and 470 mM of NaCl (with the highest peak at ~400 mM) (Figure 4-1 B). The fractions 12 and 13 were selected and combined for the final purification step. Two peaks were observed on the gel filtration chromatogram. The fractions around the peaks were analyzed on SDS-PAGE, showing the presence of JlpA in both peak fractions. The first peak contained an aggregated form of JlpA and other high molecular non-specific proteins. The elution volume of the second peak centered on around 48 mL corresponded to ~43 kDa, suggesting that JlpA adopts a monomer structure in solution (Figure 4-1 C). Subsequently, clean elution fractions (21 to 24) were concentrated for crystallization purpose. We found this protein was moderately soluble, and could routinely be concentrated to ~7 mg/mL.

Although this three-step purification approach resulted in moderately pure protein sample (~85% purity) compared with the Cj0977 (>95% purity), one crystallization condition was identified from the initial screening of native His-JlpA. The crystallization



**Figure 4-1. Purification of His-JlpA.** (A) SDS-PAGE analysis of nickel affinity chromatography (15% gel). The elution fractions E2-E5 were combined for the anion exchange chromatography. (B) Anion exchange chromatogram and SDS-PAGE analysis. A wide peak of JlpA was observed between 350 mM and 500 mM of NaCl. Due to impurities, only fractions 12 and 13 were used for the gel filtration chromatography. (C) Gel filtration chromatogram and SDS-PAGE analysis. JlpA was eluted as two pools (a large aggregated form and a monomeric form), and the second peak corresponded to the molecular mass of 43 kDa. The fractions 21-24 were collected for crystallization. M, Marker; S, Sample; FT, Flow-through; W, Wash; and E, Elution.

condition, 30% (v/v) 2-methyl-2,4-pentanediol (MPD), 0.02 M  $\text{CaCl}_2$ , and 0.1 M Na-acetate, was further optimized for crystal growth. Sizable JlpA crystals (0.15 x 0.15 x 0.1 mm) were grown at 17°C by vapor diffusion in hanging drops against a reservoir solution containing 32-35% MPD, 0.03 M  $\text{CaCl}_2$  and 0.1 M Na-acetate, pH 4.2-4.6 (Figure 4-2 A). JlpA crystals were observed within a week and grew for several weeks. For MAD phasing purpose, selenomethionine-substituted JlpA (smHis-JlpA) was also produced using the same purification steps of native His-JlpA and crystallized in the same condition as native His-JlpA (Figure 4-2 B). These JlpA crystals grown with MPD precipitant turned out to be extremely fragile and sensitive to vibration and temperature, which made crystal handling very challenging.

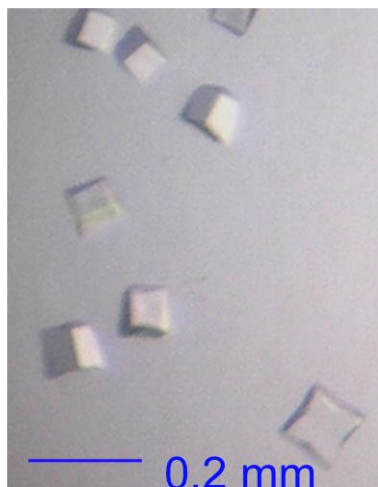
#### **4.2.2. Electron density map of JlpA crystals**

Initially, data collection, processing, and MAD phasing of JlpA crystals were performed in collaboration with Dr. Yokoyama in our laboratory. In general, JlpA crystals did not diffract well. Among many diffraction trials, one smHis-JlpA crystal diffracted to a resolution of 3.3 Å. As shown in Figure 4-3, the electron density map revealed that the JlpA structure contains many  $\beta$ -sheets, although we could not see the structure in detail at this resolution level (3.3 Å). These observations required phase improvement to solve the protein structure, which can be achieved by generating other forms of JlpA crystals (for example, JlpA crystals in different space groups or crystals of modified JlpA).

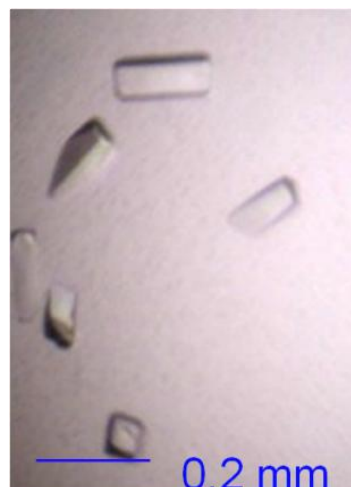
#### **4.2.3. Strategy for phase improvement: Introduction of methionine**

Despite an exhaustive crystallization screening, JlpA was found to be crystallized under only one condition (with MPD solution), as described above. Therefore, we turned

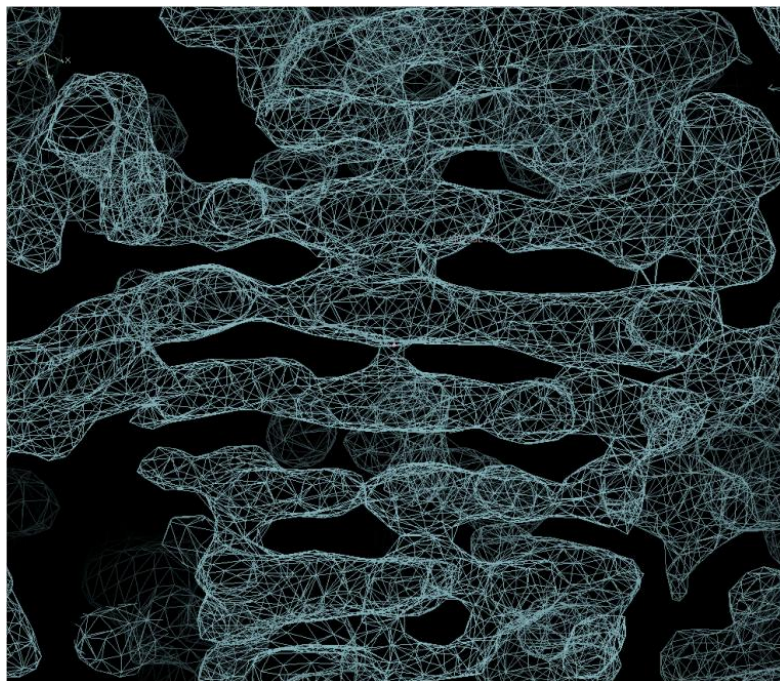
(A)



(B)



**Figure 4-2. JlpA crystals.** (A) native His-JlpA crystals. (B) smHis-JlpA crystals. Cubic shaped crystals appeared within a week and grew for several weeks. Both native and selenomethionine-derivatized His-JlpA crystals were grown at 17°C by vapor diffusion in hanging drops against a reservoir solution containing MPD, CaCl<sub>2</sub>, and Na-acetate, pH 4.2-4.6.

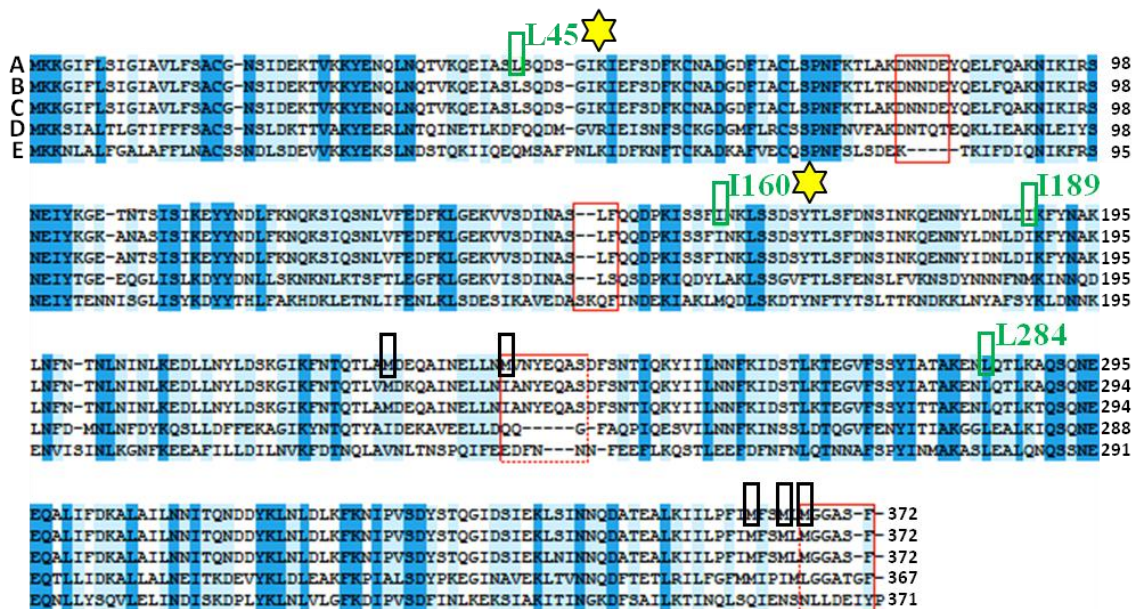


**Figure 4-3. Initial electron density map of a JlpA crystal.** JlpA crystals diffracted to 3.3 Å resolution. This view of the electron density map reveals the JlpA structure contains many  $\beta$ -sheets.

to protein engineering. Based on our previous successful experience with Cj0977 (Chapter III), our first choice was to produce GST-JlpA fusion proteins and subsequently remove the GST moiety to generate a tag-free JlpA protein. By applying the same strategy on JlpA, we hoped to grow well-ordered high quality crystals. While we could purify the JlpA protein without any affinity tag (even with a higher purity, >90%), we never observed a crystal of JlpA arising from the GST-fusion construct.

Next, we turned to another strategy: to introduce additional methionine sites in the JlpA sequence. For successful MAD phasing using ‘Se’ anomalous dispersion, one methionine site per every 100 residues is typically necessary. As JlpA contains 372 amino acid residues in total, at least 4 methionine sites would be necessary to obtain a good signal in MAD phasing. There are five methionine (Met) sites (Met-228, Met-239, Met-362, Met-365, and Met-367), excluding the first Met site within the JlpA sequence (from strain 81-176) (Figure 4-4). However, three Met sites (Met-362, Met-365, and Met-367) are clustered in the C-terminal region of JlpA, suggesting uncertain occupancy of ‘Se’ sites for these Met residues in MAD phasing. These three Met sites may not contribute much for structure determination, since it may be a flexible region not involved in core structure (often the case in a protein structure). Consequently, only two Met sites (Met-228 and Met-239) would contribute to the ‘Se’ anomalous signal in MAD.

To improve phasing (or increase the signal), additional Met sites were introduced within the JlpA sequence. Four sites, L45, I160, I189, and L284, were selected as Met introduction sites (Figure 4-4). These residues were chosen because both Leu and Ile are hydrophobic amino acids like Met, and thus these residues are interchangeable among homologous proteins. Very likely, replacing these Leu or Ile residues of JlpA with Met



**Figure 4-4. Sequence alignment of JlpA homologs.** Alignment of JlpA from multiple strains of *C. jejuni* and other *Campylobacter* spp. illustrates remarkable sequence conservation. The original five Met sites in JlpA (from strain 81-176) are indicated with black boxes. Four sites (Leu or Ile) selected for ‘Met’ mutation are indicated with green boxes. The two Met introduced sites yielding the crystal structure are highlighted with ‘★’ (adapted from Kawai *et al.*, 2012). A, *C. jejuni* 81-176; B, *C. jejuni* NCTC 11168; C, *C. jejuni* CG8486; D, *C. coli* RM2228; and E, *C. upsaliensis* RM3195.



would not affect the overall structure of JlpA. Indeed, the sequence alignment showed that L160 and I189 in JlpA from *C. jejuni* NCTC 11168 are replaced with Met in *C. upsaliensis* RM3195 (M160) and *C. coli* RM2228 (M189), respectively (Figure 4-4). However, since we did not know the 3D structure of JlpA, we could not predict with confidence the mutation effect of these sites on the solubility and the structure of the protein during purification, which could be determined empirically. Thus, three sites L45, I160, and L284 (mutagenesis of I189 was not successful) were targeted for the Met introduction and four JlpA variants containing two to three additional Met sites were produced (JlpA<sub>L45M/I160M</sub>, JlpA<sub>L45M/L284M</sub>, JlpA<sub>I160M/L284M</sub>, and JlpA<sub>L45M/I160M/L284M</sub>).

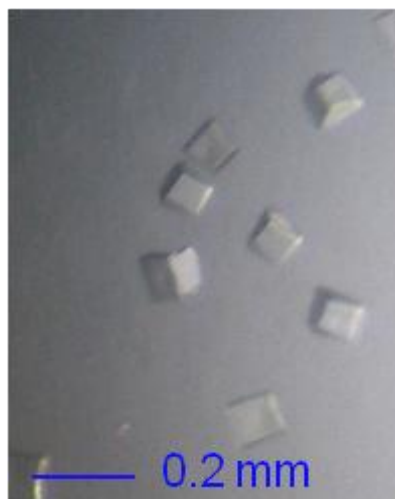
#### **4.2.4. Purification and crystallization of JlpA variants**

All four variants were expressed as N-terminal His-JlpA. Purification and crystallization were performed using the same procedure as for native His-JlpA (data not shown), as described earlier (see 4.2.1.). Despite the Met introduction into the sequence, all variants were successfully expressed and purified well. The variants behaved similarly to native His-JlpA in terms of solubility and stability during and after purification. Interestingly, although JlpA underwent alterations in sequence (two or three residues), these JlpA variants were all crystallized in the same crystallization condition as the native His-JlpA protein, suggesting that the additional Met residues did not significantly affect crystal packing and stability.

Moreover, like His-JlpA crystals, cubic shaped JlpA variant crystals were also observed within a week and grew for several weeks. Crystals of selenomethionine-derivatized JlpA variants were also produced for MAD phasing. A crystal from one of



smJlpA variants, smJlpA<sub>L45M/I160M</sub>, diffracted to 2.7 Å, allowing improved phasing (Figure 4-5).



**Figure 4-5. smHis-JlpA<sub>L45M/I160M</sub> crystals.** Crystals of all JlpA variants were grown in a reservoir solution containing 32-35% MPD, 0.03M CaCl<sub>2</sub> and 100 mM Na-acetate, pH 4.2-4.6. A crystal from JlpA<sub>L45M/I160M</sub> diffracted to 2.7 Å, allowing a successful MAD phasing.

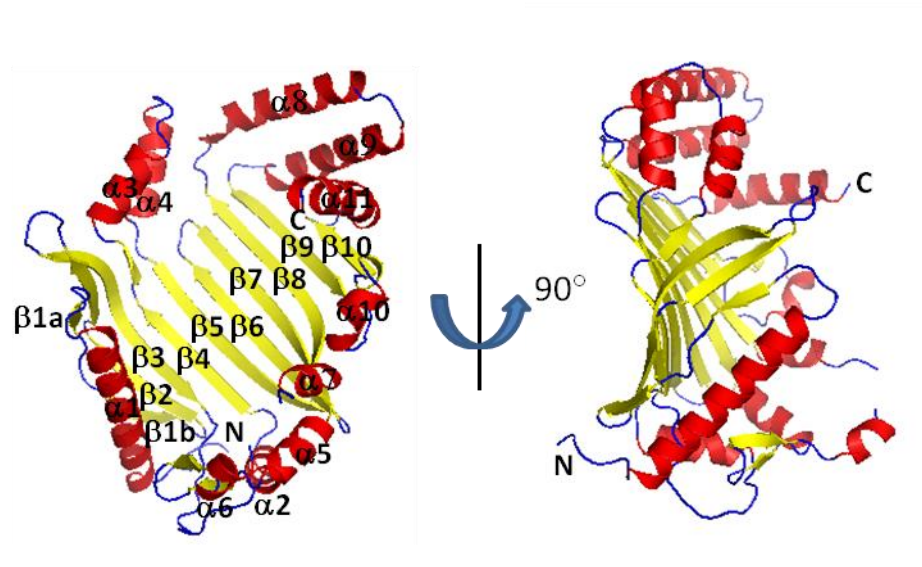
### 4.3. Discussion

JlpA is a *Campylobacter* specific cell surface-exposed lipoprotein and an adhesin virulence factor (Jin *et al.*, 2001). To gain insight into the structural basis of *C. jejuni* virulence, we aimed to determine the crystal structure of JlpA. My work in this chapter focused on establishing a purification method and obtaining high quality crystals, a bottleneck step in protein crystallography. The mature form of JlpA (Cys-18 to Phe-372) without the signal peptide was overexpressed for crystallization purpose. In contrast to Cj0977 (highly soluble protein concentrated to ~45 mg/mL), the solubility of JlpA was limited (concentrated to ~7 mg/mL), and the final protein sample was moderately pure (~85%) after three chromatography steps. Nonetheless, we succeeded in growing crystals. The native JlpA crystal diffracted to a resolution of 3.3 Å, but a structure model building was impossible because of inadequate phasing. Next, we constructed JlpA variants by introducing additional Met sites in the JlpA sequence to improve MAD phasing. We were able to grow sizable crystals from different JlpA variants, and collected a MAD data set to a resolution of 2.7 Å using a crystal of JlpA<sub>L45M/I160M</sub>. In addition to an improved resolution, two additional Met sites in the protein were very useful for building the structure model. The subsequent structure determination and refinement steps were done in collaboration with Dr. Kawai in our laboratory.

