Studies on the structure and composition of the outer membrane of *Caulobacter crescentus*

by

Michael D Jones

B. Science, Specialization in Biotechnology, University of Alberta, 2006
M. Science, Pharmaceutical Sciences, University of Alberta, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Microbiology and Immunology)

The University of British Columbia
(Vancouver)

October 2015

© Michael D Jones, 2015
Abstract

Classically, outer membranes are half lipid, half protein, and the outmost layers of Gram-negative bacteria. For *Caulobacter crescentus* the outer membrane is the penultimate layer beneath a protein surface layer (S-layer). The S-layer of the caulobacter cell envelope is an exciting platform for high density peptide display and biotechnology development. We focused on elucidating the structure of the outer membrane by crystallizing the S-layer protein, RsaA; solving the structure of the lipopolysaccharide; and characterizing a newly discovered porin, OmpW.

S-layer proteins are highly resistant to crystallization, because two-dimensional S-layer formation out competes three-dimensional crystal formation. To achieve a crystallisable form of RsaA, a C-terminal truncation version was constructed and expressed in the native host, *C. crescentus*. The secreted protein was prone to aggregation, so low agitation and slow concentration protocols had to be developed. The RsaA truncate produced large crystals that diffracted to <2.5 Å. Solving the phases proved to be a serious hurdle and the final protein structure remains unsolved.

The lipopolysaccharide of *C. crescentus* is the anchor that supports the S-layer. The structure of the lipid A portion was solved previously but structures for the core oligosaccharide and the O-polysaccharide had not been deduced. In collaboration with Dr. Evgeny Vinogradov, these remaining structures were solved. The core oligosaccharide has a branched heptasaccharide structure. The O-polysaccharide is a heptasaccharide containing the dideoxy sugar N-acetylperosamine. Additionally, a rhamnan polysaccharide was discovered and its structure
was determined.

Porins, non-specific passive protein channels, are significant components of classical Gram-negative outer membranes. Despite this, no porin had ever been identified in *C. crescentus*. We report the identification and characterization of the porin OmpW in *C. crescentus*. OmpW has low conductance of 125 pSv in 1 M KCl. That is interesting because homologous porins in other bacteria have no detectable pore-forming activity.

The cell envelopes of bacteria are remarkable structures; the work here illuminates the unique structures present in the caulobacter envelope.
Preface

“Plenum ingenni pudoris fateri per quos profeceris”
(To own up to those who were the means of one’s own achievements)
Pliny the Elder, Natural History

The contents contained herein are original work by or in collaboration with me, the primary author, except as specified below.

Chapter 2  Portions of chapter 2 have been used for the preparation of a manuscript for submission for publication. The work was done in collaboration with Dr. Anson Chan. I performed all the experiments and wrote the majority of the manuscript. Dr. Chan assisted in the X-ray diffraction data collection, the data evaluation and processing, and edited and provided feedback on the writing of the manuscript.


The work was done as a collaboration with Dr. Evgeny Vinogradov. Dr. Vinogradov was responsible for the majority of the structural work: monosaccharide analysis, methylation analysis, and NMR analysis, as well as the written summaries of those methods. Dr. John Nomellini constructed the strain used, JS1025. The cytokine and cell culture experiments were new additions not present in our publication of the LPS structures. The execution of the HEK-Blue assay was performed by Gyles Ifill. I was responsible for the preparation of all the LPS
samples, experimental design, and writing the manuscript.

*Chapter 4* A version of chapter 4 has been submitted for publication. I was responsible for the sample preparation, purification, and strain construction. The laboratory of Dr. Roland Benz was responsible for synthetic membrane conductance experiments. The writing of the resulting manuscript was a joint effort.
# Table of Contents

Abstract .................................................................................................................. ii

Preface .................................................................................................................... iv

Table of Contents ..................................................................................................... vi

List of Tables .......................................................................................................... xi

List of Figures ......................................................................................................... xii

List of Abbreviations ............................................................................................... xv

Acknowledgments .................................................................................................... xix

1 Introduction ........................................................................................................... 1

1.1 S-layers ............................................................................................................. 1
   1.1.1 S-layer structure ..................................................................................... 1
   1.1.2 S-layer proteins ..................................................................................... 6
   1.1.3 Functions of S-layers ........................................................................... 8
   1.1.4 RsaA ..................................................................................................... 11
   1.1.5 Biotechnology applications of S-layers ................................................ 14

1.2 Lipopolysaccharide .......................................................................................... 17
   1.2.1 Lipid A ................................................................................................. 18
4 OmpW of *Caulobacter crescentus* functions as an outer membrane channel

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>91</td>
</tr>
<tr>
<td>4.2 Methods</td>
<td>93</td>
</tr>
<tr>
<td>4.2.1 Growth and maintenance of bacteria</td>
<td>93</td>
</tr>
<tr>
<td>4.2.2 Generation of <em>C. crescentus</em> ∆ompW</td>
<td>94</td>
</tr>
<tr>
<td>4.2.3 Outer membrane-enriched preparations</td>
<td>95</td>
</tr>
<tr>
<td>4.2.4 Crude membrane preparations</td>
<td>96</td>
</tr>
<tr>
<td>4.2.5 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
<td>96</td>
</tr>
<tr>
<td>4.2.6 Isolation and purification of the channel-forming protein from</td>
<td>96</td>
</tr>
<tr>
<td>outer membrane-enriched preparations</td>
<td></td>
</tr>
<tr>
<td>4.2.7 Tryptic digestion and peptide sequencing</td>
<td>97</td>
</tr>
<tr>
<td>4.2.8 Lipid bilayer experiments</td>
<td>97</td>
</tr>
<tr>
<td>4.2.9 Modeling of the OmpW structure</td>
<td>98</td>
</tr>
<tr>
<td>4.3 Results</td>
<td>98</td>
</tr>
<tr>
<td>4.3.1 Protein composition of the outer membrane-enriched preparations of <em>C. crescentus</em></td>
<td>98</td>
</tr>
<tr>
<td>4.3.2 Identification of channel-forming activity in outer membrane-enriched preparations of <em>C. crescentus</em></td>
<td>99</td>
</tr>
<tr>
<td>4.3.3 Identification of the purified channel-forming protein from the enriched outer membranes</td>
<td>100</td>
</tr>
<tr>
<td>4.3.4 Partial sequencing of the 22 kDa protein of <em>C. crescentus</em></td>
<td>101</td>
</tr>
<tr>
<td>4.3.5 Analysis of the channels formed by OmpW channel-forming protein of <em>C. crescentus</em></td>
<td>102</td>
</tr>
<tr>
<td>4.3.6 The OmpW channel of <em>C. crescentus</em> is highly cation-selective</td>
<td>106</td>
</tr>
<tr>
<td>4.3.7 Knocking out <em>ompW</em> removes channel-forming activity</td>
<td>106</td>
</tr>
<tr>
<td>4.4 Discussion</td>
<td>107</td>
</tr>
</tbody>
</table>
5 Conclusions ................................................................. 117
  5.1 Relevance and contribution to the field ............................ 117
  5.2 Future directions ....................................................... 120

Bibliography ............................................................... 125

Appendix ................................................................. 154
List of Tables

Table 2.1  Summary data from the X-ray diffraction of RsaA Δ0–222 .......................... 52

Table 3.1  LPS silver stain procedure ................................................................. 58

Table 3.2  MALDI-TOF settings ................................................................. 60

Table 3.3  NMR data for *C. crescentus* PS1 ....................................................... 77

Table 3.4  NMR data for the minor components of the double oxidized non-deacylated PS 79

Table 3.5  NMR data for *C. crescentus* PS2 and its NaIO₄ oxidation product OS1 81

Table 3.6  NMR data for the core oligosaccharide ............................................. 83

Table 4.1  Average conductance through OmpW .............................................. 105
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Cross-sectional diagrams of S-layer containing cell envelopes</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>The first published image of a S-layer</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>A simple overview of S-layer symmetries</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Electrotomograph of <em>C. crescentus</em></td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>Reconstructed surface of the <em>C. crescentus</em> S-layer</td>
<td>12</td>
</tr>
<tr>
<td>1.6</td>
<td>The structure of RTX motifs</td>
<td>14</td>
</tr>
<tr>
<td>1.7</td>
<td>Simplified structure of LPS from <em>Salmonella minnesota</em></td>
<td>19</td>
</tr>
<tr>
<td>1.8</td>
<td>Kdo&lt;sub&gt;2&lt;/sub&gt;-Lipid A from <em>E. coli</em> K12</td>
<td>20</td>
</tr>
<tr>
<td>1.9</td>
<td>A generalized formula for core oligosaccharides</td>
<td>24</td>
</tr>
<tr>
<td>1.10</td>
<td>A simple overview of the Wzy-dependent pathway</td>
<td>26</td>
</tr>
<tr>
<td>1.11</td>
<td>A simple overview of LPS transport across the periplasm</td>
<td>27</td>
</tr>
<tr>
<td>1.12</td>
<td>TonB-mediated transport versus transport through a porin</td>
<td>29</td>
</tr>
<tr>
<td>2.1</td>
<td>Example of a large cluster of ‘plate’ RsaA Δ0–222 crystals</td>
<td>37</td>
</tr>
<tr>
<td>2.2</td>
<td>Edge plots for useful anomalous dispersion elements</td>
<td>38</td>
</tr>
<tr>
<td>2.3</td>
<td>Purification of RsaA Δ0–222 shown by SDS-PAGE</td>
<td>40</td>
</tr>
<tr>
<td>2.4</td>
<td>DLS plots of RsaA constructs</td>
<td>41</td>
</tr>
<tr>
<td>2.5</td>
<td>Panel of well diffracting ‘plate’ crystals of RsaA Δ0–222</td>
<td>44</td>
</tr>
<tr>
<td>2.6</td>
<td>A panel of ‘pencil crystals of RsaA Δ0–222’</td>
<td>45</td>
</tr>
<tr>
<td>2.7</td>
<td>Examples of unusable ‘dendritic’ RsaA Δ0–222 GCC723 crystals</td>
<td>45</td>
</tr>
</tbody>
</table>
Figure 2.8 ‘Dendritic’ RsaA Δ0–222 GSCC723 crystals that were used for X-ray diffraction ........................................... 46
Figure 2.9 A diffraction pattern of RsaA Δ0–222, ‘plate’ crystal ......................... 48
Figure 3.1 The lipid A of *C. crescentus* ................................................................. 54
Figure 3.2 Four different visualisations of *C. crescentus* LPS ............................ 68
Figure 3.3 Visual comparison of *E. coli* O122 LPS and *C. crescentus* LPS .......... 69
Figure 3.4 MALDI-TOF analysis of intact, whole LPS from *C. crescentus* .......... 70
Figure 3.5 1H NMR spectra of *C. crescentus* OPS .............................................. 73
Figure 3.6 Monosaccharide analysis of LPS ............................................................ 74
Figure 3.7 The structure of the *C. crescentus* OPS .............................................. 75
Figure 3.8 2D NMR of *C. crescentus* PS1 ............................................................ 76
Figure 3.9 The structure of minor component, the end caps of the OPS. ............ 78
Figure 3.10 The structure of the *C. crescentus* rhamnan. ................................... 80
Figure 3.11 The structure of the *C. crescentus* core oligosaccharide. ............... 82
Figure 3.12 Fragment of 1H-13C HSQC spectrum of the core. ......................... 84
Figure 3.13 NF-κB Assay for cellular activation by LPS ................................. 85
Figure 3.14 Cytokine profile comparing *Caulobacter crescentus* and *Salmonella enterica* var. Typhimurium LPS ......................... 86
Figure 4.1 Coomassie stained SDS-PAGE of different protein preparations of *C. crescentus* ......................................................... 99
Figure 4.2 SDS-PAGE showing OmpW from *C. crescentus* obtained by elution of the 20 kDa band ................................................. 101
Figure 4.3 Single-channel measurements with detergent treated outer membrane of C. crescentus and purified 20 kDa protein ......................... 104
Figure 4.4 SDS-PAGE of *C. crescentus* ΔompW .............................................. 108
Figure 4.5 Single-channel conductance of OmpW in 1 M salt solution as a function of the ionic radii of monovalent cations ................................. 111
Figure 4.6  Amino acid sequence alignment of OprG of *Pseudomonas aeruginosa*,
OmpW of *E. coli*, and OmpW of *C. crescentus* ........................................ 113
Figure 4.7  Structural properties of OmpW of *C. crescentus* .......................... 115
Figure 4.8  Comparison of porin channel radii ............................................. 116
Figure 5.1  The complete structure of the *C. crescentus* LPS .......................... 121
Figure A.1  RsaA, amino acid sequence ......................................................... 154
Figure A.2  OmpW, amino acid sequence ....................................................... 155
# List of Abbreviations