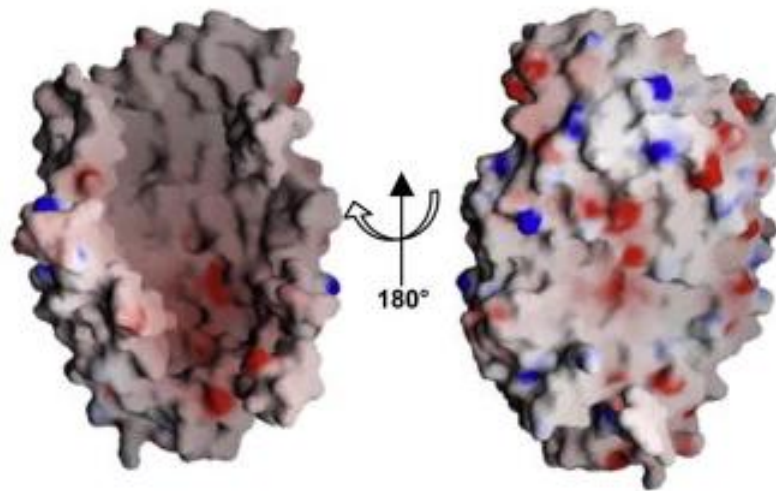
#### 4.3.1. Molecular architecture of JlpA

The JlpA crystal shows a single domain structure with a  $\alpha/\beta$ -fold composed of curved 10-stranded antiparallel  $\beta$ -sheets, which form a large central  $\beta$  sheet, and 11  $\alpha$ -helices with 3 small strands, which encircle the central  $\beta$  sheet (Figure 4-6 A). The large  $\beta$ -sheet has the strand order of  $\beta$ 1 to  $\beta$ 10. The overall structure of JlpA looks like a

(A)



(B)



**Figure 4-6. Structure of JlpA.** (A) Cartoon representation of two views of JlpA (PDB code: 3UAU). The secondary structural elements are labeled. “N” and “C” indicate the N-terminus and C-terminus of JlpA, respectively. JlpA has the strand order of  $\beta 1$  to  $\beta 10$ . (B) Surface electrostatic potential of JlpA. The surface was contoured and displayed using the program GRASP (Nicholls *et al.*, 1991). Color-coding is according to charge, with blue for the most positive regions, red for the most negative regions, and linear interpolation in between (adapted from Kawai *et al.*, 2012).

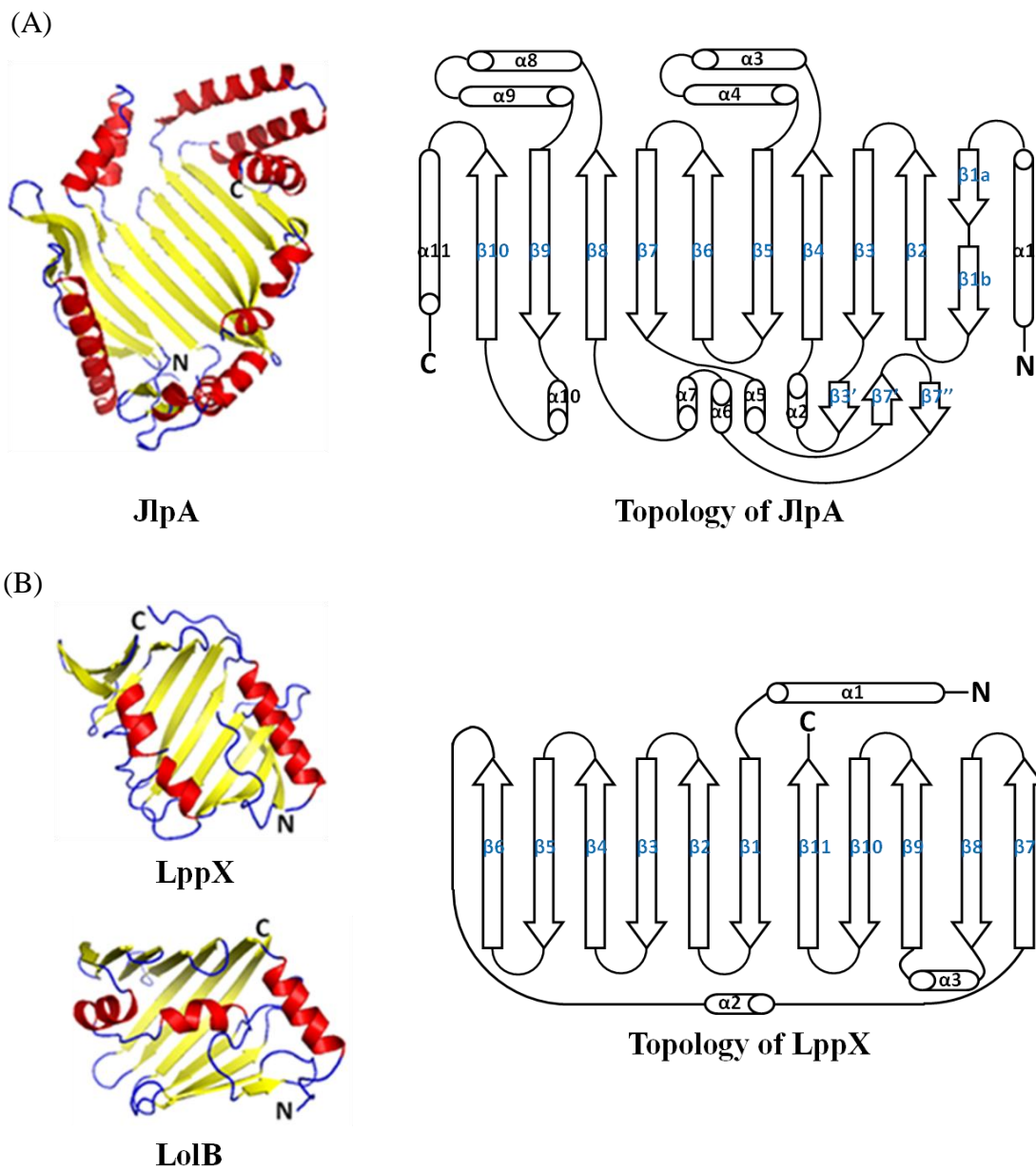
‘catcher’s mitt’ shape with dimension of  $\sim 80 \times 50 \times 25 \text{ \AA}$  (Figure 4-6 B). One remarkable feature of the JlpA structure is the partition of the molecular surface: while the concave side forms a wide hydrophobic basin with a localized acidic pocket, the convex face reveals a polar surface throughout (Figure 4-6 B). While the floor of the concave face made by the major  $\beta$ -sheet is rather flat, all  $\alpha$ -helices together with loop regions and three minor  $\beta$ -strands are arranged to create the basin walls. The hydrophobic basin is rich in Phe residues and results in a surface area of  $\sim 4041 \text{ \AA}^2$  and a volume of  $\sim 13,197 \text{ \AA}^3$ . Indeed, out of total 22 Phe residues of the mature form of JlpA, 11 Phe residues lie in the basin (Kawai *et al.*, 2012). The hydrophobic basin is empty in the crystal, which might explain the fragile characteristic of JlpA crystals in the crystallization drops. Consistent with the gel filtration result, the functional assembly of JlpA is a monomer in the crystal structure.

At the primary structure level, JlpA does not have any significant homology to proteins from other organisms. Although DALI searches for structural homologs of JlpA revealed a few proteins with Z-scores above 7, all of these proteins were aligned to JlpA with an RMSD over 4. Thus, it is unlikely that these structural homologs are functionally relevant. Indeed, all of the top-scored proteins as structural homologs were large bacterial outer membrane proteins such as FauA, FhaC, and OmpG reflecting the  $\beta$ -barrel like topology of JlpA (unclosed half barrel). On the other hand, from literature surveys, the resemblance to several bacterial lipoproteins, such as LolA, LolB and LppX, was striking. LolA (22.6 kDa) and LolB (23.5 kDa) are lipoprotein localization factors in *E. coli* (Matsuyama *et al.*, 1995 and Matsuyama *et al.*, 1997). LppX (24 kDa) is a lipoprotein of *Mycobacterium tuberculosis*, required for the translocation of complex lipids such as

phthiocerol dimycocerosates (DIM) (Sulzenbacher *et al.*, 2006). These bacterial lipoproteins are much smaller than JlpA (42 kDa) and have a common  $\beta$ -sheet topology with the strand order  $\beta 7/\beta 8/\beta 9/\beta 10/\beta 11/\beta 1/\beta 2/\beta 3/\beta 4/\beta 5/\beta 6$ . On the other hand, JlpA has the strand order of  $\beta 1$  to  $\beta 10$  (Figure 4-7 A and B). Remarkably, despite the absence of sequence homology, the apparent structure of JlpA is reminiscent of these bacterial lipoproteins. LolA, LolB, and LppX also have a hydrophobic cavity in the concave face of the molecule, similar to JlpA. Since JlpA is larger than these proteins, the hydrophobic surface is more important than for these lipoproteins. These observations lead us to speculate a role for JlpA as a *Campylobacter*-specific transporter for lipids (or lipoproteins) in addition to a role in adhesion.

#### **4.3.2. Structure-function relationship of JlpA**

JlpA was discovered as an adhesin by promoting adherence to epithelial cells and is known to be surface-exposed and lipid-anchored to the bacterial outer membrane. Indeed, the structure suggests an adhesin function, as revealed by a large hydrophobic basin and several flexible surface regions. The predicted lipid-modified N-terminus of JlpA is positioned at one end of the major axis of the central  $\beta$ -sheet (Figure 4-6). Structural aspects, such as protruding flexible surface loops and the location of the N-terminus, suggest the positioning of JlpA in the outer membrane such that these flexible surface loops are oriented toward the extracellular space (Kawai *et al.*, 2012). *C. jejuni* expresses a phase-variable polysaccharide capsule (Bacon *et al.*, 2001). Since only five unstructured N-terminal residues are available for spacing between the lapidated cysteine residue and helix  $\alpha 1$ , JlpA would not have enough capacity to protrude through the bacterial cell surface when the polysaccharide capsule is expressed. Possibly JlpA is



**Figure 4-7. JlpA and other bacterial lipoproteins.** (A) Ribbon diagram (left) and topology (right) of JlpA. (B) Ribbon diagrams of LppX and LolB (left) and topology of LppX (right). JlpA is distinguished in topology compare with other lipoproteins.

not extended beyond the polysaccharide layer, which is reminiscent of other bacterial antigenic outer membrane proteins such as the Neisserial surface protein A (NspA). NspA is a homolog of the opacity-associated (Opa) adhesin proteins that mediate adherence into host cells (Vandeputte-Rutten *et al.*, 2003). Moreover, NspA (a  $\beta$ -barrel outer membrane protein) presents a hydrophobic cleft surface at the extracellular side. Along these observations, it is conceivable that JlpA may act as an antigen and an adhesin analog to antigenic outer membrane proteins, but without being an integral membrane protein.

Adhesive activity of JlpA was previously studied by using HEp-2 cells and a recombinant GST-fusion form of JlpA (Jin *et al.*, 2001). In collaboration with Dr. Guerry, we have investigated JlpA binding to the more biologically relevant human intestinal INT407 cells. To identify key residues for ligand binding, we performed structure-guided mutagenesis, but we were unable to define any single residue required for binding to INT407 cells *in vitro* (data not shown). In addition, His-tagged JlpA was fluorescently labeled and added in increasing concentrations to INT407 cells in suspension and binding was monitored by flow cytometry (Kawai *et al.*, 2012). Unexpectedly, concentrations of JlpA as high as 300  $\mu\text{g/mL}$  failed to reach saturation and 100-fold excess of unlabeled protein was unable to compete for binding with labeled JlpA, while as a control a recombinant form of the B subunit of the heat labile enterotoxin (LTB) of enterotoxigenic *E. coli* showed saturation in binding under similar conditions. These observations suggest that JlpA may possess unusual avidity and may accommodate multiple ligands. Considering a large concave surface in the JlpA structure, we speculate that JlpA may contain multiple ligand-binding sites, which may make functional probing by



mutagenesis of a single residue ineffective. Further investigations are necessary to decipher the molecular function of the protein.

In conclusion, we determined the first high-resolution structure of a *C. jejuni* surface exposed lipoprotein, using an engineered JlpA variant. The 3D structure of JlpA revealed a unique topology among known bacterial lipoproteins. Nevertheless, JlpA has a  $\alpha/\beta$ -fold with a large curved  $\beta$ -sheet forming a hydrophobic concave surface, reminiscent of other known bacterial lipoproteins such as LppX, LolA, and LolB proteins. As these proteins serve as lipoprotein carriers or lipid translocators, we speculate a similar role for JlpA as a carrier of as yet unidentified *Campylobacter*-specific lipids. Since we were unable to define the binding site of JlpA by single residue mutation, mutation of multiple residues can be accessed in an effort to find out its binding site. To search for its binding partner, *C. jejuni* whole cell extract or host whole cell extract (such as INT407 cells) can be examined with JlpA. Our structural work provided a framework for determining the molecular function of JlpA and new strategies for the rational design of small molecule inhibitors efficiently targeting JlpA.

## **Chapter V. Searching for Novel Lipoproteins of *Campylobacter jejuni***

Contributed to:

Paek, S., Kawai, F., Choi, K.J., and Yeo, H.J. (2012) Crystal structure of the *Campylobacter jejuni* Cj0090 protein reveals a novel variant of the immunoglobulin fold among bacterial lipoproteins. *Proteins: Struct., Funct., Bioinfo.* **80**, 2804-2809.

## 5.1. Introduction

Despite being a major diarrheal pathogen and the availability of multiple genome sequences, the pathogenesis of *C. jejuni* in humans is not well understood. This is in part because *C. jejuni* largely lacks homologs of virulence factors found in other pathogens (Poly and Guerry, 2008). Along with the significant emergence of antibiotic-resistant strains, the incidence of human *C. jejuni* infection is noticeably increasing (Moore *et al.*, 2006). Thus, there is a growing interest in investigating novel virulence factors to understand better the pathogenic process of *C. jejuni* as a key step toward controlling the disease.

Bacterial lipoproteins are important components of bacterial membranes and are anchored to the membrane through lipids that covalently modify the N-terminal cysteine residue of the mature protein. Lipoproteins are known to play various important roles in bacterial survival and pathogenesis including antigenicity, colonization, transport, cell growth, and signal transduction (Marra *et al.*, 2002; Steyn *et al.*, 2003; Adu-Bobie *et al.*, 2004; Xu *et al.*, 2008; Cron *et al.*, 2009; and Yang *et al.*, 2009). Bacterial lipoproteins are characterized by the presence of a signal peptide sequence at the N-terminal end and a lipobox motif within the signal peptide sequence. The lipobox sequence motif determining lipidation is defined as [LIVMF<sub>W</sub>STAG]<sub>-3</sub>-[LIVMFYSTAGCQ]<sub>-2</sub>-[AGS]<sub>-1</sub>-C<sub>+1</sub>, in which the cysteine residue is invariant (Taylor *et al.*, 2006). Many putative lipoproteins were predicted from genome sequences of various bacteria. For example, Gram-positive *Bacillus subtilis* appears to possess more than one hundred putative lipoproteins and Gram-negative *E. coli* contains about one hundred (Tjalsma *et al.*, 1999 and Brokx *et al.*, 2004). However, most of these lipoproteins remain elusive.

Bioinformatics analyses predicted over 20 putative lipoproteins in the *C. jejuni* genome sequence, based on the presence of conserved signal peptide and lipobox (Babu *et al.*, 2006). Interestingly, however, most of these *C. jejuni* lipoproteins have no apparent sequence similarity (except for signal peptides) to other known bacterial proteins or lipoproteins in the sequence databanks and their functions remain unknown. As in other bacterial pathogens, we hypothesized that many of the putative lipoproteins contribute to the pathogenesis of *C. jejuni*.

To date, only one surface-exposed lipoprotein, JlpA, has a known crystal structure in *C. jejuni*, as described in Chapter IV (Kawai *et al.*, 2012). Despite unexpected from the JlpA sequence, the structure of JlpA is reminiscent of that of other bacterial lipoproteins. The structure revealed an unusually large hydrophobic basin with a localized acidic pocket, suggesting a possibility that JlpA may accommodate multiple ligands, and thus provide a framework for determining the molecular function of JlpA. By analogy, we aimed to reveal the architectures of other predicted lipoproteins of *C. jejuni* by X-ray crystallography and to provide a framework for investigating their functions. Therefore, we initiated structure-function studies on novel lipoproteins of *C. jejuni* to discover their functions, their contribution to pathogenesis, and their potential as drug and vaccine targets. For this, we have focused on four novel lipoproteins: Cj0090, Cj1026c, Cj1090c, and Cj1649.

The Cj0090 protein is encoded within a lipoprotein gene cluster composed of *cj0089*, *cj0090*, and *cj0091*. A recent study reported the function and regulation of the *cj0089-cj0090-cj0091* operon, demonstrating that this lipoprotein operon is activated and directly regulated by CmeR, a pleiotropic transcription regulator modulating the

expression of multiple genes including the multidrug efflux pump CmeABC (Oakland *et al.*, 2011). However, the function of Cj0090 remains unknown. The *cj1026c* gene, together with the immediate downstream gene *cj1025c*, is required for motility of *C. jejuni* but not for flagellar biosynthesis (Sommerlad and Hendrixson, 2007). The majority of the Cj1026c protein appears to be associated with the outer membrane. Moreover, the membrane localization of Cj1026c is dependent on the Cj1025c protein, suggesting that Cj1025c may be necessary for localization and subsequent stabilization of Cj1026c to contribute to the larger macromolecular flagellar complex to promote motility (Sommerlad and Hendrixson, 2007). Currently, structural and functional studies on Cj1090c and Cj1649 are totally absent.