## General abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Amino acid</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP binding cassette</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search</td>
</tr>
<tr>
<td>CFG</td>
<td>Consortium of Functional Glycomics</td>
</tr>
<tr>
<td>CLS</td>
<td>Canadian Light Source</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation spectroscopy</td>
</tr>
<tr>
<td>DLS</td>
<td>Dynamic light scattering</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco’s modified Eagle’s medium</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EPS</td>
<td>Extracellular polysaccharide</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal bovine serum</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography-mass spectroscopy</td>
</tr>
<tr>
<td>gCOSY</td>
<td>Gradient correlation spectroscopy</td>
</tr>
<tr>
<td>gHMBC</td>
<td>Gradient heteronuclear multiple bond coherence</td>
</tr>
<tr>
<td>gHSQC</td>
<td>Gradient heteronuclear single quantum coherence</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear multiple bond coherence</td>
</tr>
<tr>
<td>HMQCC</td>
<td>Heteronuclear multiple-quantum correlation spectroscopy</td>
</tr>
</tbody>
</table>
HSQC  Heteronuclear single quantum coherence
IFN-β  Interferon β
IL  Interleukin
JCSG  Joint Centre for Structural Genomics
$K_D$  Dissociation constant
LDAO  Lauryldimethylamine-oxide
LPS  Lipopolysaccharide
MAD  Multi-wavelength anomalous dispersion
MALDI-TOF  Matrix assisted laser desorption/ionization-time of flight mass spectroscopy
MAMP  Microbe-associated molecular pattern
MCP-1  Monocyte chemotactic protein 1
MS  Mass spectrometry
MW  Molecular weight
NF-κB  Nuclear factor kappa-light-chain-enhancer of activated B cells
NMR  Nuclear magnetic resonance spectroscopy
NOE  Nuclear Overhauser enhancement
NOESY  Nuclear Overhauser enhancement spectroscopy
OD$_{600}$  Optical density at 600 nm
OPS  O-specific polysaccharide
OS  Oligosaccharide
PBS  Phosphate-buffered saline
PC  Phosphatidylcholine
PCR  Polymerase chain reaction
PEG  Polyethylene glycol
pI Isoelectric point
PMA  Phorbol 12-myristate 13-acetate
PS  Polysaccharide
PYE  Peptone-yeast extract medium
ROESY  Rotating frame nuclear Overhauser effect spectroscopy
RPMI  Roswell Park Memorial Institute medium
RTX  Repeat-in-toxin
S-layer  Protein surface layer
SAD  Single-wavelength anomalous dispersion
SDS-PAGE  Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SSRL  Stanford Synchotron Radiation Light Source
T1SS  Type 1 secretion system
TFA  Trifluoroacetic acid
TLC  Thin-layer chromatography
TLR4  Toll-like receptor 4
TNF-\(\alpha\)  Tumor necrosis factor \(\alpha\)
TCOSY  Total correlation spectroscopy
UV  Ultraviolet Light

Sugar abbreviations

GalA  Galacturonic acid
GalNAc  N-acetylgalactosamine
Glc  Glucose
GlcA  Glucuronic acid
GlcN  Glucosamine
GlcNAc  N-acetylglycosamine
3-O-MeGlc  3-O-methylglucose
Kdo  3-deoxy-d-manno-octulosonic acid

xvii
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Chemical Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDHep</td>
<td>L-glycero-D-mannoheptose</td>
</tr>
<tr>
<td>LLHep</td>
<td>L-glycero-L-mannoheptose</td>
</tr>
<tr>
<td>Man</td>
<td>Mannose</td>
</tr>
<tr>
<td>PerN</td>
<td>Perosamine</td>
</tr>
<tr>
<td>Rha</td>
<td>Rhamnose</td>
</tr>
</tbody>
</table>
Acknowledgments

First and most of all I’d like to thank my family. My parents, J.C. and Dorothea, your support has been completely indispensable. I am lucky to have such fantastic parents. My wife, Elizabeth, I started this PhD before I met you but I could not have finished it without you.

I do believe I have had the best supervisory committee of anyone I know. Dr. Beatty, Dr. Murphy, and Dr. Fernandez you have been the exactly perfect level of supportive, critical, and congratulatory. Thank you.

I have been incredibly fortunate to have been in the Smit lab. It has been said that choosing your supervisor is the most important choice you have to make in grad school—I chose well in John Smit. Dr. Smit, your patience with my crazy ideas let me grow as a scientist but your lack of patience with my crazy ideas kept me in line and prevented my PhD from lasting a decade. Dr. Nomellini, I have been fortunate to have you nearby to fix my mistakes, hear me complain about my mistakes, and showing me how to prevent future mistakes. The few other members of our lab that I have shared time with, Lyngrace, Jan, Christina, you have all been great allies in the lab.

Beyond the official credits that are given, I want to acknowledge the incredible collaborators I have had since day one. Dr. Evgeny Vinogradov is the world’s preeminent expert on bacterial polysaccharide analysis, there is quite literally no one that would have been a better collaborator for our lipopolysaccharide work. The Consortium for Functional Glycomics Core H at Emory University were incredibly easy to work with and even though our results with them were ultimately negative, the service they provide is an invaluable asset to the field. Dr. Martine
Caroff has been a surprisingly helpful source of expertise and critique, especially critique.

The Murphy lab in our department of Microbiology and Immunology is the go to collaborators for protein X-ray crystallography, my experiences are evidence to why that is. All of them (Angele, Meghan, Slade, Jason, Catherine, Stephanie, Marek, Michael, Mariko) accepted me like a lab member from the start. Dr. Anson Chan was an exceptional partner for crystallography, his expertise kept us moving forward against a tough project. Especially I would like to thank Dr. Chan for staying up all night many times on our long data collection runs. Staying up late is not so bad, but staying up late for poor results deserves commendation.

Dr. Roland Benz was my only collaborator to initiate a collaboration with us. The porin project started as a side-project but it has become a nice little story that I am happy to have been apart of. Dr. Benz, I hope to actually meet you one day.

Lastly, I would like to acknowledge my department as a whole. This department has been incredibly supportive and warm. With people like Darlene, Sue, and Dr. Gold guiding the ship, I never felt like we were going to crash. A special acknowledgment goes to Dr. Jan Burian, who was always around with encouragement and advice but especially for his organization of many departmental events and teams that brought together a cadre of scientists on the dodgeball court, on the softball diamond, and beyond the confines of campus.
For my Oma, Agnes Bergen...

She pushed everyone in our family to get as much education as possible. She supported me even when she thought a PhD might be a bit too much education.

Your prayers for my sugars and crystals helped me along the way,

I love you Oma!
Chapter 1

Introduction

“But during the writing of this review, I learned how little I knew in this area, and this was a humbling and sobering experience. I am certain that I have made many mistakes due to my ignorance, and I hope that the review will be useful despite its many faults”

— Hiroshi Nikaido, 2003, my grand supervisor.

1.1 S-layers

1.1.1 S-layer structure

Cellular envelopes are the interface between a cell and its environment. Many different organisms have evolved extra coatings and membrane adaptations to enhance or provide new functionality to their cell envelopes. For example, capsules can enhance the virulence of streptococci\(^2\) and mycolic acids increase resistance to antibiotics for mycobacteria\(^3\). One such adaptation widespread in prokaryotic life is the protein surface layer (S-layer). S-layers are proteinaceous, para-crystalline coatings that surround some bacteria and archaea\(^4,5\). A good review on the subject of S-layer structure was written by Tea Pavkov-Keller, Stefan Howorka and


Walter Keller S-layers are two-dimensional sheets, being a thin layer of protein or glycoprotein sitting exterior to the outer membrane for Gram-negative bacteria, the peptidoglycan layer in Gram-positive bacteria, and the cell membrane of most archaea. Figure 1.1 shows the general, cross-sectional structure of cell envelopes of bacteria and archaea that possess an S-layer.

S-layers are composed of one or only a few proteins or glycoproteins. In most known examples, S-layers are composed of one repeated protein. In a few cases S-layers are composed of a few separate proteins, which are stacked one on the other. The S-layer from Clostridium difficile is composed of two proteins (HMW and LMW) that arise from the cleavage of a single precursor protein (SlpA) that is cleaved during secretion. E. Calabi et al. Mol Microbiol, 40: 1187–1199, 2001. Bacillus anthracis has a S-layer that is composed of two proteins (EA1 and Sap) that are individually expressed and assemble together on the surface of the bacterium. S. Mesnage et al. Mol Microbiol, 23: 1147–1155, 1997. Interestingly, Brevibacillus brevis appears to have two S-layers stacked on top of each other; the outer layer having oblique symmetry and the inner layer having hexagonal symmetry.

For S-layers that are composed of a single protein, that protein is repeated thousands of times across the surface of the cell and is arranged in a regular geometric pattern. The geometric patterns can be observed by electron microscopy. The traditional electron microscopy techniques of freeze fracture and negative stain have been the most successful techniques historically to see S-layers and S-layer geometries. The discovery of S-layers and the first image of an S-layer published was produced by electron microscopy. That first S-layer image can be seen in fig. For a historical overview of S-layer discoveries, refer to Sleytr et al., 2014.

The five major geometric symmetries found in S-layers are oblique (P1 and P2), triangular (P3), tetragonal (P4), and hexagonal (P6). Figure 1.3 features simple examples of the five

---

Figure 1.1: Cross-sectional diagrams of the cell envelopes of (A) Gram-negative bacteria, (B) Gram-positive bacteria, and (C) archaea. In all known cases the S-layer sits on the outer surface of the cell. The left side is a cartoon, artistic interpretation. The right side is a simpler, structural diagram. (This diagram was inspired by Fig. 1 from U. B. Sleytr and P. Messner. Crystalline surface layers on bacteria. Annual Reviews in Microbiology, 37: 311–339, 1983.)

major symmetries found in S-layers. Examples of bacteria with oblique S-layers are *Bacillus stearothermophilus* NRS2004/3a\(^{15}\) and *Brevibacillus (Lactobacillus) brevis*\(^{16}\) Examples of bacteria with tetragonal S-layers are *Corynebacterium diphtheriae*\(^{17}\) and *Aeromonas salmonicida*

---


Figure 1.2: The first published image of a S-layer. The hexagonal S-layer on the surface of the bacterium — probably Spirillum sp. — is visible along the edges of the cell body (centre right). The scale bar denotes one micrometre. (This image is Figure 1 from A. Houwink. A macromolecular mono-layer in the cell wall of Spirillum spec. Biochimica et Biophysica Acta, 10: 360–366, 1953. Reused with full permission from the publisher, Elsevier.)
An example of a bacterium with a triangular S-layer is *Sulfolobus acidocaldarius*. Examples of bacteria with hexagonal S-layers are *Bacillus anthracis* and *Caulobacter crescentus CB15*.

Figure 1.3: A simple overview of S-layer symmetries. P1 and P2 are oblique symmetries. P4 is a tetragonal symmetry. P3 is a triangular symmetry. P6 is a hexagonal symmetry. Bacterial examples are given on page 2.

In addition to the traditional transmission electron microscopy techniques that have been used to visualize bacterial cell surfaces in the past, a new method of cryo-electron tomography has been used recently to observe and dissect bacterial cells. Cryo-electron tomography has the advantage that it can be used to observe unfixed and unstained intact cells. Being a tomographical technique, cross sections can be reconstructed to observe the fine structure within and on a bacterial cell. No deliberate study has been taken to use cryo-electron tomography to study S-layers, but their presence is easily observed in the fantastic images produced by other studies utilizing the technique. This technique has been pioneered for use in bacteria and specifically in *Caulobacter crescentus* by the laboratories of Dr. Grant Jensen at the California Institute of Technology and Dr./Sriram Subramaniam at the National Cancer Institute. Figure 1.4 features a tomographical cross-section of a *C. crescentus* cell; visible are the inner and outer membranes and the S-layer composed of the protein RsaA.

---

**Figure 1.4:** The location of the S-layer is around the exterior of the cell, as can be seen in this image. The figure is a 19 nm cross-section through the three-dimensional reconstruction of one representative *C. crescentus* cell: **SL**, surface layer; **OM**, outer membrane; **PG**, peptidoglycan layer; **IM**, inner membrane; **St**, stalk; **Rib**, probable ribosome; **GF**, gold fiducial used to align images; **Phb**, putative poly-β-hydroxybutyrate granule. The scale bar indicates 200 nm. (This figure is a portion of Figure 1 from A. Briegel et al. Multiple large filament bundles observed in *Caulobacter crescentus* by electron cryotomography. *Molecular Microbiology*, **62**: 5–14, 2006. Reused with full permission from the publisher, Wiley. The dotted white lines highlight the location of filaments of the protein Crescentin, the focus of the originating paper.)

### 1.1.2 S-layer proteins

Discussing S-layer proteins as a whole can be difficult because S-layer proteins do not represent a monophyletic group of related molecules; S-layer proteins seem to be unrelated proteins that share a common cellular location and gross structure through convergent evolution. One of the best reviews on S-layer proteins was published in 2000 by Margit Sára and Uwe B. Sleytr.\(^5\)

S-layer proteins do tend to share similar properties and overall composition. S-layer proteins are relatively large proteins. For the S-layer proteins whose sequences are known, all are larger than 400 amino acids (\(\text{AA}s\)) and many of them larger than 1000 \(\text{AA}s\).\(^5\) The amino acids in S-layer proteins are rich in hydrophobic and negatively charged residues, low in methionines, and cysteines are almost completely absent.\(^5\)\(^{22,23}\)

The most well studied S-layer proteins are RsaA from *C. crescentus*, SbsB, and SbsC, from


**Geobacillus stearothermophilus.** RsaA is studied in part for its potential as a biotechnology platform. SbsB and SbsC also have potential for applications but they are well studied because they are easy to work with; they can be expressed in *Escherichia coli* and will readily renature from dissociation in strong chaotrope solutions to form S-layers on almost any flat surface available. The ease of working with SbsB has paid off in the form of the first, and only, atomic-level resolution crystal structure of a bacterial S-layer protein with the symmetry determining domains intact.²⁴ The protein SbsB was crystallized by inhibiting its S-layer forming ability by binding it with a nanobody, a monoclonal single domain camelid antibody. The crystal study reveals that calcium ions play an integral role in the structure of the S-layer protein. These calcium ions ‘trigger’ the folding of the protein. This mirrors the hypothesized role of repeat-in-toxin (RTX) motifs in the *C. crescentus* S-layer protein RsaA discussed in section 1.1.4. For our efforts to crystallize RsaA, see chapter 2.

S-layer proteins are almost universally glycosylated in archaea, commonly glycosylated in Gram-positive bacteria, and very rarely glycosylated in Gram-negative bacteria.⁵²⁵ For some microbes, the S-layer’s main function may be as a platform for oligo/polysaccharide chains. One study by Schäffer et al. asked the question about the glycosylation of S-layer proteins “Are S-layer Glycoproteins and Lipopolysaccharide Related?”²⁶ The study identified structural similarities between the carbohydrates in O-specific polysaccharide (OPS) and the carbohydrates on glycosylated S-layer proteins and pointed out they have more in common with each other than with eukaryotic polysaccharides. The study makes flawed leaps from sugar structures to phylogenetic and taxonomic conclusions, but the study did highlight the concept that supporting a highly glycosylated surface is a widespread feature in microbial life and it is a role that S-layer proteins serve in some organisms.

---

1.1.3 Functions of S-layers

As mentioned before, S-layers are not a single family of related proteins—likewise they do not fill the same needs in all species. Interestingly, despite the apparent convergent evolution in many disparate organisms pushing towards a widespread similarity in structure and composition, S-layers seemingly perform many different functions. Further, S-layers require a large metabolic commitment to produce and maintain. Despite their cost, they are still retained in many species where their functions are enigmatic. Due to their location on the cell surface, S-layers are undoubtedly involved with the cell’s interaction with the environment. While the classification of S-layer-possessing bacteria into pathogenic versus non-pathogenic can be somewhat arbitrary, it is a convenient distinction to compare S-layer functions.

Pathogenic bacteria For pathogenic bacteria, an S-layer is often a virulence factor. In *Bacillus anthracis* the S-layer consists of the protein BslA and it functions as an adhesin. BslA is sufficient and required for *B. anthracis* to adhere to host cells and mutants that lack the S-layer have much reduced virulence and increased time-to-death for animal models. Due to their location on the surface of the cell, S-layers functioning as adhesins is a common theme in pathogens. *A. salmonicida* has a S-layer that also functions as an adhesin. The *A. salmonicida* S-layer binds to both fibronectin and laminin on host cell surfaces but is not involved or required for host cell invasion, which is probably mediated by the lipopolysaccharide (LPS). The *A. salmonicida* S-layer, composed of the protein VapA, is sufficient to confer host cell adhesion to the bacterial cells and also to nylon microbeads *in vitro*.

For *Campylobacter fetus*, its S-layer is central to its success as a pathogen of ungulates. The S-layer of *C. fetus* is involved in disseminating the bacteria throughout the host’s body and evading the host immune system. The manner in which *C. fetus* uses its S-layer to escape

---

the immune system is remarkable. The S-layer provides the cells with an innate resistance to complement; strains with an intact S-layer bound less C3 protein than cells lacking a S-layer. By binding less C3, S-layer positive *C. fetus* cells prevent complement-mediated killing and opsonization by C3b. The S-layer would be a natural target for antibody-mediated adaptive immunity, but *C. fetus* has also evolved a strategy to escape antibody recognition. *C. fetus* possess the ability to undergo antigenic conversion by expressing different versions of its S-layer. The *C. fetus* cells can invert portions of their genomic DNA to express alternate S-layer proteins and cognate OPS anchor molecules. As an example, the *C. fetus* serotype A strain 23D encodes for eight different S-layer proteins and can switch between expressing each by moving a single active promoter in front of the S-layer protein gene to be expressed. These S-layers are not just different antigenically, they can even form different symmetry patterns (see fig. [1.3]).