Since our lipoprotein candidates were putative open reading frames (ORFs), the first goal was to test expression and purification of these proteins using a semi-high throughput approach. Using the expression system pET45/BL21(DE3), we systematically cloned, expressed, and purified target proteins. Positive results allowed us to pursue the subsequent objective, crystallization. In addition, since bacterial lipoproteins are thought to be involved in protein-protein interactions, we set out to screen ligands for our target lipoproteins employing a phage display technique. In this chapter, we present an overall status of this initiative and progress on selected target proteins (Cj0090, Cj1026c, Cj1090c, and Cj1649). Hopefully, our results will lead to a framework for determining the functions of these novel lipoproteins in the near future.

## 5.2. Results

### 5.2.1. Identification of putative lipoproteins from *C. jejuni* genome

To investigate novel lipoproteins from *C. jejuni*, 20 putative lipoproteins containing a signal peptide sequence and a lipobox motif were initially selected from the *C. jejuni* NCTC 11168 sequence (Table 5-1). The signal peptide sequence was evaluated by four different predictors (Prediction program: A, PROSITE; B, LipPred; C, LipoP; and D, PRED-LIPO). These proteins were positive in at least two predictors. The lipoboxes of these 20 lipoprotein candidates of *C. jejuni* strictly adhere to the consensus sequence ([LIVMFIRSTAG]<sub>-3</sub>-[LIVMFYSTAGCQ]<sub>-2</sub>-[AGS]<sub>-1</sub>-C<sub>+1</sub>) with the presence of a cysteine residue within the first 25 residues (Table 5-1).

The 20 genes encoding the putative lipoproteins excluding signal peptide sequence were PCR amplified and cloned into the pET45b(+) plasmid to overexpress as N-terminal His-tagged soluble proteins in the *E. coli* strain BL21(DE3). Not surprisingly, the expression and solubility level varied among these proteins (Table 5-2). For example, Cj0091 and Cj1026c were totally insoluble. To purify Cj1026c, we adopted a new expression system, where the full length Cj1026c (including signal peptide) was secreted to the cell culture *via* its signal peptide and contained a C-terminal His-tag. This Cj1026c form was soluble and suitable for purification. After rapid tests of expression, solubility, and Ni-binding affinity of these proteins, we further purified Cj0090, Cj0406c, Cj0646, Cj0842, Cj1026, Cj1090c, and Cj1649 for initial crystallization screen. Based on the crystallization promise, we further focused on Cj0090, Cj1026c, Cj1090c, and Cj1649.

No	Locus	N-terminal sequence	Lipobox prediction
1	Cj0089	MKIKVGLIFSGIAC <b>FLTAC</b> VNQ	A,B,C,D
2	Cj0090	MKKNILFILMFL <b>LSAC</b> APS	A,B,C,D
3	Cj0091	MK <b>TKILGTALIGALLFSGCA</b> QT	A,B,C,D
4	Cj0113	MKKILFTSIAALAV <b>VISG</b> CSTK	A,B,C,D
5	Cj0375	MKHCLALCFVFF <b>LCAC</b> SVK	A,B,C,—
6	Cj0396c	MKKIFLTLFCLIF <b>LCAC</b> GTK	A,B,C,—
7	Cj0406c	MKKYILLASSML <b>ILAAC</b> GGT	A,B,C,D
8	Cj0457c	MGKSFKIHCLTYIIFILQWLIF <b>LSCK</b> KH	A,B,—,—
9	Cj0497	MYRYLLFVLA <b>FFLAAC</b> GSS	A,B,C,D
10	Cj0646	MQIKTITLKLSAVSLGAL <b>FFSGCL</b> GT	A,B,C,—
11	Cj0842	MKIYSVILGA <b>FLLGAC</b> SLK	A,B,C,—
12	Cj0946	MKSCLYFTFIVL <b>FLTAC</b> STK	A,B,C,D
13	Cj0950c	MKKMLQIALAAT <b>FFAGC</b> AST	A,B,C,D
14	Cj1026c	MKKIYFMLAIAGIF <b>AGC</b> VPS	A,B,C,—
15	Cj1074c	MKKGIFLLVFLSV <b>FFSAC</b> STK	A,B,C,D
16	Cj1090c	MKKILIFCIGL <b>FLGAC</b> GYI	A,B,C,D
17	Cj1381	MSKFALILSFFIAL <b>FMSAC</b> TNA	A,B,C,D
18	Cj1483c	MLETKKSFWPYGILLSLLA <b>IIAC</b> IVT	A,B,—,—
19	Cj1649	MLKQLFCILTFIF <b>MLCGC</b> SLR	A,B,C,D
20	Cj1653c	MRHIFFIIT <b>IFFISG</b> CSFY	A,B,C,—

**Table 5-1. Signal peptide sequence and lipobox motif of 20 putative lipoproteins from *C. jejuni*.** The N-terminal sequences of putative lipoproteins are shown. The predicted signal peptide sequence is marked with gray bar and lipobox motif is shown in **bold**. The signal peptide sequence was evaluated by four predictors (Prediction program: A, PROSITE; B, LipPred; C, LipoP; and D, PRED-LIPO). Each of the proteins was positive in at least two of the four predictors (A, B, C, or D).

No	Locus	FL (ORF)	L	MW	E	S	B	P
1	Cj0089	453	434	49.4	+	O	++	-
2	<b>Cj0090</b>	122	115	13.3	+++	O	++	O
3	Cj0091	207	188	20.5	++	X	-	-
4	Cj0113	165	147	16.0	+++	O	++	-
5	Cj0375	158	143	16.7	++	O	++	-
6	Cj0396c	332	316	35.5	+++	O	++	-
7	Cj0406c	299	283	31.4	+	O	++	O
8	Cj0457c	209	185	21.4	+	O	++	-
9	Cj0497	425	409	47.6	+++	O	++	-
10	Cj0646	275	253	27.1	++	O	++	O
11	Cj0842	161	146	17.1	++	O	++	O
12	Cj0946	448	432	50.6	++	O	++	-
13	Cj0950c	144	128	14.4	+	O	+	-
14	<b>Cj1026c</b>	171	155	16.8	++	O	++	O
15	Cj1074c	215	198	23.6	+	O	++	-
16	<b>Cj1090c</b>	170	155	17.4	+++	O	++	O
17	Cj1381	176	158	18.4	+	O	++	-
18	Cj1483c	173	150	17.7	+	O	+	-
19	<b>Cj1649</b>	199	182	21.0	++	O	++	O
20	Cj1653c	152	136	15.3	+	O	++	-

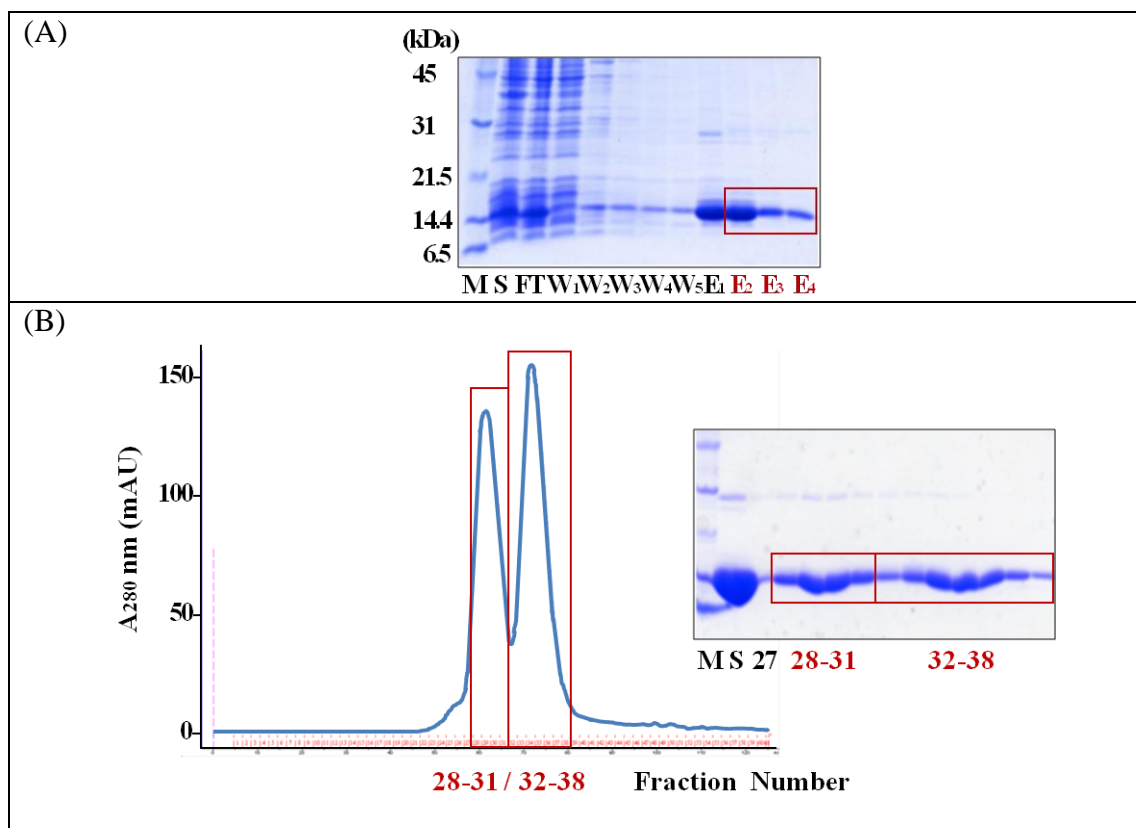
**Table 5-2. List of 20 putative lipoproteins from *C. jejuni* NCTC 11168.** Initially 20 proteins were cloned and expressed as a N-terminal His-tagged proteins in *E. coli* BL21(DE3). Since N-terminal His-tagged Cj1026 was totally insoluble, a new expression system was constructed for the Cj1026c production. In this system, Cj1026c could be purified from the culture supernatant as C-terminal His-tagged protein. The four proteins investigated in this study are in **Bold**. FL, full length of amino acids (aa), ORF, open reading frame; L, length of amino acids without signal peptide; MW, molecular weight excluding signal peptide (kDa); E, expression; S, solubility; B, binding to nickel column; P, purification; +++, Very Good; ++, Good; +, Weak; O, Yes; X, No; and -, no data.



### 5.2.2. Production of soluble proteins: Cj0090, Cj1026c, Cj1090c, and Cj1649

The mature form of N-terminal His-tagged Cj0090 (His-Cj0090<sub>16-122</sub>, ~14 kDa) lacking the signal peptide was purified by a simple two-step procedure. As expected, His-Cj0090<sub>16-122</sub> was eluted as the major protein of ~14 kDa by 125 mM imidazole during nickel affinity chromatography (Figure 5-1 A). The protein sample remained soluble and stable, which facilitated the purification procedure. In the second chromatography step, the elution profile of the gel filtration showed two peaks around 60 mL and 80 mL, which correspond to molecular weights of 28 kDa and 14 kDa, respectively (Figure 5-1 B). Consistently, the SDS-PAGE gel showed Cj0090 as two pools, and thus Cj0090 appeared to be in equilibrium between a monomer form and a dimer form in solution. Interestingly, among different purification trials, the heights of two peaks varied. When the gel filtration step was carried out immediately after the Ni-column step (in the same day), we observed that the second peak (monomer population) was much more prominent. On the other hand, when gel filtration was performed in the next day, the first peak (dimer population) was as large as the first peak.

Initially, the mature form of His-tagged Cj1026c (His-Cj1026c<sub>17-171</sub>) without the signal peptide was overexpressed. However, the protein was found in inclusion bodies in *E. coli*, hampering subsequent purification. Therefore, we adopted an alternative strategy: to express it moderately as a secreted protein *via* its own signal peptide. The full length Cj1026c<sub>1-171</sub> was produced as a C-terminal His-tagged soluble protein in the culture supernatant using the pET20b(+)/BL21(DE3) system (His-Cj1026c). The purification procedure of His-Cj1026c included three steps: ammonium sulfate precipitation, Ni-NTA affinity chromatography, and gel filtration, as detailed in Chapter II. The ammonium

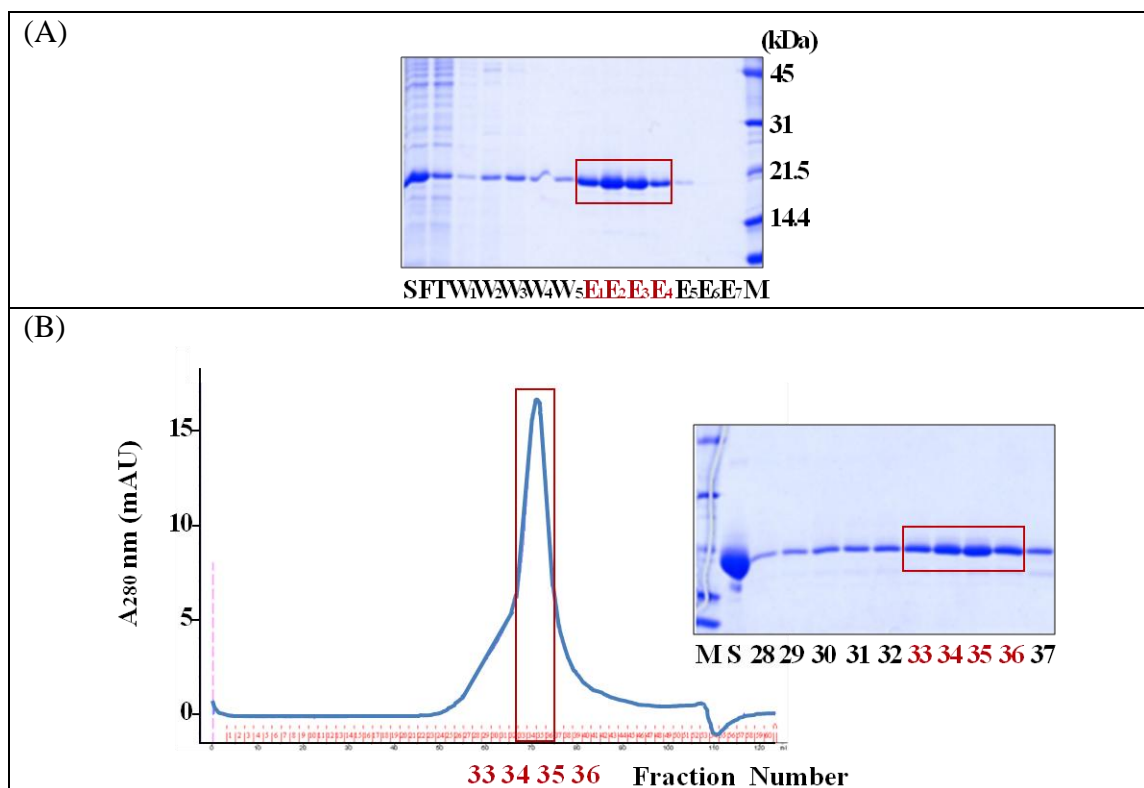


**Figure 5-1. Purification of His-Cj0090<sub>16-122</sub>.** (A) SDS-PAGE analysis of nickel affinity chromatography (15% gel). The elution fractions E<sub>2</sub>-E<sub>4</sub> contained His-Cj0090<sub>16-122</sub> as a major band (~14 kDa) and were combined for gel filtration chromatography. (B) Gel filtration chromatogram and its SDS-PAGE analysis (15% gel). Two peaks corresponding to the molecular mass of 28 and 14 kDa appeared on the chromatogram, indicating that Cj0090 exists as both monomer and dimer structures in solution. M, Marker; S, Sample; FT, Flow-through; W, Wash; and E, Elution.