**Non-pathogenic bacteria**  The roles of S-layers in non-pathogenic bacteria can often be mysterious in lab settings. While virulence can be a clear, quantifiable phenotype, the fitness gain derived from an S-layer in the wild is not often abundantly obvious in a test tube. S-layers have been shown to act as barriers to the outside world. S-layers acting as a physical barrier, protecting the cell, is a function primarily exemplified by *C. crescentus* and other Gram-negative bacteria possessing an S-layer. Susan Koval of the University of Western Ontario has done extensive research into the biology and life cycle of *Bdellovibrio* spp. and *Bdellovibrio*-like species, predatory bacteria that prey on Gram-negative bacteria such as *C. crescentus*. Koval showed that the S-layer of *C. crescentus* had a protective effect against *Bdellovibrio* predation. This protective effect was seen for *A. salmonicida*, *Aquaspirillum serpens*, *Aquaspirillum sinuosum*, *C. fetus*, and *C. crescentus*. On the other hand, S-layers do not confer any protection from predation by protozoans (e.g. *Tetrahymena thermophila* and *Paraphysomonas vestita*). The barrier capacity of S-layers extends to blocking abiotic threats, as was demonstrated by De la

Fuente-Núñez et al. in *C. crescentus* in 2012. *C. crescentus* cells that possessed an S-layer were found to have protection against the effects of antimicrobial peptides compared to cells that lack an S-layer.\textsuperscript{33} The authors note that the peptides were smaller than the pores in *C. crescentus* S-layer (see fig. \textsuperscript{[1,5]} and yet the protective effect existed. The S-layer may act as a sponge, not just a physical barrier, adsorbing harmful molecules before they can reach the cell. The S-layer of *C. crescentus* is not glycosylated but many other species’ S-layers are glycosylated and those carbohydrates would add another layer of coating to the cell to ward off intrusion.

A protein coating will impart to a cell surface with different properties than a ‘naked’ cell; these differences may be a conserved function of an S-layer. It has been noted that many bacteria with an S-layer are more hydrophobic than isogenic variants lacking an S-layer.\textsuperscript{34} A noted exception to the hydrophobicity observations was for *C. crescentus* strains CB2A and CB2NY66R; its S-layer covers beneath it a more hydrophobic outer membrane (see chapter \textsuperscript{3} for our research on the *C. crescentus* LPS). Beyond changing the general chemical properties of the cell surface, S-layers also change the cell surface to provide a specific site for extracellular proteins to anchor. Enzymes that are secreted, but remain associated with the cell, can use S-layers as display platforms. The enzyme pullulanase in *Clostridium thermosulfurogenes* anchors itself into the S-layer with three C-terminal regions that have a significant sequence homology to the *C. thermosulfurogenes* S-layer protein.\textsuperscript{35} It is thought that the *C. thermosulfurogenes* pullulanase integrates itself in the lattice of the S-layer in order to anchor itself. Integration into the S-layer is not the strategy used by *G. stearothermophilus* DSM2358 and its amylases. The high molecular weight amylase from *G. stearothermophilus* binds the cell surface but without any detectable interaction with peptidoglycan, instead it associates with S-layer in a manner that is specifically not lattice integration.\textsuperscript{36}

1.1.4 RsaA

RsaA, named for regular surface array,\textsuperscript{27} is a 1026 AA protein from \textit{C. crescentus} with an average molecular weight (\textit{MW}) of 98001.70 Da and an isoelectric point (\textit{pI}) of 3.85. The amino acid sequence of RsaA can be found in the appendix on page\textsuperscript{154}. The S-layer of \textit{C. crescentus} comprises roughly 45000 copies\textsuperscript{38} of RsaA arranged in an \textit{p6} symmetry (see fig. 1.3). RsaA is a type 1 secretion system (T\textit{1SS}) secreted protein containing five to six RTX motifs.

Figure 1.5 displays a computer reconstructed view of the surface of the S-layer of \textit{C. crescentus} as determined by tilt-tomography electron microscopy.\textsuperscript{39} This figure is the highest resolution image available for the structure of RsaA and the S-layer, a resolution of about 20 Å. In a very elegant series of experiments in 2010, Amat et al. were able to determine the configuration of the RsaA monomers in S-layer structure\textsuperscript{38} Different constructs of RsaA were produced that contained single cysteine residues at different locations throughout the protein. Labelling those cysteines with nano-gold particles \textit{in situ} in the S-layer, led to noticeable changes in electron density in the hexagonal unit cell of the S-layer. These densities corresponded to the physical locations of the engineered cysteine residues. This work proved that the N-termini of the RsaA monomers are located in the axis of three-fold symmetry, while the C-termini are centered around the axis of the six-fold symmetry (refer to fig. 1.5)

\textit{Type 1 secretion} RsaA is secreted by a T\textit{1SS}. T\textit{1SS} secreted proteins are secreted in one step across both the inner and outer membranes of Gram-negative bacteria.\textsuperscript{40} A T\textit{1SS} requires three components: an inner membrane spanning ATP binding cassette (ABC) transporter, a periplasm spanning membrane fusion protein, and an outer membrane protein. For RsaA’s secretion from \textit{C. crescentus}, these proteins are RsaD, RsaE, and RsaF (either RsaFa or RsaFb)\textsuperscript{41,42}

\begin{itemize}
\item \textsuperscript{38} F. Amat et al. \textit{J Bacteriol}, \textbf{192}: 5855–5865, 2010.
\end{itemize}
**Figure 1.5:** The surface structure of the S-layer from *C. crescentus*, reconstructed from tilt-series electron microscopy. **A.** A section of reconstructed surface of the *C. crescentus* S-layer. The image shown is in a left-handed configuration, except the handedness of the S-layer could not be determined in the original study, therefore this image may be a mirror image of the true configuration of the S-layer. **B.** Interpretations of the micrograph. A monomer of RsaA is highlighted in green (*B.1*). The unit cell of the S-layer is six copies of RsaA arranged in a hexagonal pattern (*B.2*). The center of three-fold symmetry, red, is comprised of the N-termini of RsaA (*B.3*). The C-termini of the RsaA monomers face towards each other in the center of the point of six-fold symmetry, blue (*B.4*). Possible pores through the S-layer are highlighted in orange (*B.5*). (This image is derived from Figure 6 from J. Smit et al. The S-layer of *Caulobacter crescentus*: three-dimensional image reconstruction and structure analysis by electron microscopy. *Journal of Bacteriology*, **174**: 6527, 1992.)

All T1SS secreted proteins, RsaA is secreted starting from the C-terminus to the N-terminus. As proteins are translated from the N-terminus to the C-terminus, all T1SS secreted proteins (including RsaA) exist in the cytoplasm as fully translated polypeptides prior to secretion. This is in contrast to protein secreted by N-terminus first means, which allow for the possibility of secretion to begin on a nascent peptide before translation is complete. The diameter of T1SS pore is not wide enough to accommodate the secretion of a fully folded protein, so T1SS secreted proteins often have a C-terminus that is intrinsically disordered prior to secretion; this disordered state is mediated by RTX motifs.43

Repeat-in-toxin motifs  First identified in the *E. coli* protein α-haemolysin (HlyA),\(^{44}\) RTX motifs are a nearly ubiquitous component of T1SS secreted proteins. Despite the name, repeat-in-toxin are not exclusively found in toxins, they are found in many classes of protein, lipases, proteases, S-layers, etc. RTX motifs are small peptide sequences, typically 6–9 AA in length, the amino acid sequence is GGXGXD\(^{45}\) or GGXGDXDLX (where G is glycine, D is aspartate, X can be any amino acid, and L is leucine but is often substituted by valine, isoleucine, phenylalanine, or tyrosine).\(^{43,46}\) The short sequences are always present in multiples (6–50+), often found in tandem repeats. Each RTX motif forms two half sites for binding Ca\(^{2+}\), which can collaborate with other neighbouring RTX sites to coordinate ions and stabilize a connection between disparate regions of a protein. Upon binding Ca\(^{2+}\), RTX motifs assume a parallel β-roll fold,\(^ {47}\) rolling around Ca\(^{2+}\) ions at the turns. An example of a RTX β-roll structure can be found in fig. 1.6.

RTX motifs are always on the C-terminal side of T1SS proteins, and are thus are in the leading part of the polypeptide to be loaded into the secretion machinery and the first part to be secreted. It should be noted that bacterial intracellular Ca\(^{2+}\) concentrations are extremely low, 90±10 nM, and extracellular Ca\(^{2+}\) concentrations are higher, 0.1–1 mM.\(^ {49,50}\) It is thought that the roles of RTX motifs in T1SS proteins are to induce an intrinsically disordered C-terminus prior to secretion and to quickly bind Ca\(^{2+}\) and begin folding once the protein has started to emerge from the secretion machinery.\(^ {51,52}\) The calcium-induced folding may help the thermodynamics of secretion, possibly assisting by pulling the protein out of the pore. Despite the apparent importance of RTX motifs in T1SS proteins, they are not sufficient nor even necessary for secretion\(^ {44}\) but in at least the case of HlyA, in *E. coli*, the RTX repeats are required for proper folding, activity, and maximal secretion.\(^ {53}\) It has been thought that the folding of the RTX motifs can act as an ‘intramolecular’ chaperone; they fold quickly first then act as a nucleation point

---

Figure 1.6: The structure of RTX motifs. This view is down the barrel of an RTX β-fold. The blue spheres are calcium ions coordinated by aspartate residues which are shown. This structure is of MIS38, an extracellular lipase from *P. aeruginosa* (PDB:2Z8X). The structure here depicts five RTX sites binding five Ca$^{2+}$ ions, RsaA from *C. crescentus* also has five RTX motifs in close vicinity and may have a similar structure to the presented structure above. See the sequence of RsaA on page 154. RTX sequences are bolded. Generated using the program Chimera (E. F. Pettersen et al. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, **25**: 1605–1612, 2004.).

for the rest of the protein to continue to fold on.

### 1.1.5 Biotechnology applications of S-layers

S-layers’ ability to self-assemble into large, complex structures makes them an attractive platform for biotechnological applications. Their density and regular arrangement over relatively large distances is unique in biology. There have been several groups pursuing technologies that utilize S-layers assembled off of the surface of the cell and groups that focus on S-layer technologies on the cell surface. Many different technologies have been developed. One of the significant defining differences among all S-layer applications is the location where the S-layer is assembled. The potential sites for assembly are:

- In suspension
- On a solid surface
• On a liquid interface of lipid membrane

• On a lipid nanoparticle

• On a cell surface

Very little interest or work has been done on developing applications for S-layers that assemble and exist in liquid suspension. S-layers are inherently two-dimensional and unrestricted freedom in liquid suspension does not highlight their unique structure. Assembling S-layers in solution is most often a means to getting the S-layers onto a solid surface. Once an assembled S-layer is on a solid surface it is a potent structure for biotechnology development. The S-layer itself can act as a filter or the S-layer can be a platform for functionalized molecules and peptides.