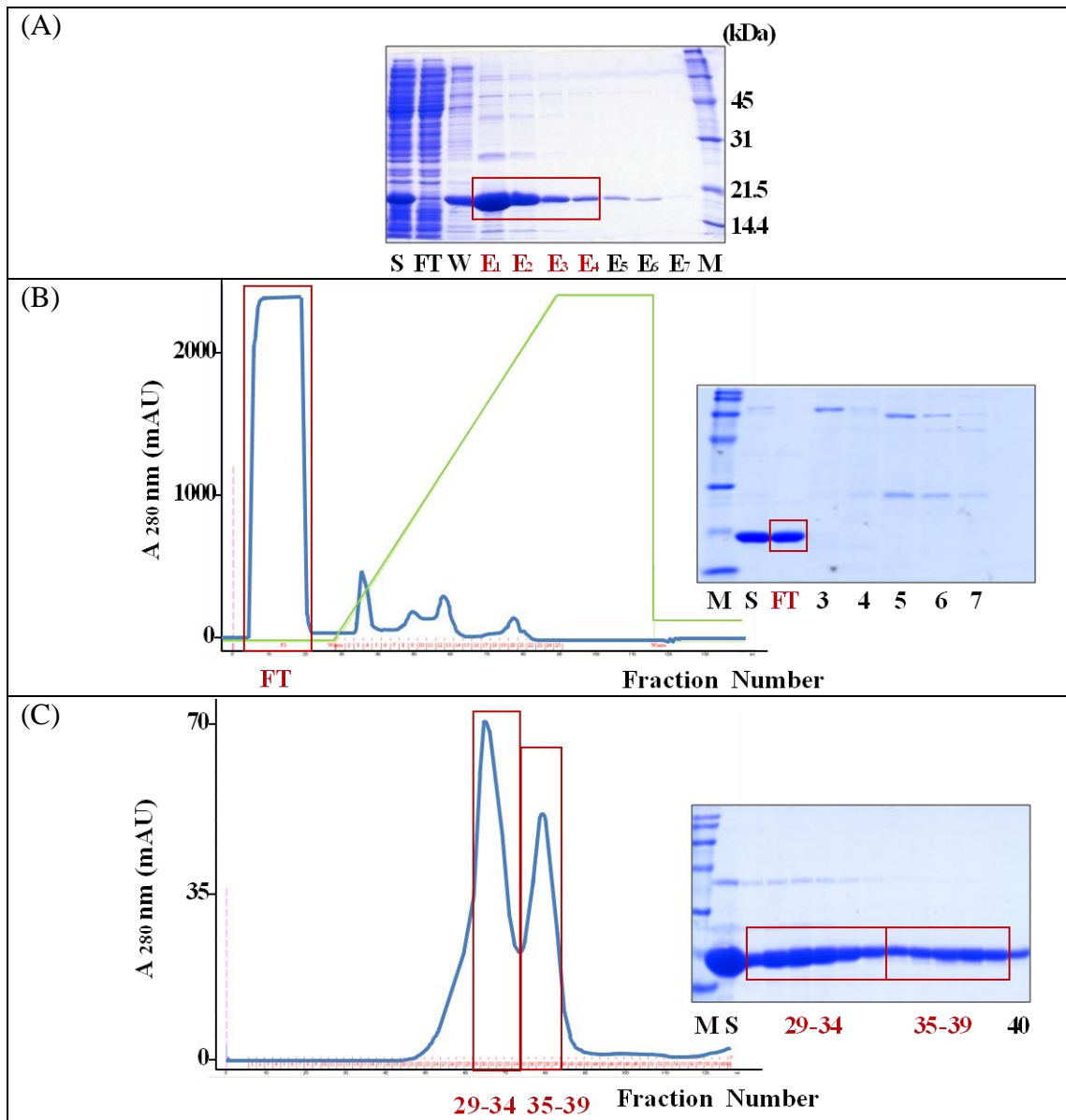
sulfate precipitation procedure was very useful for concentrating proteins from the large volume of cell culture (typically 2 L). The affinity column step further enriched His-Cj1026c, as the SDS-PAGE gel showed a single band at ~17 kDa (Figure 5-2 A). The protein size corresponded to the mature form of Cj1026c, indicating that the signal peptide was removed in the course of secretion in *E. coli*. We also confirmed the new N-terminus by N-terminal sequencing. At the final step of gel filtration, the protein was eluted as a single peak (~17 kDa), suggesting that Cj1026c adopts a monomer structure in solution (Figure 5-2 B).

The mature form of N-terminal His-tagged Cj1090c<sub>16-170</sub> (His-Cj1090c<sub>16-170</sub>, ~18 kDa) was purified by a three-step procedure. His-Cj1090c<sub>16-170</sub> was eluted as the major protein of ~18 kDa at 125 mM imidazole during nickel affinity chromatography (Figure 5-3 A). Purity of the elution fractions were moderate, compared to His-Cj0090<sub>16-122</sub> above (Figure 5-1 A). Thus, a second step included an anion exchanger. While His-Cj1090c<sub>16-170</sub> did not bind to the column, other impurities bound well to the column, resulting in highly pure His-Cj1090c<sub>16-170</sub> in the flow-thorough fraction (Figure 5-3 B). At the final chromatography step, the elution profile of gel filtration showed two peaks corresponding to molecular weight of ~36 kDa and ~18 kDa (Figure 5-3 C). Thus, Cj1090c appears to form both monomeric and dimeric forms in solution like Cj0090 (Figure 5-1 A).

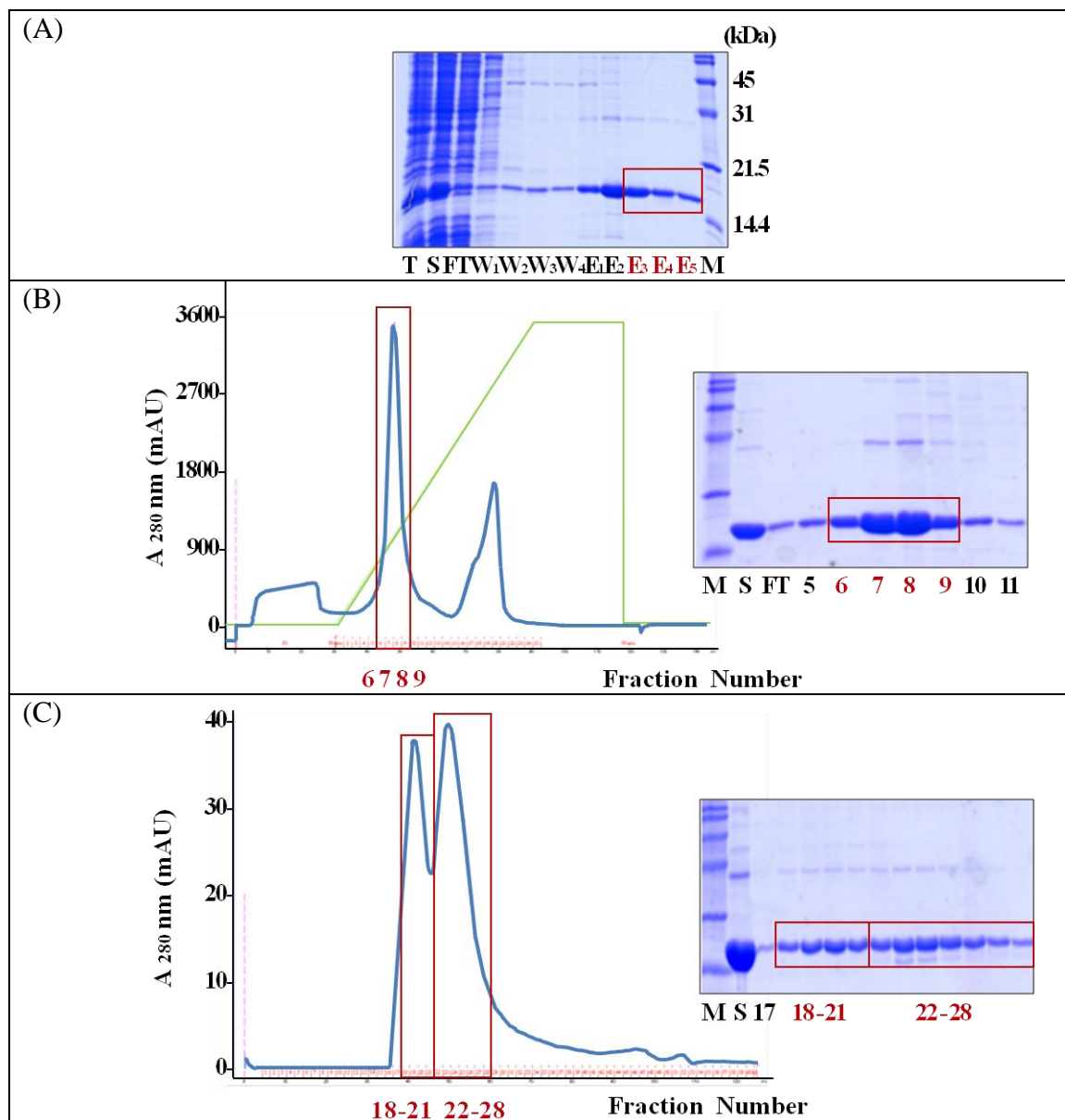
Finally, the mature form of N-terminal His-tagged Cj1649 (His-Cj1649<sub>18-199</sub>, ~23 kDa) was also purified by a three two-step procedure. A moderate expression level was observed. His-Cj1649<sub>18-199</sub> was eluted as the major protein of ~23 kDa at 125 mM imidazole during nickel affinity chromatography (Figure 5-4 A). At the second chromatography step, the bound His-Cj1649<sub>18-199</sub> was eluted from the HiTrapQ column



**Figure 5-2. Purification of His-Cj1026c.** (A) SDS-PAGE analysis of nickel affinity chromatography (15% gel). The elution fractions E1-E4 showed highly enriched His-Cj1026c (~17 kDa) and were combined for the final purification step. (B) Gel filtration chromatogram and its SDS-PAGE analysis (15% gel). A single peak was observed at the elution volume of around 70 mL (an approximate molecular mass of 17 kDa), indicating that Cj1026c adopts a monomer structure. S, Sample; FT, Flow-through; M, Marker; W, Wash; and E, Elution.



**Figure 5-3. Purification of His-Cj1090c<sub>16-170</sub>.** (A) SDS-PAGE analysis of nickel affinity chromatography (15% gel). The elution fractions E<sub>1</sub>-E<sub>4</sub> were combined for the next purification step. (B) Anion exchange chromatogram and SDS-PAGE analysis. While His-Cj1090c<sub>16-170</sub> did not bind to the column, other impurities bound to the column and were separated from the target protein, resulting in highly pure His-Cj1090c<sub>16-170</sub> in FT. (C) Gel filtration chromatogram and SDS-PAGE analysis. Two peaks corresponding to 36 kDa and 18 kDa were observed on the chromatogram. Cj1090c exists as monomers and dimers in solution. S, Sample; FT, Flow-through; M, Marker; Wash; and E, Elution.



**Figure 5-4. Purification of His-Cj1649<sub>18-199</sub>.** (A) SDS-PAGE analysis of nickel affinity chromatography (15% gel). Only fractions E3-E5 with highly pure His-Cj1649<sub>18-199</sub> (~ 23 kDa) were combined for the next step. (B) Anion exchange chromatogram and SDS-PAGE analysis. A peak of His-Cj1649<sub>18-199</sub> was observed between 150 mM and 300 mM of NaCl concentration. (C) Gel filtration chromatogram and SDS-PAGE analysis. Two major peaks corresponding to 46 kDa and 23 kDa were observed on the chromatogram. Cj1649 exists as monomers and dimers in solution. T, Total lysate; S, Sample; FT, Flow-through; M, Marker; Wash; and E, Elution.

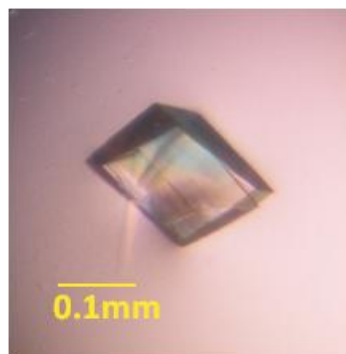
around 150 mM to 300 mM of NaCl (Figure 5-4 B). Finally, the elution profile of gel filtration also showed two relatively broad peaks corresponding to the molecular weight of ~46 kDa and ~23 kDa (Figure 5-4 C). Thus, Cj1649 also appears to form equilibrium between monomeric and dimeric forms in solution like Cj0090 and Cj1090c.

### **5.2.3. Crystallization of Cj0090, Cj1090c, and Cj1649**

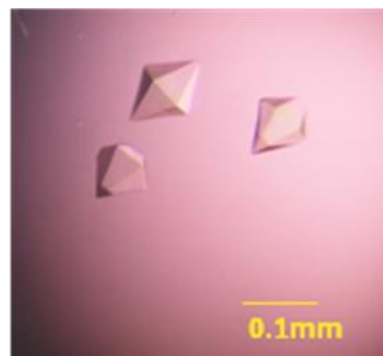
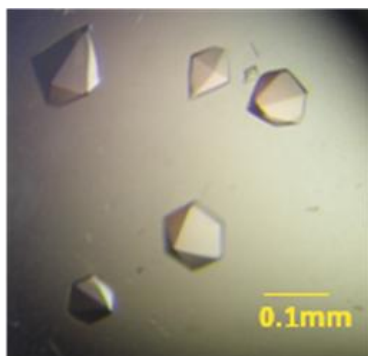
For crystallization of Cj0090, the highly pure concentrated His-Cj0090<sub>16-122</sub> sample (~12 mg/mL) was used for initial crystal screenings. Various promising conditions were observed from the initial screening. Among those, four conditions were optimized to grow sizable crystals. Rectangular-shaped crystals were grown at 17°C by vapor diffusion in hanging drops against a reservoir solution containing 0.8-1.0 M NaCl, HEPES pH 7.0 for about 2 months (Figure 5-5 A). Crystals in straw-like rectangular shape also grew with three different conditions: (i) 18-24% PEG3350/0.2-0.3 M magnesium formate, (ii) 17-25% PEG3350/0.1-0.3 M potassium nitrate, and (iii) 19-24% 2K MME/0.2-0.45 M potassium bromide/0.1 M MES pH 6.3-6.5 within one week. Since His-Cj0090<sub>16-122</sub> tends to form both monomers and dimers in solution (Figure 5-1 B), we thought different oligomeric states might affect crystallization of the protein. However, both forms yielded crystals with the same conditions.

For crystallization of Cj1090c, the concentrated His-Cj1090c<sub>16-170</sub> sample (~10 mg/mL) was submitted for initial crystal screenings. One crystallization condition with PEK 8K precipitant was identified. From both native and selenomethionine-derivatized protein samples, bipyramidal shaped crystals were grown in drops containing 16-18% PEG 8K, 0.3-0.35 M calcium acetate, 10-20% glycerol, and 0.08 M sodium cacodylate or 0.1 M MES (pH6.5) (Figure 5-5 B).

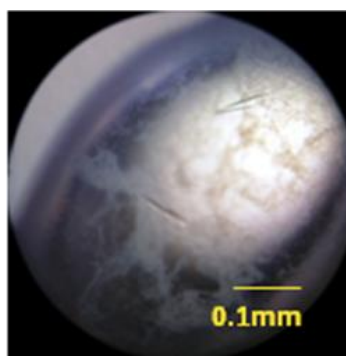
(A)



(B)



(C)



**Figure 5-5. Crystal pictures of Cj0090, Cj1090c, and Cj1649.** (A) A crystal of His-Cj0090<sub>16-122</sub>. Crystals were grown at 17°C by vapor diffusion in drops containing 0.8-1.0 M NaCl and 0.1 M HEPES (pH 7.0). (B) Crystals of His-Cj1090c<sub>16-170</sub>. Crystals of native His-Cj1090c<sub>16-170</sub> (left) were grown with 16-18% PEG 8K, 0.3-0.35 M calcium acetate, 10-20% glycerol, and 0.08 M sodium caccodylate (pH 6.5). Crystals of selenomethionine-derivatized His-Cj1090c<sub>16-170</sub> (right) were grown with 13-16% PEG 8K, 0.3-0.35 M calcium acetate, 10-20% glycerol, and 0.1 M MES (pH 6.5). (C) Crystals of His-Cj1649<sub>18-199</sub>. Crystals were grown with 5-11% PEG 6K, 0.1 M Ammonium acetate, 0-5% glycerol, and 0.1 M Bis-tris (pH 5.5).



Similarly, the Cj1649 protein sample (~10 mg/mL) was screened. Rod-shaped crystals were grown with 5-11% PEG 6K, 0.1 M Ammonium acetate, 0-5% glycerol, and 0.1 M Bis-tris (pH5.5) in about 3 days (Figure 5-5 C). In fact, the similar crystals were observed with a wide range of PEG precipitants (PEG 3350, 4K, 6K, 8K, and 10K).

#### **5.2.4. Search for potential ligands for lipoproteins**

##### **5.2.4.1. Screening of small peptides by phage display experiment**

To discover peptides that specifically interact with Cj0090, Cj1026c, Cj1090c, and Cj1649, a M13 phage display library was screened. After the fifth round of *in vitro* panning, 20-40 plaques were randomly selected and submitted for sequence analysis to identify specific peptides. In the screen with Cj0090, 40 plaques were selected for sequence analysis and sequence results were obtained for 37 clones. 9 out of 37 sequenced clones showed the same peptide sequence WHGTPWEVTVIG (Cj90P-1). This peptide represented the highest frequency with 24.3% [9/37] (Table 5-3). In the case of Cj1026c, 20 plaques were picked out for sequence analysis and all showed sequence results. 19 out of 20 showed the same peptide sequence KTINPWDGMLYW (Cj1026cP-1) [19/20 (95.0%)] (Table 5-3). To confirm whether Cj1026cP-1 is the true specific peptide to Cj1026c, the screen was repeated. The sequence analysis of the second experiment revealed that all of 6 candidates also showed the same peptide sequence of KTINPWDGMLYW as the sequence identified in the first screening [6/6 (100.0%)] (Table 5-3). Because the same peptide was identified from both independent screenings with high frequency, this peptide (KTINPWDGMLYW) is considered as a true specific peptide sequence for Cj1026c (95.0% from the 1<sup>st</sup> screening and 100% from the 2<sup>nd</sup> screening) (Table 5-3). A specific peptide WGITVETAYGTA (Cj1090cP-1) was

**Table 5-3. Peptide identified by phage display.**

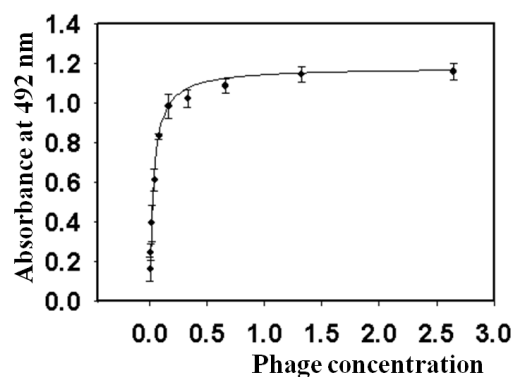
<b>Target protein</b>	<b>Peptide sequence</b>	<b>Frequency</b>
Cj0090	WHGTPWEVTVIG (Cj90P-1)	<b>9/37 (24.3%)</b>
Cj1026c	KTINPWDGMLYW (Cj1026cP-1)	<b>19/20 (95.0%)</b>
		<b>1<sup>st</sup> experiment</b>
		<b>6/6 (100.0%)</b>
		<b>2<sup>nd</sup> experiment</b>
Cj1090c	WGITVETAYGTA (Cj1090cP-1)	<b>6/18 (33.3%)</b>

identified for Cj1090c. 6 out of 18 clones showed the same peptide sequence of WGITVETAYGTA [6/18 (33.3%)] (Table 5-3). Although the fifth round of *in vitro* panning was carried out for Cj1649, we could not detect any specific peptide. The database searches of these peptides did not reveal any significant known protein or peptide sequence, suggesting that these are novel peptides. It is speculated that the artificial peptide may be a structural mimetic of the natural ligand and not bear high enough sequence similarity to give a hit in the blast search. Regardless, these novel peptides with high specificity can be useful for developing inhibitors of the target protein for example.