Using S-layers as ultrafilters has been a tempting idea since not long after they were discovered. As mentioned in section 1.1.3, one of the possible natural functions of S-layers is to act as a physical barrier against large, intrusive foreign entities. Unlike conventional ultrafilters made of inorganic polymers, S-layers are isoporous with very stringent molecular weight cutoffs. In its most basic inception, an S-layer ultrafilter could be constructed by assembling an S-layer on the inner surface of a traditional ultrafiltration membrane. These S-layer ultrafiltration membranes (sometimes called SUMS) do have exquisite selectivity of filtering and when they are crosslinked with glutaraldehyde, they are highly resistant to solvents, acids, and bases. These S-layer filters were covered by at least three US patents, the last of which expired in 2011. The apparent lack of enthusiasm in the field of ultrafiltration using S-layers in the last few years suggests that the early results were not easily translatable into a large-scale commercial product.

Having the S-layer function as a platform for functional molecules and peptides is the most popular approach to using S-layers. The most uncomplicated way to introduce a functional structure into an S-layer is to genetically clone it into the S-layer gene itself. Molecules of interest can always be chemically linked to an intact S-layer, but the opportunity for S-layer to

inherently display peptides of interest dominates the field. The best few examples of molecules of interest being attached to an S-layer after it is assembled are cases where coupling domains were first cloned into the S-layer protein. One such example is streptavidin cloned into SbsB (an S-layer protein from \textit{G. stearothermophilus}) so that biotinylated molecules could be loaded on to it.\textsuperscript{55} Another example is from the S-layer of \textit{C. crescentus}, where protein G has been expressed so that thousands of copies of any IgG antibody can be loaded across its surface.\textsuperscript{56}

As mentioned above, genetically modifying S-layer genes so that the proteins contain valuable peptides and proteins is the showpiece technology for S-layers. These exogenous peptides can be displayed at a high density across a relatively large distance (\textit{e.g.} the size of a cell or larger) by being displayed in an S-layer. This approach has been successfully pursued in \textit{C. crescentus}' S-layer. The \textit{C. crescentus} S-layer has displayed vaccines,\textsuperscript{57} enzymes,\textsuperscript{58} HIV-binding proteins,\textsuperscript{59,60} and peptides for biosorption of heavy metals.\textsuperscript{61}

All S-layers are biological entities that are ultimately tethered to a cell membrane, but if that arrangement is flipped, a lipid membrane can be formed and supported by an S-layer.\textsuperscript{62} Two-dimensional lipid-membrane structures can also be turned into three-dimensional S-layer-coated nanoparticles.\textsuperscript{63} The two-dimensional structures can be assembled across the surface of a Langmuir-Blodgett apparatus. The majority of the interest in these structures is in the lipid layer;\textsuperscript{64} the S-layer only acts as a stabilizer. On the other hand, lipid nanoparticles coated in S-layers garner a great deal of excitement for the S-layer itself. The promise of lipid nanoparticles (\textit{i.e.} vesicles, liposomes and solid-cored particles) is that all the above technology already developed or being currently developed can be translated to an abiotic particle that highlights the strengths of S-layers. Some S-layers will readily assemble on nearly any surface, those S-

layers will coat standard lipid vesicles that are present, while other S-layers require specialized lipids to assemble upon a nanoparticle. For instance, \textit{C. crescentus} S-layer can be reassembled on to lipid vesicles containing \textit{C. crescentus} LPS.

Two good reviews on the subject of S-layer applications were written by Smit (2008) and Ilk, Egelseer, and Sleytr (2011). The long-range uniformity of S-layers whet the interest for the structures that can be built on top of them. In this vein, using S-layers to support mineralization, nanoparticle arrangement, and lipid membranes all share in common using S-layers as physical structures. Many S-layers can self-assemble into long-range regular structures \textit{in vitro}, unattached to a cell surface. These \textit{in vitro} S-layers” are a hotbed for biotechnology development because they provide a uniquely dense and patterned protein surface that can be divorced from contaminants of whole cells.

1.2 Lipopolysaccharide

The most abundant non-protein component of the Gram-negative outer membrane is LPS. The LPS of \textit{E. coli} and \textit{Salmonella enterica} var. Typhimurium are regarded as the “canonical” examples. Those LPS are the molecules that all others are generally compared to and this thesis will do the same. An excellent review of LPS was written by Raetz and Whitfield (2002).

LPS resides in the outer leaflet of the Gram-negative outer membrane. In physiological conditions, in wild-type coliform bacteria LPS is the only lipid component in the outer leafle.

LPS is synthesized in the inner membrane. While in wild type cells there exists detectable levels of LPS in the inner membrane, the vast majority of LPS resides in the outer membrane.

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{65} M. H. Ucisik et al. \textit{Colloid Surface B}, \textbf{128}: 132–139, 2015.
\item \textsuperscript{72} A. Polissi and C. Georgopoulos. \textit{Mol Microb}, \textbf{20}: 1221–1233, 1996.
\item \textsuperscript{73} D. Ding and S Kaplan. \textit{Prep Biochem}, \textbf{6}: 61–79, 1976.
\end{itemize}
\end{footnotesize}
Fully formed LPS is transported across the periplasm to the outer membrane via the proteins LptABCDE\textsuperscript{74,75}. The exact details of how a large LPS molecule can traverse the periplasm and be inserted into the outer leaflet are still uncertain; the difficulty is in no small part due to the fact that all the required proteins are essential for viability.

Figure [1.7] highlights a simplified LPS structure from \textit{Salmonella minnesota} as reported by Vukajlovich, Hoffman, and Morrison (1987)\textsuperscript{76}. There are three fundamental components of LPS: lipid A, the core oligosaccharide, and the O-polysaccharide.

### 1.2.1 Lipid A

**Structure** Lipid A is the ‘base’ of the LPS molecule. The backbone of canonical LPS is two glucosamine sugars bound to each other by an $\beta 1\rightarrow 6$ bond. The diglucosamine molecule has two O-linked beta hydroxy acyl chains at the C-3 and C-3’ positions and two N-linked beta hydroxy acyl chains at the C-2 and C-2’ positions. The two ‘prime’ acyl chains are modified with subsequent acyl chains attached at the beta hydroxy groups. All these acyl chains are fully saturated.

The C-1 and C-4’ positions of the diglucosamine are phosphorylated. These phosphates give the lipid A a net negative charge and they are thought to bind divalent cations which form electrostatic bridges with neighbouring lipid A molecules and stabilize the outer membrane\textsuperscript{77,78}. The C-6’ position is ligated to the core oligosaccharide. The canonical structure has the C-6’ position occupied with an $\alpha 2\rightarrow 6$ bond to a Kdo sugar. Figure [1.8] shows this canonical structure as it is found in \textit{E. coli} K12.

There are numerous Lipid A structures that differ from the canonical structure. Examples exist of modified structures at all the possible locations: the phosphates, the glucosamines, and

Figure 1.7: The wild type LPS from *Salmonella minnesota*. The lipid A is at the bottom of the figure (in blue). It comprises a backbone of two glucosamine residues. Each glucosamine has an attached phosphate moiety and three fully saturated acyl chains. Above the lipid A is the core oligosaccharide (in green). This core oligosaccharide is a hendecasaccharide. Extending from the core oligosaccharide is the O-polysaccharide (in purple). The structure of this O-polysaccharide is a variable number of pentasaccharide subunits (5–40). Abbreviations: Glc, glucose; GlcN, glucosamine; GlcNAc, N-acetylglucosamine; Gal, galactose; GalNAc, N-acetylgalactosamine; Kdo, 3-deoxy-D-manno-octulosonic acid; Hep, L-glycero-D-mannoheptose; P, phosphate; and EtN, ethanolamine.
Figure 1.8: Kdo₂-Lipid A from E. coli K12. The lipid A portion comprises the phosphates (green), the glucosamines (blue), and the acyl chains (black). The two Kdo sugars present are the start of core oligosaccharide.

the lipids. An excellent review on lipid A modifications is by Raetz et al. (2007).  

The phosphate modifications are often extra chemical groups linked directly to the phosphates. Cationic attachments are common, such as ethanolamine (e.g. Salmonella minnesota, see fig. 1.7) and glucosamine (e.g. Bordetella pertussis). Neutral attachments to the phosphate are also possible such as methyl groups (e.g. Leptospira interrogans). The phosphates can also be absent; examples of lipid A molecules missing a phosphate are Rhizobium etli, Bacteroides thetaiotaomicron, Prevotella intermedia, Helicobacter pylori, and Francisella tularensis. Removed phosphates can be replaced with ethanolamine (e.g. H. pylori) or anionic sugars such as galactosamine (e.g. C. crescentus and R. etli).

Modifications to the diglucosamine backbone are less common than other lipid A modifications. The most frequent alteration seen in the backbone is the replacement of the either one or both glucosamines for a D-2,3-diamino-2,3-dideoxy-glucose. Di-D-2,3-diamino-2,3-dideoxy-glucose backbones are seen in *C. crescentus*. One of the most extreme modifications to the diglucosamine backbone is the unique example of *Rhizobium leguminosarum*. *R. leguminosarum*’s lipid A contains no phosphate, instead a galacturonate sugar is present at the 4’ position; and a 28 carbon long acyl chain. The extra long acyl chain is attached to the C-5 position which is only possible because the reducing glucosamine’s ring has been opened up, converting the glucosamine into a linear 2-amino-2-deoxygluconate. Another uniquely modified lipid A is found in *Bdellovibrio bacteriovorus*. The *B. bacteriovorus* lipid A’s backbone consists is 2,3-diamino-2,3-dideoxy-D-glucose disaccharide and contains no charges; the canonical phosphates are replaced with D-mannose moieties.

Lipid modifications come in the form of the number of acyl chains, the length of the acyl chains, and the inner structure of the acyl chains. Six acyl chains is the standard found in canonical lipid A molecules in enteric bacteria, but the number of fatty acid groups can be as high as eight (*e.g.* Vibrio cholerae O139) and as low as four (*e.g.* *H. pylori*). The lengths of the acyl chains in lipid A do not vary as much as in other lipids due to more stringent acyltransferases, but they do vary among bacteria. An extreme example is the 28 carbon long acyl chain mentioned above in *R. leguminosarum*. The acyl chains in lipid A are usually fully saturated but unsaturated examples are not uncommon. Bacteria with unsaturations in their lipid A include *C. crescentus*, *Yersinia pestis*, and *Rhodobacter capsulatus*.

The canonical lipid A biosynthesis pathway has been fully determined. The best sources for information about lipid A synthesis are reviewed by the field’s most prominent investigator,
Effects of LPS on the immune system  LPS is a potent immune stimulatory molecule and it is the lipid A that is the primary structure responsible for the innate immune response. It was this inflammatory activity of LPS that was identified before the molecule was isolated, giving LPS its alternate, original name “endotoxin”.

Lipid A is a microbe-associated molecular pattern (MAMP), a highly conserved structure that is specifically recognized by the vertebrate innate immune system. There are a number of excellent papers and reviews on the immunological aspects of lipid A. The review by Alexander and Rietschel (2001) is a good general review of the innate immunity against LPS. The LPS-toll-like receptor 4 (TLR4)-MD2 complex has been solved by Park et al. (2009). The review by Lu, Yeh, and Ohashi (2008) covers the signal transduction pathways, downstream of LPS/TLR4 recognition.

In the immune system, two proteins, LPS-binding protein and CD14, are responsible for the initial binding of LPS and delivery of it to the recognition machinery. That recognition machinery, once it binds LPS, dimerizes and is comprised of two copies of TLR4 and two copies of the protein MD2. The LPS is held in the cup of MD2 and two of these complexes are sandwiched between two TLR4s. All of the canonical lipid A acyl chains a buried within MD2 except for the acyl chain attached at the C-2 position that is positioned in the hydrophobic interface between MD2 and TLR4. This ‘outside’ acyl group aids in the dimerization of TLR4 and is likely the reason that lipid A molecules with fewer acyl chains are less immunogenic. The phosphates present on the lipid A interact with positively charged residues on both TLR4s and MD2. The presence of phosphates increases the immunogenicity of lipid A about a 100 fold.

---

The crystal structure of the LPS-TLR4-MD2 complex revealed space around the phosphates that may accommodate the presence of bulkier anionic substitutions.\(^9\)

Once LPS is bound and the TLR4-MD2 receptor dimerizes, a signal transduction pathway is activated. The two signaling pathways are the MyD88-dependent pathway and the TRIF pathway (MyD88-independent). Briefly, the MyD88 pathway leads to the expression of proinflammatory cytokines and activation of acquired immunity. The TRIF pathway leads to the expression of type 1 interferons and the activation of the innate immune system.\(^9\) These two pathways’ signals do overlap at a few places, specifically in the activation of the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB). Modulating the immune response to LPS is an avenue for vaccine/adjuvant development. The two pathways can be differentially activated depending on the structure of the bound lipid A molecules.\(^8\) See section \[3.2.16\] for our investigation into the immunogenicity of LPS from \textit{C. crescentus}.

1.2.2 Core oligosaccharide

Structure On the subject of core oligosaccharide structures there are excellent reviews by Frirdrich and Whitfield (2005)\(^{101}\) and Holst (1999)\(^{102}\). Core oligosaccharides, relative to lipid A, are highly variable, even within a species. As an example, there are five known core oligosaccharides in \textit{E. coli} \(^{103}\). Yet compared to OPSs, core oligosaccharides are highly conserved; there are 186 known OPS structures in \textit{E. coli} \(^{104}\). Figure 1.7 shows the LPS from \textit{S. minnesota}; its core oligosaccharide has many “canonical” features (e.g. Kdos, heptoses, hexoses, and phosphates) that are used as comparators for other core oligosaccharides.

The only universally conserved feature of core oligosaccharides is the eight carbon sugar that sits at the base of the structure, Kdo. Kdo is always found as the sugar bonded to the C-6\(^{\prime}\) position of the lipid A by a \(\alpha 2\rightarrow 6\) bond (see fig. 1.8). The only known exception to the Kdo

rule is *Acinetobacter calcoaceticus*, where the Kdo is replaced with a 2-manno-octulosonic acid, Ko. This Kdo-lipid A bond is a ketosidic bond, usually the only one present in an LPS, and it is easily cleaved by heating in the presence of a mild acid. This cleavage makes the separation of the poly/oligosaccharides and the lipid A a simple procedure.

There are often multiple Kdo sugars present in the core, as seen in fig. 1.7. These adjunct Kdo sugars are bonded to the central Kdo by an α2→4 bond. That C-4 position on the central Kdo is often bonded to an anionic moiety, such as Kdo, Ko (*Y. pestis*), phosphate (*V. cholerae*), or galacturonic acid (*Rhizobium trifolii*). The C-5 position on the central Kdo is usually bonded to a mannO-configured sugar; the canonical and common example is LD-heptose. The mannO-sugar is usually the first sugar in an elongated oligosaccharide. Three heptoses are attached to that C-5 position in enteric bacterial core oligosaccharides, these heptoses along with the Kdos represent the ‘inner’ core of enteric cores. The ‘inner’ and ‘outer’ core delineations do not apply well to non-canonical LPSs. Figure 1.9 shows these general core oligosaccharide characteristics. Many core oligosaccharides are phosphorylated at one or many locations. Presumably, the phosphates and other negative charges fulfill similar functions as the canonical lipid A phosphates. LPSs that terminate at the core oligosaccharide, or before, are termed “rough,” while those that possess an OPS are called “smooth.”

![Figure 1.9](image-url)

**Figure 1.9:** This is a general formula for core oligosaccharide structures. R¹ is usually a mannO sugar, specifically it is most often a heptose. R², when present, is substituted with a negatively charged moiety, examples are Kdo and phosphate. (This figure is a recreation of fig 2 from O. Holst. Chemical structure of the core region of lipopolysaccharides. *Endotoxin in Health and Disease*, 1: 115–154, 1999.)

---

Synthesis The synthesis of the core oligosaccharide happens in the inner leaflet of the inner membrane. The structure is built on to lipid A, one sugar at a time.\textsuperscript{101} It is worth noting that the transfer of the central Kdo is the first step for the core oligosaccharide and a required step for the complete synthesis of lipid A.\textsuperscript{109} The completed core-lipid A is then flipped across the inner membrane by the flippase MsbA.\textsuperscript{69}

1.2.3 O-polysaccharide

Structure OPSs are highly variable in structure between species and even strains. For an examples of this variability, an exhaustive review of the OPSSs in \textit{E. coli} has been assembled by Stenutz, Weintraub, and Widmalm (2006).\textsuperscript{104} There are 186 known OPSSs in \textit{E. coli} alone.

OPSSs are synthesized with a repeating structure, formed from subunits. For \textit{E. coli}, the OPSSs are built of subunits that range in size from 2–7 monosaccharides (3 examples of 2 sugar subunits, 1 example of a 7 sugar subunit). Even smaller configurations exist, such as that are homopolymer of a single repeating sugar in a single configuration (\textit{e.g.} \textit{V. cholerae}).\textsuperscript{110} There are also OPSSs in \textit{E. coli} that are comprise a single repeating sugar (O8, O101) but they have multiple configurations or intersaccharide bonds. These ‘simpler’ OPSSs are often characteristic of an OPSS assembled by the ABC transporter-dependent pathway (see below).

OPSSs can be synthesized from a large range of sugars. Well known sugars, such as glucose and mannose are regular components of OPSSs but so are rare sugars such as talose, colitose, and pseudaminic acid.\textsuperscript{104} See chapter 3 for our determination of the structure of the OPSS from \textit{C. crescentus} which contains the rare sugar N-acetylperosamine.

Synthesis Unlike the core oligosaccharide, the OPSS is produced on a lipid carrier in the inner membrane and then ligated onto the nascent LPS as a fully completed polymer. There are

three known biosynthetic pathways that produce OPS: the Wzy-dependent pathway, the ABC transporter-dependent pathway, and the synthase-dependent pathway.

**Figure 1.10:** A simple cartoon overview of the Wzy-dependent pathway of OPS synthesis. Wzx is the ‘O-antigen/OPS subunit flippase.’ Wzy is the ‘O-antigen polymerase.’ Wzz is the ‘O-antigen length determining protein.’ WaaL is the ‘O-antigen ligase.’ MsbA is the ‘LPS ABC flippase.’

The Wzy-dependent pathway is the classic subunit based synthesis pathway for OPSS. The subunits, consisting of 2–7 sugars, are assembled on a undecaprenyl-pyrophosphate carrier lipid in the inner leaflet of the inner membrane. The complete subunits are flipped across the inner membrane by the protein Wzx. The subunits are then polymerized by the enzyme Wzy in the outer leaflet of the inner membrane. The length of the growing OPS polymer can vary but it is influenced by the chain length regulator protein Wzz. When the OPS has reached a suitable length, the enzyme WaaL transfers the entire OPS from the carrier lipid and ligates it on to the core oligosaccharide, completing LPS synthesis. Figure 1.10 is an overview of this pathway. LPSs with OPSS made via the Wzy-dependent pathway have a characteristic ladder pattern on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels. The ladder pattern comes from the variable chain length introduced during subunit polymerization. The separation between the bands is equivalent to the size of one OPS subunit. Figure 3.3 has an example of that Wzy-dependent pathway ladder pattern.

The defining feature of the ABC transporter-dependent pathway is that it assembles OPSS
one sugar at a time by glycosyltransferase enzymes. Because the polymer grows one sugar at a time there is no specific polymerase enzyme (like Wzy) and the chains are non-branched. The OPS polymer is fully assembled in the inner leaflet of the inner membrane before it is transported across the inner membrane by a ABC transporter. Once the OPS is in the periplasm it is ligated on to the core via WaaL. A unique feature of ABC transporter-dependent pathway synthesized OPSs is the addition of a unique end cap on the non-reducing end.\textsuperscript{111}

\textbf{Figure 1.11:} A simple cartoon overview of LptABCDE mediated transport of \textit{LPS} across the periplasm. LptA is depicted here as shuttling the \textit{LPS} across the periplasm but it is possible that LptA forms filaments across the periplasm.

The synthase-dependent pathway is a rare pathway only found in the plasmid-encoded OPS:54 of \textit{Salmonella enterica} serovar Borreze.\textsuperscript{112} The synthase-dependent pathway involves a single protein that processively assembles the OPS in the inner membrane.

Once the \textit{LPS} is fully formed in the outer leaflet of the inner membrane it is transported across the periplasm to the outer membrane by the LptABCDE proteins. Ruiz, Kahn, and Silhavy (2009) covers the subject of \textit{LPS} transport.\textsuperscript{113} Figure\textsuperscript{1.11} is an overview of \textit{LPS} transport

across the periplasm.