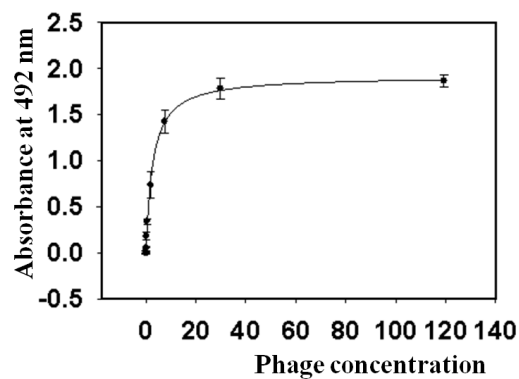
#### **5.2.4.2. Specificity of the potential ligands**

To confirm the binding specificity of the highest frequency peptides to the target proteins, the binding affinity was examined by phage ELISA. Four-fold serial dilutions of phages were incubated with a target protein coated in the microplate. As a negative control, BSA was coated to assess nonspecific binding of phages. After washing the microplate, bound phages were detected with anti-M13 antibody. The specific curve for the peptide phage was obtained by subtracting the nonspecific curve from total curve (Figure 5-6). The calculated apparent dissociation constants ( $K_d^{app}$ ) were  $0.040 \pm 0.002$  pM,  $2.64 \pm 0.33$  pM, and  $0.10 \pm 0.01$  pM, for Cj90P-1, Cj1026cP-1 and Cj1090cP-1, respectively.

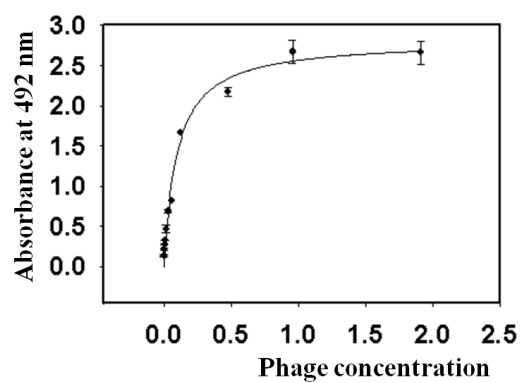
(A) Binding curve of Cj90P-1 to Cj0090



(B) Binding curve of Cj1026cP-1 to Cj1026c



(C) Binding curve of Cj1090cP-1 to Cj1090c



**Figure 5-6. Binding phage peptides.** Binding affinity was determined using the phage ELISA method. Each experiment was done in triplicate.

### 5.3. Discussion

Bacterial lipoproteins carry out various essential roles in host-pathogen interactions. Most *C. jejuni* lipoproteins are putative proteins and have no sequence homology with other known proteins. As the first step to discover structural and biochemical properties of these lipoproteins, 20 putative lipoproteins were initially selected from the genome of *Campylobacter jejuni* NCTC 11168. In this study, we made progress on four proteins Cj0090, Cj1026c, Cj1090c and Cj1649. The crystal growth of Cj0090 allowed phasing using the SAD method on Br that was incorporated to the crystal. The subsequent structural work was done in collaboration with Dr. Kawai in our laboratory (Paek *et al.*, 2012). We also obtained diffracting crystals of Cj1090c and the completion of structural refinement is in progress. The Cj1649 crystals diffracted to a low resolution, hampering phasing. With the highly pure protein samples in our hand, we attempted to discover possible ligand peptides using phage display. For Cj0090, Cj1026c, and Cj1090c, we found at least one specifically binding peptide.

#### 5.3.1. Crystal structure of Cj0090, a novel immunoglobulin fold

Cj0090 protein revealed a single domain structure composed of seven  $\beta$ -strands forming two  $\beta$  sheets ( $\beta$  sandwich), one containing 4 strands (EDGA as sheet 1) and the other containing 3 strands (FCB as sheet 2), with a Greek key motif (Figure 5-7 A and B). The Cj0090 structure represents an immunoglobulin (Ig)-like  $\beta$  sandwich fold, but differs from the most common Ig-like superfamily folds referred to as C-, V-, H-, and I-type folds (Bork *et al.*, 1994 and Halaby *et al.*, 1999). Current classification systems of Ig-like folds are mainly based on the number of  $\beta$  strands distributed in two sheets with characteristic topology and connectivity (Figure 5-8). According to the classification of

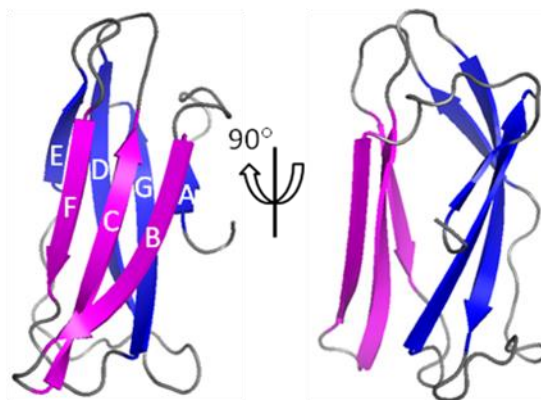
(A)

A 1	-MKKNI--LFILMFLLSACAPSYQINSNQ---NTVILGSNLPKSLVKQFQKRINSNGYLEFEVILRSTF	
B 1	-MNKGL-VLACLLGLSACAPHTGGIMIS-STGEVRVDNGSPHSDVDVSAVTTQAEAGFLRARGTIISK	
C 1	-MRTWL-LAVLTALLLVGCSANTAGLRVDGASQQLFNDLSKSLSDIEDISTTEVDGHTRGVRLQSNQ	
D 1	MIKGC I-AVSLG L L L L T G C R S H P E I P V S D --A Q S L V M E S T V L A A G V T A E P P E L T A S D I Q P S A S S R V Y N E R	
E 1	MRKGC F G L V S L A L L L L V G C R S H P E I P V N D --E Q S L V M E S S L L A A G I S A E K P V L S T S D I Q P S A S S T L Y N E R	

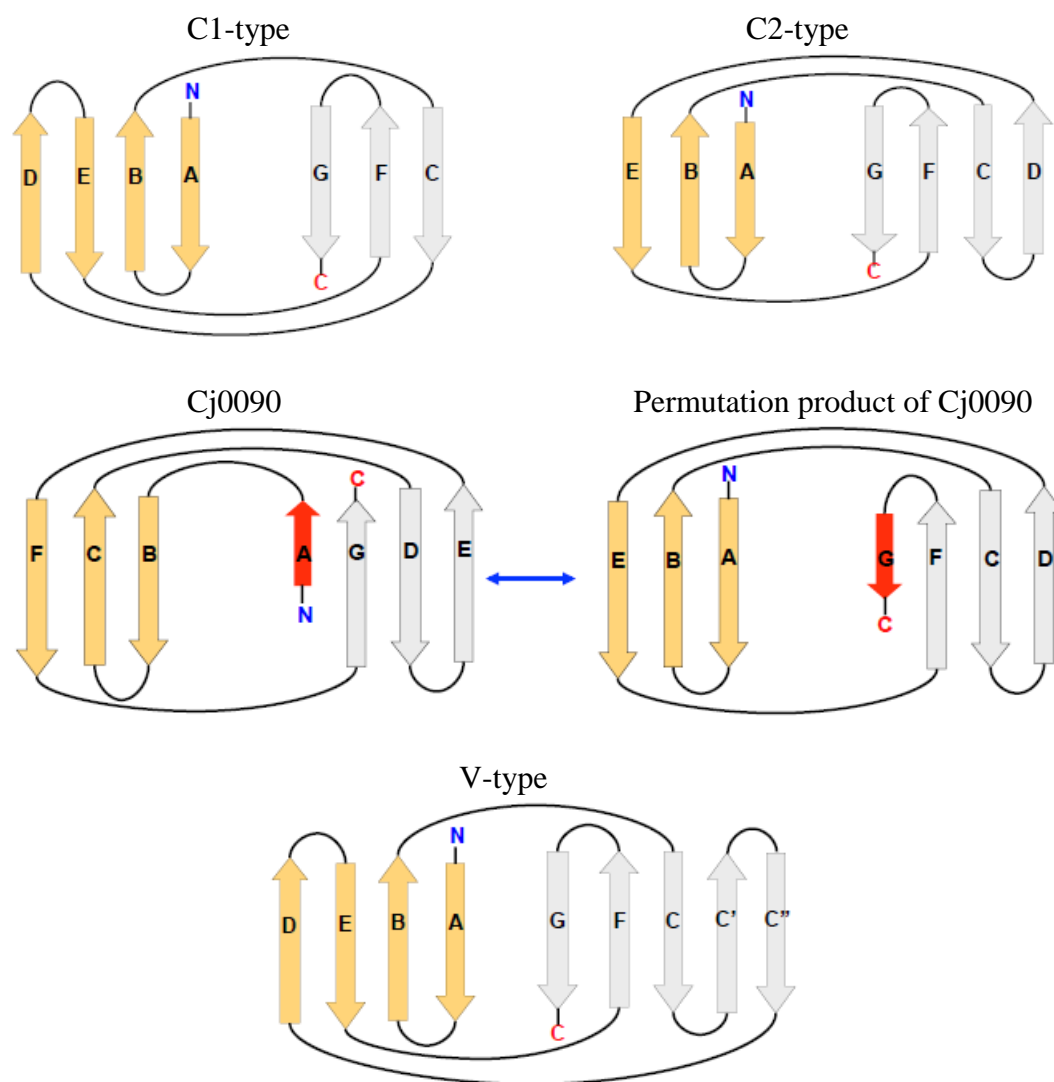
  

64	AKD--VIYKVDWLDKDGFLRDVLNEDYQALRIPAGQEVILRKLASDTRANDFRLEIKAKN--	122
68	PKDQRLQYKFTWYDINGATVEDEGVSW-KSLKLHGKQQMQVTALSPNATAVRCELYVREAI SN	129
69	KSDVHVQYRFPYWDNDGLEVNTKLSPW-KTIILRGMETVSLTEVSVNPNKGQFRVQIRES DQ-	129
67	QEPITVHYRFPYWDARGLEMHPLEAP--RSVTIPARSSVTLFGSANYLGAHKVRLYL L Y L ----	124
69	QEPVTVHYRFPYWDARGLEMHPLEP--RSVTIPAHSAVTLYGSANFLGAHKVRLYL L Y L ----	125

(B)



**Figure 5-7. Structure of Cj0090.** (A) Sequence alignment of Cj0090 and representative putative lipoproteins: A, Cj0090 from *Campylobacter jejuni*; B, Shew\_1734 from *Shewanella loihica*; C, VC1895 from *Vibrio cholerae*; D, STM1206 from *Salmonella enterica*; and E, Z1743 from *Escherichia coli*. The Cj0090 secondary structures are indicated above the primary sequences. Identical residues, strongly conserved residues, and weakly conserved residues are highlighted in blue hues. (B) Cartoon representation of the Cj0090 structure containing two  $\beta$ -sheets (blue, AGDE and magenta, BCF) (PDB code: 4GIO). The  $\beta$  strands are labeled as A to G (adapted from Paek *et al.*, 2012).



**Figure 5-8. Topology diagrams of the Ig fold and comparison with the Cj0090 fold.**

The  $\beta$ -sandwich structure is formed by 7 and 9 strands in C-type and V-type, respectively. In contrast to the C1-type, the C2-type fold has three strands for sheet 1 and four strands for sheet 2. A particular case of C2-type is referred to as the fibronectin type III (FN3) fold. The fold of Cj0090 can be seen as a circular permutation product of the C2-type fold. Following ligation of the N and C termini, the new termini are formed by cleavage between strands A and B. The resulting product becomes C2-type (or FN3) (adapted from Paek *et al.*, 2012).

Ig-like folds, Cj0090 is a variant of type-C that is distinct from the type-V in the number of strands (typically, 7 strands for type-C and 9 strands for type-V). In contrast to the C1-type (that is the classical C-type), the C2-type (also referred to as S-type) fold has three strands for sheet 1, and four strands for sheet 2. A particular case of C2-type is referred to as the fibronectin type III (FN3) fold that is found in many animal proteins and is involved in ligand binding, since the FN3 structure does not contain the conventional disulfide bond that is commonly presented in other C2-type Ig related proteins (Bork *et al.*, 1994 and Halaby *et al.*, 1999). Thus, the fold of Cj0090 can be considered as a circular permutation product of the C2-type fold. Indeed, if strand A were to be at the C-terminus by permutation, strands of Cj0090 would occur as EBA (from FCB) and DCFG (from EDGA), which is a typical C2-type topology (or FN3 fold) (Figure 5-8). Taken together, the fold of Cj0090 is unique and a novel variant of Ig-like superfamily.

At the primary structure level, Cj0090 is highly conserved within *Campylobacter* species and has homologs with low sequence identities from various organisms. The function of these proteins is unknown. Structure homology searches using the DALI server revealed a large number of solved structures belonging to the Ig superfamily (Holm and Rosenström, 2010). However, none of these proteins has sequence identity over 17% with Cj0090. Among these, two bacterial proteins were noteworthy: (1) an uncharacterized protein referred to as Shew\_1734 and (2) ApaG. Both proteins are similar to Cj0090 in size. The crystal structure of Shew\_1734 from *Shewanella loihica* represents the closest structural homolog of Cj0090. Shew\_1734 is also a putative bacterial lipoprotein and contains the signal peptide sequence (Figure 5-7 A). Despite the low sequence identity, Cj0090 and Shew\_1734 clearly share the same topology except



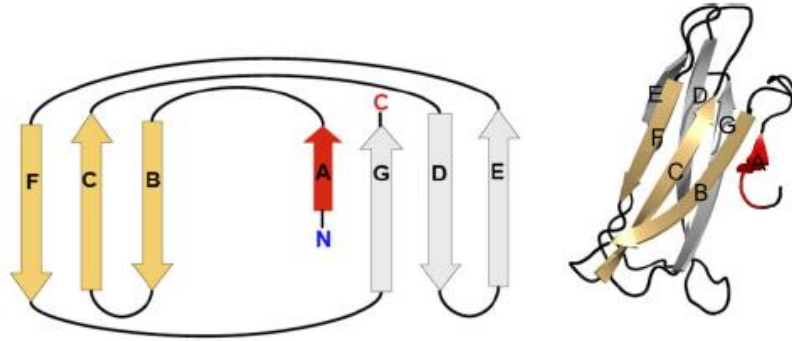
for the N-terminus (strands A1 and A2) (Figure 5-9 A and B). Interestingly, strands A1 and A2 contribute to a 5-stranded  $\beta$  sheet with strands FCB in Shew\_1734, whereas strand A contributes to a 4-stranded  $\beta$  sheet with strand GDE in Cj0090. Thus, this novel variant of Ig fold observed in Cj0090 may be conserved with a minor difference among bacterial lipoproteins. The next structural homology, despite a low sequence identity of 6%, was found in the ApaG proteins (~14 kDa) that belong to a conserved protein family from Gram-negative bacteria. While the function of the ApaG proteins is unknown, it has been suggested that ApaG may be involved in the metabolism of nucleotide polyphosphates (Cicero *et al.*, 2007). Unlike Cj0090 and Shew\_1734, the ApaG proteins do not contain the signal peptide and thus are not lipoproteins. The structure is characterized by seven antiparallel  $\beta$  strands forming two  $\beta$  sheets, one with three strands (ABE) and the other with four strands (GFCD) (Figure 5-9 C). Thus, ApaG adopts a FN3 fold that is a particular case of C2-type, as described above.

In summary, having an Ig-like fold suggests that Cj0090 is involved in mediating protein-protein interaction, consistent with a possible role for bacterial lipoproteins.