1.3 Porins

The outer membrane of Gram-negative bacteria is a particularly impermeable membrane to lipophilic and hydrophilic molecules. In 1976 a protein complex was first isolated from the outer membrane of *S. enterica* that produced transmembrane channels which non-specifically transported solutes; this complex was the first to be called a ‘porin’ \(^{114}\). Today, the term porin refers to outer membrane proteins that transport molecules across a membrane passively. This definition is for non-specific protein channels \(^{115}\), such as OmpF from *E. coli*, but sometimes the definition is widened to include structure-specific transport channels, such as LamB in *E. coli* \(^{116,117}\).

Porins act as sieves in the outer membrane allowing small solutes to pass from the environment to the periplasm. Without a channel for them to pass, the outer membrane would be impermeable to these usually hydrophilic nutrient molecules. In their simplest form, porins can be thought of as water filled holes in the membrane, allowing anything that will fit through. In reality, these holes are complex structures with changing shapes and charges. The activity of these holes is defined by their size and selectivity.

The size of a pore can be an absolute measure of the pore diameter, as measured by X-ray crystallography, or as a derived figure based on permeability studies. The selectivity is judged on the relative diffusion rates of different molecules and how they are discriminated by size, charge, and hydrophobicity. An excellent example of these properties, as seen in *E. coli* porins, is assembled in a review by Nikaido \(^{115}\) In *E. coli* the porins OmpF and OmpC both have a preference for cationic molecules over anionic molecules, but both will allow the passage of uncharged molecules like sugars. On the other hand, the *E. coli* porin PhoE strongly prefers

---


Figure 1.12: A simple cartoon overview of how TonB-mediated active transport differs from passive transport through a porin. On the left is a simplified TonB-mediated uptake system consisting of a TonB-dependent outer membrane protein (OMP), and the TonB inner membrane complex (TonB and ExbBD). Energy consumed at the inner membrane is transduced to the outer membrane through the action of TonB. Not shown is the globular domain of the TonB-dependent outer membrane protein called the ‘plug.’ On the right is a porin, passively allowing solutes to pass through it.

...anions over cations. There are theoretical models that suggest that ion selectivity is related to ionic concentration, the lower the concentration the greater the selectivity.\textsuperscript{118} One important difference between OmpF and OmpC is the size; OmpF has a larger pore diameter, allowing more solutes to be transported and a broader range of solutes. This discrimination in size is an important aspect of porin function. Porins give outer membranes a characteristic molecular weight cutoff of permeability, also known as an exclusion limit.\textsuperscript{119} This exclusion limit can be analyzed by measuring the rate of diffusion of differing sized carbohydrates through the pores in a liposome swelling assay.\textsuperscript{119} OmpF and OmpC have severe reductions in permeability for solutes above a size around 200 Da.

Porins increase the communication between the periplasm and the outside world allowing selective permeability take place at the cytoplasmic membrane by proteins that have access.

to cellular energetics. There does exist active transport systems in the outer membrane, such as TonB-dependent transporters. These active transport systems are often associated with the uptake of scarce nutrients like iron and vitamin B<sub>12</sub>. The alphaproteobacterium *C. crescentus* encodes for 67 such TonB-dependent outer membrane proteins, possibly as an adaptation to living in nutrient poor environments. TonB-mediated transport systems utilize energy consumed at the inner membrane to power specific uptake systems in the outer membrane. Figure 1.12 illustrates TonB-mediated transport systems and a porin. To date no published study has characterized a general porin in *C. crescentus*. See chapter 4 for our work to characterize OmpW in *C. crescentus*.

### 1.4 Aims and Objectives

The bacterium *C. crescentus* is a well studied model organism for its dimorphic life cycle, cell shape, and regulatory pathways. Despite having a unique S-layer, an unsolved LPS structure, and no known porins in its outer membrane, the cellular envelope of *C. crescentus* has not attracted widespread attention. This thesis covers my work to illuminate the composition and the structures of the three different components of the *C. crescentus* cell envelope: the S-layer protein, RsaA (chapter 2); the LPS (chapter 3); and the porin, OmpW (chapter 4).

---

Chapter 2

Crystallization and X-ray diffraction of the S-layer protein RsaA

“Please forgive me for presenting, on such a great occasion, results which are still in the making, but the glaring sunlight of certain knowledge is dull and one feels most exhilarated by the twilight and expectancy of the dawn.”
— Max Perutz, 1962 Nobel lecture
The father of protein X-ray crystallography.

2.1 Introduction

Protein surface layers (S-layer) are para-crystalline coatings that surround microbial cells\[34\]. They have various functions, such as adhesion and immune evasion in pathogenic bacteria;\[27,125\] protection against phage and predation;\[31,32\] and as a pseudo-membrane in archaea.\[126\] S-layers are composed of either one or a small number of secreted proteins that self assemble into a contiguous sheet in the extracellular environment.

Despite their unique features, S-layers have not received the intensive structural investigations that most other protein families have received, in no small part due to their resistance to 3D-crystallization. The reason why S-layers resist crystallization is not completely un-

derstood, but it is definitely related to their proclivity to self-assemble into large polymeric two-dimensional structures \((i.e. \) S-layers). That S-layer formation plays a role in the crystallization problems seemed confirmed by the fact that the only groups to successfully crystallize an S-layer protein have done so by specifically inhibiting or removing the S-layer formation ability of the proteins\(^{24,127}\). It is thought that perhaps the 2D-S-layer formation out competes the 3D-crystallization or maybe the 2D-S-layer structure is incompatible with 3D-crystal growth. Possibly there are conditions that exist that would allow for the crystallization of an intact S-layer protein, but those conditions are currently unknown.

The Gram-negative bacterium \textit{Caulobacter crescentus} possesses an S-layer that is composed of thousands of copies of a single 98.6 kDa protein, RsaA\(^{12}\). This particular S-layer has been pursued as a potent platform for peptide display and protein expression\(^{67}\). RsaA is secreted from the cell via a type 1 secretion system\(^{128}\).

RsaA shares little to no sequence similarity to any other proteins, except for in the conserved repeat-in-toxin (RTX) motifs canonically found in all type 1 secreted proteins\(^{43,129}\). When the entire RsaA protein sequence is compared against other proteins with a basic local alignment search (BLAST) analysis the highest hits returned are unknown proteins from \textit{Skermanella aero-}
lata (Score: 1063, Identity: 33.7\%) and \textit{Pseudomonas syringae} (Score: 958, Identity: 31.5\%). Interestingly, the protein BLAST also returns hits to the S-layer protein from \textit{Campylobacter fetus}, SapA (Score: 598, Identity: 29.0\%). Closer investigation of the sequence alignment between RsaA and \textit{C. fetus} SapA shows few extended regions of sequence identity. Of the few regions of consistent homology most were RTX motifs. There are no tantalizing clues to be learned from BLAST analysis at this point, but it may be good platform to translate any future structural information to other unsolved proteins.

RsaA further isolates itself in its sequence by featuring a uniquely limited amino acid composition that is biased towards small, simple residues. This bias towards amino acids that are energetically less costly to synthesize is thought to be an evolutionary strategy for secreted proteins conserved across microbial life and especially in *C. crescentus*.\(^{130}\)

All crystallographic studies of S-layer proteins appear to require inhibition of S-layer formation. In the past co-crystallization with nanobodies\(^{24}\) and truncation\(^{127}\) approaches have resulted in monodisperse, crystallizable proteins. Here we describe the expression of an 804 amino acid (AA) C-terminal fragment of RsaA (RsaA Δ0–222) from its native host, *C. crescentus*. The untagged protein was purified from culture supernate and crystallized by hanging-drop vapor diffusion. Crystals were grown that diffracted to 2.5 Å. Diffraction data were collected at both the Canadian Light Source (CLS) and the Stanford Synchotron Radiation Light Source (SSRL). RsaA is the first Gram-negative S-layer protein to be crystallized.

Crystallization and X-ray diffraction are the two essential first steps of any structural determination but the next hurdle of solving the phasing problem is an equally crucial and fundamental problem in crystallography. Solving the phases has been the hurdle with which this project had particular difficulties. Many strategies have been pursued to complete the structural solution of RsaA. No successful solution to the phase problem has been determined. The atomic-level structure of RsaA remains elusive.

### 2.2 Methods

#### 2.2.1 Strain and plasmid construction

**Expression strain** The strain used to produce RsaA truncates was JS4032\(^{131}\). This strain can also be labeled *C. crescentus* CB2A ΔsapA ΔmanB ::repABC.

---


Expression vectors  The plasmid used initially in this study, pUC8CVX-RsaA Δ0–222, was a construct of the plasmid pUC8CVX containing an open-reading frame encoding for the C-terminal 804 AA of RsaA. pUC8CVX, digested with BamHI and HindIII, was ligated to the 2819 bp fragment from a similarly digested plasmid containing an open-reading frame encoding for RsaA with a BamHI endonuclease recognition sequence at the 666 bp position created in a previous study. The plasmid, pUC8CVX-RsaA Δ0–222, was then electroporated into the expression strain.

The same protocol as above was followed to generate an expression vector to produce RsaA Δ0–551, a smaller C-terminal truncate. To construct this plasmid, pUC8CVX-RsaA Δ0–551, the rsaA-truncate gene was swapped for a version encoding for only the C-terminal 475 AA of RsaA. An annotated sequence of full-length, wild-type RsaA is found in the appendix on page 154.

The plasmid that was used to produce RsaA Δ0–222 GSCC723, pUC8CVX-RsaA Δ0–222 GSCC723, was produced by replacing the rear portion of RsaA Δ0–222 in pUC8CVX-RsaA Δ0–222 with a preexisting engineered RsaA fragment. The fragment contained an in-frame insertion of the C-Myc-tag antigen at the 723rd codon. This replacement was done by exchanging 2630 bp BstXI–HindIII fragments. The C-Myc-tag is a decapeptide, EQKLISEEDL, that was derived from the oncogenic human protein Myc. In RsaA we predicted that the added charges might improve the ability of the protein to bind ions useful for the phasing problem. Our ‘GSCC’ cassette contains the C-myc-tag as well as a few other restriction sites that can be utilized for the instillation of future peptide cassettes, specifically a BgII site and a PstI site. The translated peptide sequence of the GSCC cassette is GSRSVNNASEQKLISEEDLRPSADGS.

2.2.2 Protein production

The same protocol, once it was finalized, was used for all the different constructs of RsaA (Δ0–222, Δ0–551, and Δ0–222 GSCC). The cells were grown in M16HIGG medium at 30°C for 72 hr. Cultures, 250 ml, were used in 2800 ml wide-bottom Fernbach flasks, resulting in shallow cultures that did not require shaking for sufficient aeration, as agitation of the culture led to macroaggregation of the secreted protein.