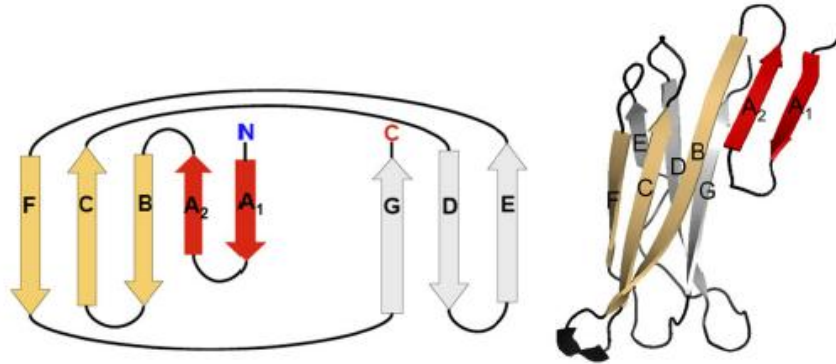
### **5.3.2. Biochemical assays of lipoproteins of *C. jejuni***

In the absence of a functional assay for the lipoproteins, we sought a general technique for the study protein-protein interactions based on the hypothesis that lipoproteins interact with other molecules. Using a phage display approach, we found specific peptides for three target proteins: Cj0090 (Cj90P-1, WHGTPWEVTVIG), Cj1026c (Cj1026cP-1, KTINPWDGMLYW), and Cj1090c (Cj1090P-1, WGITVETAYGTA) (Table 5-3). Sequence searches suggest that these peptides represent novel peptides. It is known that a binding peptide identified in a particular panning experiment does not

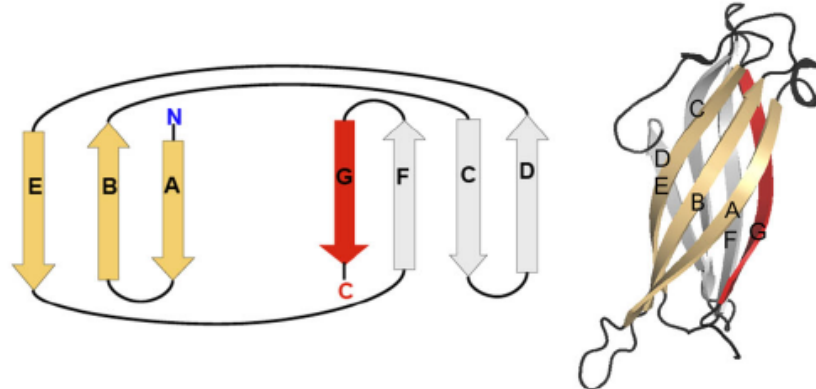
(A)



(B)



(C)



**Figure 5-9. Topology diagrams of Cj0090 and close structural homologs with Ig fold variants.** The corresponding cartoon representation of the structure is shown to the right. The conserved core strands are labeled and colored in gold and silver. The structurally variable strands are highlighted in red (adapted from Paek *et al.*, 2012). (A) Cj0090. (B) Putative lipoprotein Shew\_1734. (C) ApaG protein.

necessarily correspond to a natural ligand for the target, since the libraries consist of fully randomized peptides displayed on phages and an isolated peptide is only a structural mimetic of the ligand and does not bear high enough sequence similarity to give a hit in the BLAST search (Tipps *et al.*, 2010).

The specificity of the peptides was assessed by phage ELISA, resulting in picomolar-binding affinities (see 5.2.4.2.). Although the tight binding suggests specific interactions between target proteins and peptides, these  $K_d^{\text{app}}$  values are not true references for the affinity, since multiple peptides are displayed on each phage and an avidity effect exists. Here, avidity is the combined synergistic strength of multiple bond affinities. In addition, since the peptide Cj1026c-P1 was the single type (100% frequency) identified from two independent experiments (Table 5-3), the Biacore assay based on surface plasmon resonance (SPR) was used to estimate the binding affinity in the Cj1026c and Cj1026cP-1 peptide interaction. The immobilized Cj1026c on the sensor surface was reacted with increasing concentrations of synthetic Cj1026cP-1 peptide. Sensograms of association and dissociation phases for the interaction between Cj1026c and Cj1026cP-1 at each concentration showed a similar dose-dependent pattern (data not shown). Kinetic analysis of these data indicated that a specific binding response was observed. As suspected above, the affinity calculated from this experiment was lower than the  $K_d^{\text{app}}$  value from phage ELISA. Further experiments are necessary to obtain true affinities. As a long-term goal, these peptides may serve as new tools for the rational design of small molecule inhibitors or markers efficiently targeting the specific proteins.

In summary, this study has established molecular tools and methods useful for further characterization of novel lipoproteins in *C. jejuni*.

## **Chapter VI. Concluding Remarks**

*Campylobacter jejuni*, a Gram-negative motile bacterium, is a leading cause of human gastrointestinal infections worldwide. Our understanding of the *C. jejuni* pathogenesis in humans is still rudimentary. This organism appears to dispose many distinct virulence determinants, compared with other significant pathogens. This dissertation concerns structural studies on *C. jejuni* virulence proteins (JlpA and Cj0977) or potential virulence factors (novel lipoproteins). As our ultimate goals of research, knowledge of structural features of our target proteins may be useful for (1) a direct implication in development of therapeutics (vaccines and small molecule inhibitors) and (2) a better understanding of the *C. jejuni* pathogenesis.

The Cj0977 protein is not a component of the flagellum, but is expressed by a  $\sigma^{28}$  promoter that controls late genes in the flagellar regulon. Cj0977 is a virulence factor associated with invasion, as a Cj0977 null mutant was fully motile but significantly reduced in invasion of intestinal epithelial cells *in vitro* (Goon *et al.*, 2006). These observations suggest a regulatory role of Cj0977 linked to flagella and/or virulence. In an effort to understand the mechanism by which Cj0977 contributes to virulence of *C. jejuni* at the molecular level, we set out to determine the structure of Cj0977. **Chapter III** describes my work focusing on obtaining high quality crystals, a bottleneck step in protein crystallography. Although the initial full-length protein yielded crystals, they did not diffract beyond 6 Å. To improve crystal diffraction, we constructed several truncated versions of Cj0977, based on limited proteolysis as well as survey of protein stability. One variant of Cj0977, corresponding to a structural core domain, yielded a crystal that diffracted to 2.6 Å and led up to solving the structure in our laboratory. The Cj0977 structure reveals homodimeric “hot dog” fold architecture. Since the characteristic hot

dog fold is found in various coenzyme A (CoA) compound binding proteins, structural comparison with other known hot dog fold proteins locates a putative binding site for an acyl-CoA compound in Cj0977 protein. Structure-based site-directed mutagenesis followed by invasion assays indicates that key residues in the putative binding site are indeed essential for the Cj0977 virulence function, suggesting a possible function of Cj0977 as an acyl-CoA binding regulatory protein.

Thus, our structural study provided some clues on how Cj0977 plays a regulatory role. In future work, one of the most important aims would be to discover the identity of the CoA derived ligand using *C. jejuni* cell extracts. Once we know the ligand, we can pursue crystallization and structure determination of Cj0977 bound to its CoA derived ligand, which will help design inhibitors of this virulence protein. These studies would represent a significant step toward the understanding of the structural basis for *C. jejuni* virulence.

JlpA is a *Campylobacter* specific cell surface-exposed lipoprotein and an adhesin virulence factor (Jin *et al.*, 2001). To gain insight into the structural basis of *C. jejuni* virulence, we aimed to determine the crystal structure of JlpA. In **Chapter IV**, my work focused on establishing experimental procedures of purification and crystal growth. The mature form of JlpA (Cys-18 to Phe-372) without the signal peptide was overexpressed for crystallization. Reflecting the adhesin function of JlpA (sticky protein interacting with macromolecules), we obtained moderately pure (~85%) protein sample after three chromatography steps. Despite this level of purity, we succeeded in growing crystals. The initial JlpA crystals diffracted to a resolution of 3.3 Å, resulting in a low phasing power. By engineering JlpA variants (introduction of additional Met sites in the JlpA sequence to

improve MAD phasing), we were able to grow improved crystals that diffracted to a resolution of 2.7 Å. As the subsequent structure determination was completed in our laboratory, the JlpA structure revealed a catcher's mitt-shaped unclosed half  $\beta$ -barrel with a unique topology among known bacterial lipoproteins. Nonetheless, the observation that JlpA has an  $\alpha/\beta$ -fold with a large curved  $\beta$ -sheet forming hydrophobic concave surface, characteristics of LolA/LolB proteins and LppX-like proteins, raises a question about a similar role for JlpA as a lipid carrier or translocator. In future, screening of cell extract based ligands would be an interesting approach.

Although the Lol pathway established in *E. coli* is considered as highly conserved and a major lipoprotein sorting system in various Gram-negative bacteria, some bacteria do not encode a LolB homolog including *C. jejuni*. The next question is whether *Campylobacter* employs a different lipoprotein sorting mechanism other than the Lol pathway or conservation of the *C. jejuni* LolB protein is not strong enough to be detected from the genome (meaning that *C. jejuni* employs a structural homolog of LolB for lipoprotein sorting, but we cannot know this from the sequence). If the latter is the case, structure determination will be necessary to identify the LolB homolog in *Campylobacter*. Compared with *E. coli* containing about 100 lipoproteins, the *C. jejuni* genome appears to contain over 20 lipoproteins and the majority of them are uncharacterized. In a future study, the systematic structural analysis approach of lipoproteins may help to identify a structural homolog of LolB in the absence of a known sequence homolog. This led up to one of the motivations for our initiatives on novel lipoproteins of *C. jejuni* (**Chapter V**).

Bacterial lipoproteins carry out various essential roles in host-pathogen interactions. Most *C. jejuni* lipoproteins are putative proteins and have no sequence

homology with other known proteins. As the first step to discover structural and biochemical properties of these lipoproteins, 20 putative lipoproteins were initially selected from the genome of *Campylobacter jejuni* NCTC 11168. In **Chapter V**, we describe progress on four proteins Cj0090, Cj1026c, Cj1090c and Cj1649 in various stages. The crystal growth of Cj0090 allowed phasing using the SAD method on Br that was incorporated to the crystal. The subsequent structural work was done in our laboratory, revealing a new variant of Ig superfamily fold that appears to be conserved among bacterial lipoproteins. Having an Ig-like fold suggests that Cj0090 is involved in mediating protein-protein interactions, consistent with a possible role for bacterial lipoproteins. We also obtained diffracting crystals of Cj1090c and the completion of structural refinement is in progress. The Cj1649 crystals diffracted to a low resolution, hampering phasing. With the highly pure protein samples in our hand, we attempted to discover possible ligand peptides using phage display. For Cj0090, Cj1026c, and Cj1090c, we found at least one specifically binding peptide.

Thus, **Chapter V** presents precious molecular tools and methods useful for further characterization of novel lipoproteins in *C. jejuni*. In the near future, we aim to complete the Cj1090c structure. By analogy to Cj0977, JlpA, Cj0090, the Cj1090c structure will provide some clues on the function of the protein. Regarding Cj1649, we will need to troubleshoot the problem with low-resolution crystals. In addition, other novel lipoproteins (from 20 targets) are ready to be exploited. As *in vivo* approaches, generating and testing null mutants of our selected lipoproteins might lead to understanding better its functions in the organism.



In conclusion, the crystal structures of three *C. jejuni* proteins were completed along with my dissertation. Both Cj0977 and JlpA are known virulence factor proteins and Cj0090 is a potential virulence factor protein. By pairing crystallography with functional assays, we can discover unprecedented characteristics of *C. jejuni* lipoproteins and specifically interacting ligand molecules. The structures will further serve as framework for addressing unresolved fundamental questions such as the role of each lipoprotein in virulence as well as the mechanisms underlying the transport of lipoproteins in *Campylobacter* spp.. The preliminary result of peptide identification is very encouraging, since this approach is transferable to almost any target expressed in an adhering system and is particularly useful for surface-exposed lipoprotein adhesins that currently have no known receptor modulators.

## References

- Adu-Bobie, J., Lupetti, P., Brunelli, B., Granoff, D., Norais, N., Ferrari, G., Grandi, G., Rappuoli, R., and Pizza, M. (2004) GNA33 of *Neisseria meningitidis* is a lipoprotein required for cell separation, membrane architecture, and virulence. *Infect. Immun.* **72**, 1914-1919.
- Alloing, G., de Philip, P., and Claverys, J.P. (1994) Three highly homologous membrane-bound lipoproteins participate in oligopeptide transport by the Ami system of the gram-positive *Streptococcus pneumoniae*. *J. Mol. Biol.* **241**, 44-58.
- Altekruse, S.F., Stern, N.J., Fields, P.I., and Swerdlow, D.L. (1999) *Campylobacter jejuni*-an emerging foodborne pathogen. *Emerg. Infect. Dis.* **5**, 28-35.
- Ang, C.W., De Klerk, M.A., Endtz, H.P., Jacobs, B.C., Laman, J.D., van der Meché, F.G., and van Doorn, P.A. (2001) Guillain-Barré syndrome- and Miller Fisher syndrome-associated *Campylobacter jejuni* lipopolysaccharides induce anti-GM1 and anti-GQ1b Antibodies in rabbits. *Infect. Immun.* **69**, 2462-2469.
- Ashgar, S.S., Oldfield, N.J., Wooldridge, K.G., Jones, M.A., Irving, G.J., Turner, D.P., and Ala'Aldeen, D.A. (2007) CapA, an autotransporter protein of *Campylobacter jejuni*, mediates association with human epithelial cells and colonization of the chicken gut. *J. Bacteriol.* **189**, 1856-1865.
- Aspinall, G.O., McDonald, A.G., and Pang, H. (1992) Structures of the O chains from lipopolysaccharides of *Campylobacter jejuni* serotypes O:23 and O:36. *Carbohydr. Res.* **231**, 13-30.
- Aspinall, G.O., Lynch, C.M., Pang, H., Shaver, R.T., and Moran, A.P. (1995) Chemical structures of the core region of *Campylobacter jejuni* O:3 lipopolysaccharide and an associated polysaccharide. *Eur. J. Biochem.* **231**, 570-578.
- Babu, M.M., Priya, M.L., Selvan, A.T., Madera, M., Gough, J., Aravind, L., and Sankaran, K. (2006) A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J. Bacteriol.* **88**, 2761-2773.
- Bachtiar, B.M., Coloe, P.J., and Fry, B.N. (2007) Knockout mutagenesis of the *kpsE* gene of *Campylobacter jejuni* 81116 and its involvement in bacterium-host interactions. *FEMS Immunol. Med. Microbiol.* **49**, 149-154.
- Bacon, D.J., Szymanski, C.M., Burr, D.H., Silver, R.P., Alm, R.A., and Guerry, P. (2001) A phase-variable capsule is involved in virulence of *Campylobacter jejuni* 81-176. *Mol. Microbiol.* **40**, 769-777.

- Baloda, S.B., Faris, A., Fröman, G., and Waldström, T. (1985) Fibronectin binding to *Salmonella* strains. *FEMS Microbiol. Lett.* **28**, 1-5.
- Baumgartner, M., Karst, U., Gerstel, B., Loessner, M., Wehland, J., and Jansch, L. (2007) Inactivation of Lgt allows systematic characterization of lipoproteins from *Listeria monocytogenes*. *J. Bacteriol.* **189**, 313-324.
- Beer, W., Adam, M., and Seltman, G. (1986) Monosaccharide composition of lipopolysaccharides from *Campylobacter jejuni* and *Campylobacter coli*. *J. Basic Microbiol.* **26**, 201-204.
- Black, R.E., Levine, M.M., Clements, M.L., Hughes, T.P., and Blaser, M.J. (1988) Experimental *Campylobacter jejuni* infection in humans. *J. Infect. Dis.* **157**, 472-479.
- Bork, P., Holm, L., and Sander, C. (1994) The immunoglobulin fold. Structural classification, sequence patterns and common core. *J. Mol. Biol.* **242**, 309-320.
- Brightbill, H.D., Libraty, D.H., Krutzik, S.R., Yang, R.B., Belisle, J.T., Bleharski, J.R., Maitland, M., Norgard, M.V., Plevy, S.E., Smale, S.T., Brennan, P.J., Bloom, B.R., Godowski, P.J., and Modlin, R.L. (1999) Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science* **285**, 732-736.
- Brokx, S.J., Ellison, M., Locke, T., Bottorff, D., Frost, L., and Weiner, J.H. (2004) Genome-wide analysis of lipoprotein expression in *Escherichia coli* MG1655. *J. Bacteriol.* **186**, 3254-3258.
- Buelow, D.R., Christensen, J.E., Neal-Mckinney, J.M., and Konkel, M.E. (2010) *Campylobacter jejuni* survival within human epithelial cells is enhanced by the secreted protein CiaI. *Mol. Microbiol.* **80**, 1296-1312.
- Carrillo, C.D., Taboada, E., Nash, J.H., Lanthier, P., Kelly, J., Lau, P.C., Verhulp, R., Mykytczuk, O., Sy, J., Findlay, W.A., Amoako, K., Gomis, S., Willson, P., Austin, J.W., Potter, A., Babiuk, L., Allan, B., and Szymanski, C.M. (2004) Genome-wide expression analyses of *Campylobacter jejuni* NCTC11168 reveals coordinate regulation of motility and virulence by flhA. *J Biol Chem.* **279**, 20327-20338.
- Chandonia, J.M. and Brenner, S.E. (2006) The impact of structural genomics: expectations and outcomes. *Science* **311**, 347-351.
- Chen, Y.H., Poly, F., Pakulski, Z., Guerry, P., and 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. *Carbohydr. Res.* **343**, 1034-1040.

Christensen, J., Pacheco, S., and Konkel M. (2009). Identification of a *Campylobacter jejuni* virulence protein that is secreted from the flagellum and is required for maximal invasion of host cells. *Mol. Microbiol.* **73**, 650-662.

Cicero, D.O., Contessa, G.M., Pertinhez, T.A., Gallo, M., Katsuyama, A.M., Paci, M., Farah, C.S., and Spisni, A. (2007) Solution structure of ApaG from *Xanthomonas axonopodis* pv. *citri* reveals a fibronectin-3 fold. *Proteins* **67**, 490-500.

Cortes-Bratti, X., Chaves-Olarte, E., Lagergard, T., and Thelestam, M. (2000) Cellular internalization of cytolethal distending toxin from *Haemophilus ducreyi*. *Infect. Immun.* **68**, 6903-6911.

Cortes-Bratti, X., Karlsson, C., Lagergard, T., Thelestam, M., and Frisan T. (2001) The *Haemophilus ducreyi* cytolethal distending toxin induces cell cycle arrest and apoptosis via the DNA damage checkpoint pathways. *J. Biol. Chem.* **276**, 5296-5302.

Cron, L.E., Bootsma, H.J., Noske, N., Burghout, P., Hammerschmidt, S., and Hermans, P.W. (2009) Surface-associated lipoprotein PpmA of *Streptococcus pneumoniae* is involved in colonization in a strain-specific manner. *Microbiology* **155**, 2401-2410.

Dawson, J.R. and Ellen, R.P. (1990) Tip-oriented adherence of *Treponema denticola* to fibronectin. *Infect. Immun.* **58**, 3924-3928.

Dawson, J.R. and Ellen, R.P. (1994) Clustering of fibronectin adhesions toward *Treponema denticola* tips upon contact with immobilized fibronectin. *Infect. Immun.* **62**, 2214-2221.

Dillon, S. C. and Bateman, A. (2004) The hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* **5**, 109.

Dorrell N., Mangan, J.A., Laing, K.G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B.G., Parkhill, J., Stoker, N.G., Karlyshev, A.V., Butcher, P.D., and Wren, B.W. (2001) Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Res.* **11**, 1706-1715.

Dyson, H.J. and Wright, P.E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**, 197-208.

Elwell, C.A. and Dreyfus, L.A. (2000) DNase I homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest. *Mol. Microbiol.* **37**, 952-963.

Endtz, H.P., Ruijs, G.J., van Klingeren, B., Jansen, W.H., van der Reyden, T., and Mouton, R.P. (1991) Quinolone-resistance in *Campylobacter* isolated from man and poultry following the introduction of fluoroquinolones in veterinary medicine. *J. Antimicrob. Chemother.* **27**, 199-208.

- Feng, W., Pan, L., and Zhang, M. (2011) Combination of NMR spectroscopy and X-ray crystallography offers unique advantages for elucidation of the structural basis of protein complex assembly. *Sci. China. Life. Sci.* **54**, 101-111.
- Flanagan, R.C., Neal-McKinney, J.M., Dhillon, A.S., Miller, W.G., and Konkel, M.E. (2009) Examination of *Campylobacter jejuni* putative adhesins leads to the identification of a new protein, designated FlpA, required for chicken colonization. *Infect. Immun.* **77**, 2399-2407.
- Fröman, G., Switalski, L.M., Faris, A., Wadström, T., and Höök, M. (1984) Binding of *Escherichia coli* to fibronectin. *J. Biol. Chem.* **259**, 14899-14905.
- Fry, B.N., Feng, S., Chen, Y.Y., Newell, D.G., Coloe, P.J., and Korolik, V. (2000) The *galE* gene of *Campylobacter jejuni* is involved in lipopolysaccharide synthesis and virulence. *Infect. Immun.* **68**, 2594-2601.
- Gilbert, M., Mandrell, R.E., Parker, C.T., Li, J., and Vinogradov, E. (2007) Structural analysis of the capsular polysaccharide from *Campylobacter jejuni* RM1221. *Chembiochem.* **8**, 625-631.
- Golden, N.J. and Acheson, D.W. (2002) Identification of motility and autoagglutination *Campylobacter jejuni* mutants by random transposon mutagenesis. *Infect. Immun.* **70**, 1761-1771.
- Goon, S., Kelly, J.F., Logan, S.M., Ewing, C.P., and Guerry, P. (2003) Pseudaminic acid, the major modification on *Campylobacter* flagellin, is synthesized via the Cj1293 gene. *Mol. Microbiol.* **50**, 659-671.
- Goon, S., Ewing, C.P., Lorenzo, M., Pattarini, D., Majam, G., and Guerry, P. (2006) A sigma28-regulated nonflagella gene contributes to virulence of *Campylobacter jejuni* 81-176. *Infect. Immun.* **74**, 769-772.
- Guerry, P., Alm, R.A., Power, M.E., Logan, S.M., and Trust, T.J. (1991) Role of two flagellin genes in *Campylobacter* motility. *J. Bacteriol.* **173**, 4757-4764.
- Guerry, P., Szymanski, C.M., Prendergast, M.M., Hickey, T.E., Ewing, C.P., Pattarini, D.L., and Moran, A.P. (2002) Phase variation of *Campylobacter jejuni* 81-176 lipooligosaccharide affects ganglioside mimicry and invasiveness *in vitro*. *Infect. Immun.* **70**, 787-793.
- Guerry, P., Ewing, C.P., Schirm, M., Lorenzo, M., Kelly, J., Pattarini, D., Majam, G., Thibault, P., and Logan, S. (2006) Changes in flagellin glycosylation affect *Campylobacter* autoagglutination and virulence. *Mol. Microbiol.* **60**, 299-311.

- Guerry, P. and Herrmann, T. (2011) Advances in automated NMR protein structure determination. *Q. Rev. Biophys.* **44**, 257-309.
- Halaby, D.M., Poupon, A., and Mornon, J. (1999) The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng.* **12**, 563-571.
- Hamilton, A., Robinson, C., Sutcliffe, I.C., Slater, J., Maskell, D.J., Davis-Poynter, N., Smith, K., Waller, A., and Harrington, D.J. (2006) Mutation of the maturase lipoprotein attenuates the virulence of *Streptococcus equi* to a greater extent than does loss of general lipoprotein lipidation. *Infect. Immun.* **74**, 6907-6919.
- Hanniffy, O.M., Shashkov, A.S., Moran, A.P., Prendergast, M.M., Senchenkova, S.N., Knirel, Y.A., and Savage, A.V. (1999) Chemical structure of a polysaccharide from *Campylobacter jejuni* 176.83 (serotype O:41) containing only furanose sugars. *Carbohydr. Res.* **319**, 124-132.
- Hantke, K., and Braun, V. (1973) Covalent binding of lipid to protein Diglyceride and amide-linked fatty acid at the N-terminal end of the murein lipoprotein of the *Escherichia coli* outer membrane. *Eur. J. Biochem.* **34**, 284-296.
- Hendrixson, D.R., Akerley, B.J., and DiRita, V.J. (2001) Transposon mutagenesis of *Campylobacter jejuni* identifies a bipartite energy taxis system required for motility. *Mol. Microbiol.* **40**, 214-224.
- Hendrixson, D.R. and DiRita, V.J. (2003) Transcription of sigma54-dependent but not sigma28-dependent flagellar genes in *Campylobacter jejuni* is associated with formation of the flagellar secretory apparatus. *Mol Microbiol.* **50**, 687-702.
- Hendrixson, D.R. and DiRita, V.J. (2004) Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol. Microbiol.* **52**, 471-484.
- Holm, L. and Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478-480.
- Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545-549.
- Hunte, C. and Michel, H. (2002) Crystallisation of membrane proteins mediated by antibody fragments. *Curr. Opin. Struct. Biol.* **12**, 503-508.
- Hutchings, M.I., Palmer, T., Harrington, D.J., and Sutcliffe, I.C. (2009) Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. *Trends Microbiol.* **17**, 13-21.

- Inzana, T.J., Seifert, W.E., and Williams, R.P. (1985) Composition and antigenic activity of the oligosaccharide moiety of *Haemophilus influenzae* type b lipooligosaccharide. *Infect. Immun.* **48**, 324-330.
- Jaffe, J., Natanson-Yaron, S., Caparon, M. G., and Hanski, E. (1996) Protein F2, a novel fibronectin-binding protein from *Streptococcus pyogenes*, possesses two binding domains. *Mol. Microbiol.* **21**, 373-384.
- Jenkinson, H. F. (1994) Cell surface protein receptors in oral streptococci. *FEMS Microbiol. Lett.* **121**, 133-140.
- Jennings, H.J., Bhattacharjee, A.K., Kenne, L., Kenny, C.P., and Calver, G. (1980) The R-type lipopolysaccharides of *Neisseria meningitidis*. *Can. J. Biochem.* **58**, 128-136.
- Jin, S., Joe, A., Lynett, J., Hani, E.K., Sherman, P., and Chan, V.L. (2001) JlpA, a novel surface-exposed lipoprotein specific to *Campylobacter jejuni*, mediates adherence to host epithelial cells. *Mol. Microbiol.* **39**, 1225-1236.
- Jin, S., Song, Y.C., Emili, A., Sherman, P.M., and Chan, V.L. (2003) JlpA of *Campylobacter jejuni* interacts with surface-exposed heat shock protein 90alpha and triggers signalling pathways leading to the activation of NF-kappaB and p38 MAP kinase in epithelial cells. *Cell Microbiol.* **5**, 165-174.
- Johnson, W.M. and Lior, H. (1988) A new heat-labile cytolethal distending toxin (CLDT) produced by *Campylobacter* spp. *Microb. Pathog.* **4**, 115-126.
- Kanipes, M.I., Tan, X., Akelaitis, A., Li, J., Rockabrand, D., Guerry, P., and Monteiro, M.A. (2008) Genetic analysis of lipo-oligosaccharide core biosynthesis in *Campylobacter jejuni* 81-176. *J. Bacteriol.* **190**, 1568-1574.
- 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.
- Karlyshev, A.V., McCrossan, M.V., and Wren, B.W. (2001) Demonstration of polysaccharide capsule in *Campylobacter jejuni* using electron microscopy. *Infect. Immun.* **69**, 5921-5924.
- Karlyshev, A.V., Champion, O.L., Churcher, C., Brisson, J.R., Jarrell, H.C., Gilbert, M., Brochu, D., St Michael, F., Li, J., Wakarchuk, W.W., Goodhead, I., Sanders, M., Stevens, K., White, B., Parkhill, J., Wren, B.W., and Szymanski, C.M. (2005) Analysis of *Campylobacter jejuni* capsular loci reveals multiple mechanisms for the generation of structural diversity and the ability to form complex heptoses. *Mol. Microbiol.* **55**, 90-103.

- Kawai, F., Grass, S., Kim, Y., Choi, K.J., St Geme, J.W. 3rd, and Yeo, H.J. (2011) Structural insights into the glycosyltransferase activity of the *Actinobacillus pleuropneumoniae* HMW1C-like protein. *J. Biol. Chem.* **286**, 38546-38557.
- Kawai, F., Paek, S., Choi, K.J., Prouty, M., Kanipes, M.I., Guerry, P., and Yeo, H.J. (2012) Crystal structure of JlpA, a surface-exposed lipoprotein adhesin of *Campylobacter jejuni*. *J. Struct. Biol.* **177**, 583-588.
- Kendrew, J.C., Dickerson, R.E., Strandberg, B.E., Hart, R.G., Davies, D.R. Phillips, D.C., and Shore, V.C. (1960) Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422-427.
- Kervella, M., Pagès, J.M., Pei, Z., Grollier, G., Blaser, M.J., and Fauchère, J.L. (1993) Isolation and characterization of two *Campylobacter* glycine-extracted proteins that bind to HeLa cell membranes. *Infect. Immun.* **61**, 3440-3448.
- Khandavilli, S., Homer, K.A., Yuste, J., Basavanna, S., Mitchell, T., and Brown, J.S. (2008) Maturation of *Streptococcus pneumoniae* lipoproteins by a type II signal peptidase is required for ABC transporter function and full virulence. *Mol. Microbiol.* **67**, 541-557.
- Konkel, M.E., Garvis, S.G. Tipton, S.L. Anderson, D.E. Jr., and Cieplak, W. Jr. (1997) Identification and molecular cloning of a gene encoding a fibronectin-binding protein (CadF) from *Campylobacter jejuni*. *Mol. Microbiol.* **24**, 953-963.
- Konkel, M.E., Kim, B.J., Rivera-Amill, V., and Garvis, S.G. (1999) Bacterial secreted proteins are required for the internalization of *Campylobacter jejuni* into cultured mammalian cells. *Mol. Microbiol.* **32**, 691-701.
- Kovacs-Simon, A., Titball, R.W., and Michell, S.L. (2011) Lipoproteins of bacterial pathogens. *Infect. Immun.* **79**, 548-561.
- Kunishima, N., Asada, Y., Sugahara, M., Ishijima, J., Nodake, Y., Sugahara, M., Miyano, M., Kuramitsu, S., Yokoyama, S., and Sugahara, M. (2005) A novel induced-fit reaction mechanism of asymmetric hot dog thioesterase PAAI. *J. Mol. Biol.* **352**, 212-228.
- Kuusela, P. (1978) Fibronectin binds to *Staphylococcus aureus*. *Nature* **276**, 718-720.
- Lampen, J.O. and Nielsen, J.B. (1984) N-terminal glyceride-cysteine modification of membrane penicillinases in gram-positive bacteria. *Methods Enzymol.* **106**, 365-368.
- Lara-Tejero, M. and Galán, J.E. (2000) A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. *Science* **290**, 354-357.
- Lara-Tejero, M. and Galán, J.E. (2001) CdtA, CdtB, and CdtC form a tripartite complex that is required for cytolethal distending toxin activity. *Infect. Immun.* **69**, 4358-4365.



Lecuit, M., Abachin, E., Martin, A., Poyart, C., Pochart, P., Suarez, F., Bengoufa, D., Feuillard, J., Lavergne, A., Gordon, J.I., Berche, P., Guillevin, L., and Lortholary, O. (2004) Immunoproliferative small intestinal disease associated with *Campylobacter jejuni*. *N. Engl. J. Med.* **350**, 239-248.

Lee, R.B., Hassane, D.C., Cottle, D.L., and Pickett, C.L. (2003) Interactions of *Campylobacter jejuni* cytolethal distending toxin subunits CdtA and CdtC with HeLa cells. *Infect. Immun.* **71**, 4883-4890.

Leesong, M., Henderson, B.S., Gillig, J.R., Schwab, J.M., and Smith, J.L. (1996) Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **4**, 253-264.

Marra, A., Lawson, S., Asundi, J.S., Brigham, D. and Hromockyj, A.E. (2002) *In vivo* characterization of the psa genes from *Streptococcus pneumoniae* in multiple models of infection. *Microbiology* **148**, 1483-1491.

Marsden, G.L., Li, J., Everest, P.H., Lawson, A.J., and Ketley, J.M. (2009) Creation of a large deletion mutant in *Campylobacter jejuni* reveals the lipooligosaccharide gene cluster is not required for viability. *J. Bacteriol.* **191**, 2392-2399.

Mathiopoulos, C., Mueller, J.P., Slack, F.J., Murphy, C.G., Patankar, S., Bukusoglu, G., and Sonenshein, A.L. (1991) A *Bacillus subtilis* dipeptide transport system expressed early during sporulation. *Mol. Microbiol.* **5**, 1903-1913.

Matsuyama, S., Tajima, T., and Tokuda, H. (1995) A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. *EMBO J.* **14**, 3365-3372.

Matsuyama, S., Yokota, N., and Tokuda, H. (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. *EMBO J.* **16**, 6947-6955.

McNally, D.J., Jarrell, H.C., Li, J., Khieu, N.H., Vinogradov, E., Szymanski, C.M., and 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. *FEBS J.* **272**, 4407-4422.

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.

- Misawa, N., Kawashima, K., Kondo, F., Allos, B.M., and Blaser, M.J. (2001) DNA diversity of the *wla* gene cluster among serotype HS:19 and non-HS:19 *Campylobacter jejuni* strains. *J. Endotoxin Res.* **7**, 349-358.
- Monteville, M.R., Yoon, J.E., and Konkel, M.E. (2003) Maximal adherence and invasion of INT 407 cells by *Campylobacter jejuni* requires the CadF outer-membrane protein and microfilament reorganization. *Microbiology* **149**, 153-165.
- Moore, J.E., Barton, M.D., Blair, I.S., Corcoran, D., Dooley, J.S., Fanning, S., Kempf, I., Lastovica, A.J., Lowery, C.J., Matsuda, M., McDowell, D.A., McMahon, A., Millar, B.C., Rao, J.R., Rooney, P.J., Seal, B.S., Snelling, W.J., and Tolba, O. (2006) The epidemiology of antibiotic resistance in *Campylobacter*. *Microbes Infect.* **8**, 1955-1966.
- Moran, A.P., Prendergast, M.M., and Appelmelk, B.J. (1996) Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease. *FEMS Immunol. Med. Microbiol.* **16**, 105-115.
- Morooka, T., Umeda, A., and Amako, K. (1985) Motility as an intestinal colonization factor for *Campylobacter jejuni*. *J. Gen. Microbiol.* **131**, 1973-1980.
- Moser, I., Schroeder, W., and Salnikow, J. (1997) *Campylobacter jejuni* major outer membrane protein and a 59-kDa protein are involved in binding to fibronectin and INT 407 cell membranes. *FEMS Microbiol. Lett.* **157**, 233-238.
- Muldoon, J., Shashkov, A.S., Moran, A.P., Ferris, J.A., Senchenkova, S.N., and Savage, A.V. (2002) Structures of two polysaccharides of *Campylobacter jejuni* 81116. *Carbohydr. Res.* **337**, 2223-2229.
- Müller, A., León-Kempis, M. del R., Dodson, E., Wilson, K.S., Wilkinson, A.J., and Kelly, D.J. (2007) A bacterial virulence factor with a dual role as an adhesin and a solute-binding protein: the crystal structure at 1.5 Å resolution of the PEB1a protein from the food-borne human pathogen *Campylobacter jejuni*. *J. Mol. Biol.* **372**, 160-171.
- Myhre, E. B. and Kuusela, P. (1983) Binding of human fibronectin to group A, C, and G *Streptococci*. *Infect. Immun.* **40**, 29-34.
- Nachamkin, I., Yang, X.H., and Stern, N.J. (1993) Role of *Campylobacter jejuni* flagella as colonization factors for three-day-old chicks: analysis with flagellar mutants. *Appl. Environ. Microbiol.* **59**, 1269-1273.
- Nachamkin, I., Liu, J., Li, M., Ung, H., Moran, A.P., Prendergast, M.M., Sheikh, K. (2002) *Campylobacter jejuni* from patients with Guillain-Barré syndrome preferentially expresses a GD(1a)-like epitope. *Infect. Immun.* **70**, 5299-5303.

- Naess, V. and Hofstad, T. (1984) Chemical composition and biological activity of lipopolysaccharides prepared from type strains of *Campylobacter jejuni* and *Campylobacter coli*. *Acta Pathol. Microbiol. Immunol. Scand. B* **92**, 217-222.
- Naito, M., Fridrich, E., Fields, J.A., Pryjma, M., Li, J., Cameron, A., Gilbert, M., Thompson, S.A., and Gaynor, E.C. (2010) Effects of sequential *Campylobacter jejuni* 81-176 lipooligosaccharide core truncations on biofilm formation, stress survival, and pathogenesis. *J. Bacteriol.* **192**, 2182-2192.
- Narita, S., S. Matsuyama, and H. Tokuda. (2004) Lipoprotein trafficking in *Escherichia coli*. *Arch. Microbiol.* **182**, 1-6.
- Nicholls, A., Sharp, K.A., and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.
- Oakland, M., Jeon, B., Sahin, O., Shen, Z., and Zhang, Q. (2011) Functional characterization of a lipoprotein-encoding operon in *Campylobacter jejuni*. *PLoS One.* **6**, e20084.
- Okuda, S. and Tokuda, H. (2011) Lipoprotein sorting in bacteria. *Annu. Rev. Microbiol.* **65**, 239-259.
- Ostermeier, C., Iwata, S., Ludwig, B., and Michel, H. (1995) Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase. *Nat. Struct. Biol.* **2**, 842-846.
- Paek, S., Kawai, F., Choi, K.J., and Yeo, H.J. (2012) Crystal structure of the *Campylobacter jejuni* Cj0090 protein reveals a novel variant of the immunoglobulin fold among bacterial lipoproteins. *Proteins: Struct., Funct., Bioinfo.* **80**, 2804-2809.
- Parker, C.T., Gilbert, M., Yuki, N., Endtz, H.P., and Mandrell, R.E. (2008) Characterization of lipooligosaccharide-biosynthetic loci of *Campylobacter jejuni* reveals new lipooligosaccharide classes: evidence of mosaic organizations. *J. Bacteriol.* **190**, 5681-5689.
- 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.
- Pei, Z. and Blaser, M. J. (1993) PEB1, the major cell-binding factor of *Campylobacter jejuni*, is a homolog of the binding component in gram-negative nutrient transport systems. *J. Biol. Chem.* **268**, 18717-18725.

Pei, Z., Burucoa, C., Grignon, B., Baqar, S., Huang, X.Z., Kopecko, D.J., Bourgeois, A.L., Fauchere, J.L., and Blaser, M.J. (1998) Mutation in the *peb1A* locus of *Campylobacter jejuni* reduces interactions with epithelial cells and intestinal colonization of mice. *Infect. Immun.* **66**, 938-943.

Perego, M., Higgins, C.F., Pearce, S.R., Gallagher, M.P., and Hoch, J.A. (1991) The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol. Microbiol.* **5**, 173-185.

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G., and North, A.C. (1960) Structure of haemoglobin: A three-dimensional fourier synthesis at 5.5 Å resolution, obtained by X-Ray analysis. *Nature* **185**, 416-422.

Petit, C.M., Brown, J.R., Ingraham, K., Bryant, A.P., and Holmes, D.J. (2001) Lipid modification of prelipoproteins is dispensable for growth *in vitro* but essential for virulence in *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.* **200**, 229-233.

Poly, F., Ewing, C., Goon, S., Hickey, T.E., Rockabrand, D., Majam, G., Lee, L., Phan, J., Savarino, N.J., and Guerry, P. (2007) Heterogeneity of a *Campylobacter jejuni* protein that is secreted through the flagellar filament. *Infect. Immun.* **75**, 3859-3867.

Poly, F. and Guerry, P. (2008) Pathogenesis of *Campylobacter*. *Curr. Opin. Gastroenterol.* **24**, 27-31.

Ray, A., Redhead, K., Selkirk, S., and Poole, S. (1991) Variability in LPS composition, antigenicity and reactogenicity of phase variants of *Bordetella pertussis*. *FEMS Microbiol. Lett.* **63**, 211-217.

Rinaldi, S. and Willison, H.J. (2008) Ganglioside antibodies and neuropathies. *Curr. Opin. Neurol.* **21**, 540-546.

Rydén, C., Rubin, K., Speziale, P., Höök, M., Lindberg, M., and Wadström, T. (1983). Fibronectin receptors from *Staphylococcus aureus*. *J. Biol. Chem.* **258**, 3396-3401.

Schirm, M., Schoenhofen, I.C., Logan, S.M., Waldron, K.C., and Thibault, P. (2005) Identification of unusual bacterial glycosylation by tandem mass spectrometry analyses of intact proteins. *Anal. Chem.* **77**, 7774-7782.

Schorey, J. S., Holsti, M. A., Ratliff, T. L., Allen, P. M. and Brown, E. J. (1996). Characterization of the fibronectin-attachment protein of *Mycobacterium avium* reveals a fibronectin-binding motif conserved among mycobacteria. *Mol. Microbiol.* **21**, 321-329.

Schujman, G.E., Guerin, M., Buschiazzi, A., Schaeffer, F., Llarrull, L.I., Reh, G., Vila, A.J., Alzari, P.M., and de Mendoza, D. (2006) Structural basis of lipid biosynthesis regulation in Gram-positive bacteria. *EMBO J.* **25**, 4074-4083.

Scott, N.E., Bogema, D.R., Connolly, A.M., Falconer, L., Djordjevic, S.P., and Cordwell, S.J. (2009) Mass spectrometric characterization of the surface-associated 42 kDa lipoprotein JlpA as a glycosylated antigen in strains of *Campylobacter jejuni*. *J. Proteome Res.* **8**, 4654-4664.

Shore, D.A., Teyton, L., Dwek, R.A., Rudd, P.M., and Wilson, I.A. (2006) Crystal structure of the TCR co-receptor CD8 $\alpha\alpha$  in complex with monoclonal antibody YTS 105.18 Fab fragment at 2.88 Å resolution. *J. Mol. Biol.* **358**, 347-354.

Shore, D.A., Issafras, H., Landais, E., Teyton, L., and Wilson, I.A. (2008) The crystal structure of CD8 in complex with YTS156.7.7 Fab and interaction with other CD8 antibodies define the binding mode of CD8  $\alpha\beta$  to MHC class I. *J. Mol. Biol.* **384**, 1190-1202.

Snyder, D.A., Chen, Y., Denissova, N.G., Acton, T., Aramini, J.M., Ciano, M., Karlin, R., Liu, J., Manor, P., Rajan, P.A., Rossi, P., Swapna, G.V., Xiao, R., Rost, B., Hunt, J., and Montelione, G.T. (2005) Comparisons of NMR spectral quality and success in crystallization demonstrate that NMR and X-ray crystallography are complementary methods for small protein structure determination. *J. Am. Chem. Soc.* **127**, 16505-16511.

Solmaz, S.R. and Hunte, C. (2008) Structure of complex III with bound cytochrome c in reduced state and definition of a minimal core interface for electron transfer. *J. Biol. Chem.* **283**, 17542-17549.

Sommerlad, S.M. and Hendrixson, D.R. (2007) Analysis of the roles of FlgP and FlgQ in flagellar motility of *Campylobacter jejuni*. *J. Bacteriol.* **189**, 179-186.

Song, Y.C., Jin, S., Louie, H., Ng, D., Lau, R., Zhang, Y., Weerasekera, R., Al Rashid, S., Ward, L.A., Der, S.D., Chan, V.L. (2004) FlaC, a protein of *Campylobacter jejuni* TGH9011 (ATCC43431) secreted through the flagellar apparatus, binds epithelial cells and influences cell invasion. *Mol. Microbiol.* **53**, 541-553.

Steyn, A.J., Joseph, J., and Bloom, B.R. (2003) Interaction of the sensor module of *Mycobacterium tuberculosis* H37Rv KdpD with members of the Lpr family. *Mol. Microbiol.* **47**, 1075-1089.

Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F., Daffé, M., Cambillau, C., Gicquel, B., Bourne, Y., and Jackson, M. (2006) LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO J.* **25**, 1436-1444.

Sutcliffe, I.C. and Russell, R.R.B. (1995) Lipoproteins of Gram-positive bacteria. *J. Bacteriol.* **177**, 1123-1128.

Szymanski, C.M., Burr, D.H., and Guerry, P. (2002) *Campylobacter* protein glycosylation affects host cell interactions. *Infect. Immun.* **70**, 2242-2244.

Szymanski, C.M., St Michael, F., Jarrell, H.C., Li, J., Gilbert, M., Larocque, S., Vinogradov, E. and Brisson, J.R. (2003). Detection of conserved N-linked glycans and phase variable lipo-oligosaccharides and capsules from *Campylobacter* cells by mass spectrometry and high resolution magic angle spinning NMR spectroscopy. *J. Biol. Chem.* **278**, 24509-24520.

Taylor, P.D., Toseland, C.P., Attwood, T.K., and Flower, D.R. (2006) LIPPRED: A web server for accurate prediction of lipoprotein signal sequences and cleavage sites. *Bioinformatics* **1**, 176-179.

Thibault, P., Logan, S.M., Kelly, J.F., Brisson, J.R., Ewing, C.P., Trust, T.J., and Guerry, P. (2001) Identification of the carbohydrate moieties and glycosylation motifs in *Campylobacter jejuni* flagellin. *J. Biol. Chem.* **276**, 34862-34870.

Thoden, J.B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H.M. (2003) The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU. *J. Biol. Chem.* **278**, 43709-43716.

Thomas, D. D., Baseman, J. B. and Alderete, J. F. (1985). Fibronectin mediates *Treponema pallidum* cytoadherence through recognition of fibronectin cell-binding domain. *J. Exp. Med.* **161**, 514-525.

Tidhar, A., Flashner, Y., Cohen, S., Levi, Y., Zauberman, A., Gur, D., Aftalion, M., Elhanany, E., Zvi, A., Shafferman, A., and Mamroud, E. (2009) The NlpD lipoprotein is a novel *Yersinia pestis* virulence factor essential for the development of plague. *PLoS One* **4**, e7023.

Tipps, M.E., Lawshe, J.E., Ellington, A.D., Mihic, S.J. (2010) Identification of novel specific allosteric modulators of the glycine receptor using phage display. *J. Biol. Chem.* **285**, 22840-22845.

Tjalsma, H., Kontinen, V.P., Prágai, Z., Wu, H., Meima, R., Venema, G., Bron, S., Sarvas, M., and van Dijk, J.M. (1999) The role of lipoprotein processing by signal peptidase II in the Gram-positive eubacterium *Bacillus subtilis*. Signal peptidase II is required for the efficient secretion of alpha-amylase, a non-lipoprotein. *J. Biol. Chem.* **274**, 1698-1707.

- Tokuda, H. (2009) Biogenesis of outer membranes in Gram-negative bacteria. *Biosci. Biotechnol. Biochem.* **73**, 465-473.
- Tu, Q.V., McGuckin, M.A., and Mendz, G.L. (2008) *Campylobacter jejuni* response to human mucin MUC2: modulation of colonization and pathogenicity determinants. *J. Med. Microbiol.* **57**, 795-802.
- Vandeputte-Rutten, L., Bos, M.P., Tommassen, J., and Gros, P. (2003) Crystal structure of Neisserial surface protein A (NspA), a conserved outer membrane protein with vaccine potential. *J. Biol. Chem.* **278**, 24825-24830.
- van Putten, J. P., Duensing, T. D. and Cole, R. L. (1998). Entry of OpaA<sup>+</sup> gonococci into HEp-2 cells requires concerted action of glycosaminoglycans, fibronectin and integrin receptors. *Mol. Microbiol.* **29**, 369-379.
- Visai, L., Bozzini, S., Petersen, T.E., Speciale, L., and Speciale, P. (1991) Binding sites in fibronectin for an enterotoxigenic strain of *E. coli* B342289c. *FEBS Lett.* **290**, 111-114.
- Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F., and Jones, D.T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138-2139.
- Wassenaar, T.M., Bleumink-Pluym, N.M., and van der Zeijst, B.A. (1991) Inactivation of *Campylobacter jejuni* flagellin genes by homologous recombination demonstrates that *fla A* but not *flaB* is required for invasion. *EMBO J.* **10**, 2055-2061.
- Wassenaar, T.M., van der Zeijst, B.A., Ayling, R., and Newell, D.G. (1993) Colonization of chicks by motility mutants of *Campylobacter jejuni* demonstrates the importance of flagellin A expression. *J. Gen. Microbiol.* **139**, 1171-1175.
- Whitehouse, C.A., Balbo, P.B., Pesci, E.C., Cottle, D.L., Mirabito, P.M., and Pickett, C.L. (1998) *Campylobacter jejuni* cytolethal distending toxin causes a G2-phase cell cycle block. *Infect. Immun.* **66**, 1934-1940.
- Wösten, M.M., Wagenaar, J.A., van Putten, J.P. (2004) The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J. Biol. Chem.* **279**, 16214-16222.
- Xu, Q., McShan, K., and Liang, F.T. (2008) Essential protective role attributed to the surface lipoproteins of *Borrelia burgdorferi* against innate defences. *Mol. Microbiol.* **69**, 15-29.
- Yamaguchi, K., Yu, F., and Inouye, M. (1988) A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **53**, 423-432.

Yamaguchi, S., Fujita, H., Ishihara, A., Aizawa S., Macnab R.M. (1986) Subdivision of flagellar genes of *Salmonella typhimurium* into regions responsible for assembly, rotation, and switching. *J. Bacteriol.* **166**, 187-193.

Yang, X., Wu, Z., Wang, X., Yang, C., Xu, H. and Shen, Y. (2009) Crystal structure of lipoprotein GNA1946 from *Neisseria meningitidis*. *J. Struct. Biol.* **168**, 437-443.

Yee, A.A., Savchenko, A., Ignachenko, A., Lukin, J., Xu, X., Skarina, T., Evdokimova, E., Liu, C.S., Semesi, A., Guido, V., Edwards, A.M., and Arrowsmith, C.H. (2005) NMR and X-ray crystallography, complementary tools in structural proteomics of small proteins. *J. Am. Chem. Soc.* **127**, 16512-16517.

Yeo, H.J. (2013) Production and crystallization of bacterial type V secretion proteins. *Methods Mol. Biol.* **966**, 205-222.

Yokoyama, T., Paek, S., Ewing, C.P., Guerry, P., and Yeo, H.J. (2008) Structure of a sigma28-regulated nonflagellar virulence protein from *Campylobacter jejuni*. *J. Mol. Biol.* **384**, 364-376.

Yonath, A. (2011) X-ray crystallography at the heart of life science. *Curr. Opin. Struct. Biol.* **21**, 622-626.

Young, K.T., Davis, L.M., and Dirita, V.J. (2007) *Campylobacter jejuni*: molecular biology and pathogenesis. *Nat. Rev. Microbiol.* **5**, 665-679.

Yuki, N. (2005) Carbohydrate mimicry: a new paradigm of autoimmune diseases. *Curr. Opin. Immunol.* **17**, 577-582.

Zilbauer, M., Dorrell, N., Wren, B.W., and Bajaj-Elliott, M. (2008) *Campylobacter jejuni*-mediated disease pathogenesis: an update. *Trans R. Soc. Trop. Med. Hyg.* **102**, 123-129.

Ziprin, R.L., Young, C.R., Stanker, L.H., Hume, M.E., and Konkel, M.E. (1999) The absence of cecal colonization of chicks by a mutant of *Campylobacter jejuni* not expressing bacterial fibronectin-binding protein. *Avian Dis.* **43**, 586-589.