The cultures were centrifuged at 6300 rpm in JA-10 bottles for 35 min, the supernates were recovered and recentrifuged. The supernates were then filtered through a 0.22 µm microfilter to ensure no cells remained. The protein-containing supernate was initially concentrated by lyophilization and rehydration to 1/10 the original volume. The rehydrated protein solution was dialyzed against distilled water and microfiltered. Purification was performed over a 26/60 S100 Sephadex size-exclusion chromatography column and mobile phase of 200 mM NaCl and 10 mM Tris pH 7.5. The first major elution peak was pooled and dialyzed against distilled water and concentrated by placing the dialysis bags on polyethylene glycol (PEG) with an average molecular weight of 20 000 Da (Sigma-Aldrich). The protein solution was slowly concentrated to a final concentration of 3.5–4 mgml⁻¹. Dialysis against PEG was slow and gentle enough to prevent the protein from aggregating during concentration. Other methods of concentration, such as ultrafiltration caused RsaA Δ0–222 to come out of solution as unusable aggregates.

2.2.3 Dynamic light scattering spectroscopy

Dynamic light scattering (DLS) was used as a qualitative measure of protein quality. DLS has an accessible and rapid ability to determine the degree of aggregation and monodispersity of a purified protein sample. Prior to DLS analysis all protein samples were concentrated to at least 1 mgml⁻¹ and filtered through a 0.22 µm syringe filter. Sixty microlitres of filtered protein solution was pipetted into a glass-bottomed 384-well plate well. Standard instrument procedures were followed (Wyatt, DynaPro Plate ReaderTM).

---

2.2.4 Crystallization of RsaA constructs

RsaA Δ0–222 was initially crystallized at a concentration of 9 mg ml$^{-1}$. Sparse-matrix screening was performed using the Joint Centre for Structural Genomics (JCSG) Core I, II, III and IV screens (Qiagen).

Initial crystal formation conditions were not reproducible in a hanging drop vapor apparatus. Contrary to conventional wisdom, thin plates readily formed once the concentration of the protein and the PEG 8000 were both reduced. Thicker, more robust crystals formed when MgCl$_2$ was replaced with other Group 2, alkaline earth metal salts, specifically CaCl$_2$, SrCl$_2$, or BaCl$_2$. Although calcium is almost certainly the natural ion involved with RsaA folding, SrCl$_2$ was ultimately deemed the most successful at producing crystals. Final crystallization conditions were 8.5% (w/v) PEG 8000, 100 mM Tris pH 7.4, and 150 mM SrCl$_2$. Two crystal forms were observed, planar ‘plate’ crystals and hexagonal ‘pencil’ crystals. Occasionally the ‘plate’ crystals would grow in clusters where it was advantageous to break off a section of the cluster for diffraction (like the crystal in fig. 2.1). The ‘pencil’ crystals always formed as individual crystals.

Mother liquor (*i.e.* the crystallization well solution) amended with 25% glycerol was used as a cryo-protectant for all crystals prior to flash freezing in liquid nitrogen or a cryo-nitrogen stream.

2.2.5 Data collection and processing

X-ray diffraction data was collected, multiple times, at two synchrotron light sources: the CLS in Saskatoon, Saskatchewan and SSRL at Stanford University, California. Both facilities proved suitable for our data collection needs. The beamlines at CLS offered a slightly larger spectrum of available X-ray wavelengths; some of the high energy wavelengths proved useful for accessing the K-edges of strontium (0.7699 Å) and bromine (0.9202 Å). For our best diffracting crystal, diffraction was performed at the ‘CMCF 08ID-1’ beamline at CLS. A wavelength of 1.54 Å...
Figure 2.1: An example of a large cluster of ‘plate’ crystals of RsaA Δ0–222. The cluster was just over one millimeter in width from tip to tip. The largest crystal, on the right, was broken off and diffracted using synchrotron radiation, resulting in a resolution $>10$ Å. The color is imparted by a polarizing filter on the microscope, the crystals are naturally colourless.

(equal to the copper K-edge) was used to maximize the overall anomalous signal of iodide. As reference, fig. 2.2 shows theoretical anomalous scattering coefficients for most of the elements used in this study. Data were processed using HKL3000\textsuperscript{134} or XDS\textsuperscript{135} (see table 2.1); phasing attempts were performed using Shelx\textsuperscript{136} or AutoSol in Phenix\textsuperscript{137}. Data were collected through a full 360° of rotation, due the low symmetry found in the ‘plate’ crystals.

2.3 Results

2.3.1 Initial crystallization and optimization

Protein was produced in JS4032. The strain’s genomic copy of rsaA is inactivated by a nonsense frame-shift mutation at the 353\textsuperscript{rd} codon. The gene sapA has been deleted; it encodes for an S-layer associated protease (SapA) that has been implicated in degrading modified versions of RsaA.\textsuperscript{138} The surface polysaccharides, extracellular polysaccharide (EPS), polyrhamnan, and

Figure 2.2: Theoretical plots of anomalous scattering coefficients for various elements used in this study. The K-edges and L1-edges that are present within the range of this graph were indicated. Selenium was included because it is often used as a landmark to compare with other elements. The data presented here was obtained and plotted by the Edgeplots web tool provided by the Biomolecular Structure Center at the University of Washington (http://skuld.bmsc.washington.edu/scatter)

O-specific polysaccharide (OPS), were all eliminated by disrupting manB. manB encodes for phosphomannomutase, an important enzyme in the bio-synthesis of many deoxysugars (see chapter 3 for our investigation on the structure of the lipopolysaccharide (LPS)). Removing these carbohydrate structures causes RsaA to no longer associate with the cell surface and aids in downstream processing, giving the cells favorable centrifugability. The operon repABC has also been knocked into this strain to allow for the replication of small broad-host range plasmids that we use for protein expression.\textsuperscript{139}

The protein construct, RsaA Δ0–222, was expressed in both a traditional *Escherichia coli* expression system and in our lab’s *C. crescentus* protein expression system. All attempts to produce protein in *E. coli* resulted in highly insoluble inclusion bodies. Expression levels of unmodified RsaA in its native host is notably high. After numerous efforts to produce protein in an appropriate form for crystallization, that is monodisperse, soluble protein, we were forced to conclude that other approaches were needed. A number of counter-intuitive methods had to be used to eventually produce a crystallizable form of RsaA from *C. crescentus*. To begin with the protein encoding gene had to be genetically altered to produce a truncated form of RsaA, the cells had to be cultured under low agitation, and the resulting secreted protein had to be concentrated very slowly.

We had observed that cells cultured in shaken test tubes produced aggregated protein in the form of large insoluble floating masses, while cells cultured in rolled test tubes produced soluble protein.

The reasons for an N-terminal truncation of RsaA are two-fold. Deletion of the N-terminus results in protein that no longer anchors itself to outer membrane, resulting in soluble supernatant protein; and the N-terminus has been identified as the center of three-fold symmetry; its deletion should prevent wide-scale S-layer formation and aggregation. This is our strategy for preventing S-layer formation as discussed in section 2.1.

Attempts to crystallize RsaA Δ0–222 started with the pre-assembled crystallization screens JCSG Core I, II, III, and IV. Higher protein concentrations were sought because it is generally believed that higher protein concentrations would lead to better results and literature sources recommended protein concentration ranges centering on 10 mg ml\(^{-1}\) [142] The RsaA Δ0–222 truncate was readily concentrated to 5–7 mg ml\(^{-1}\), but when concentrated beyond 7–9 mg ml\(^{-1}\) aggregation began to occur, either visual macroaggregation or microaggregation assessed by DLS. DLS was used as a qualitative measure of protein quality and guided our protein production.

---

Figure 2.3: The purification of RsaA Δ0–222 as monitored by SDS-PAGE. A. 7 µl of culture supernate of our *C. crescentus* expression strain expressing RsaA Δ0–222. B. Each lane has 15 µl from separate size-exclusion chromatography elutions from the major peak containing RsaA Δ0–222. The contents of the tubes represented by these 4 bands were pooled, dialyzed, concentrated, and crystallized.

and purification protocols. Higher concentrations of RsaA Δ0–222 would exhibit lower levels of monodispersity and larger fractions of high-molecular weight aggregates, both predictive signs of reduced crystallizability. Figure 2.4 shows DLS plots of RsaA Δ0–222 at 9 mg ml\(^{-1}\) and 3 mg ml\(^{-1}\).

For RsaA Δ0–222, initial screens were assembled at 9 mg ml\(^{-1}\). Small crystalline shapes were observed after one week at room temperature. Those first crystals grew in JCGS Core I, well 28 with conditions of 20% PEG 8000, 200 mM MgCl\(_2\), and 100 mM Tris pH 7.

Figure 2.4: The DLS plots of various constructs and preparations of RsaA. The data are shown as histograms representing the population of particle sizes observed in the protein solutions. 

A. RsaA Δ0–222 at 3.5 mg ml\(^{-1}\). This protein sample was used successfully to produce well-diffracting crystals. 

B. RsaA Δ0–222 at 9 mg ml\(^{-1}\). 

C. RsaA Δ0–551 at 5 mg ml\(^{-1}\). 

D. Full-length wild-type RsaA at 2 mg ml\(^{-1}\).
The first attempts to transfer from a sitting-drop strategy, as used in initial screening, to a
hanging-drop strategy, to be used in optimization, were unsuccessful due to widespread protein
precipitation/aggregation. Shifting our divalent cation from Mg\(^{2+}\) to Ca\(^{2+}\) led to small crystals in
our drops. Crystals that were grown with Ca\(^{2+}\) were very thin, but sizable (>0.5 mm) in the other
two dimensions. The biggest improvements came from shifting the cation again to Sr\(^{2+}\) and
lowering the concentrations of the protein and precipitants. Due to low protein concentrations,
larger drop sizes (6–10 µl) were utilized and larger ratios between protein and mother liquor
were (1.5:1–3:1) used so that initial conditions would not cause protein precipitation but protein
concentrations would eventually get high enough to be supersaturated. Fine tuning of protein
concentration and drop ratios were mostly used to optimize nucleations in the drops. RsaA
\(\Delta 0–222\) always nucleated well, leading to more drops full of small crystals than clear drops
with no crystals. A protein concentration in the range of 3.5–4.5 mg ml\(^{-1}\), mixed in a ratio of
1.5:1 with mother liquor, had the best results, averaging around 3 nucleations per drop. Further
optimization to the crystallization conditions, such as lowering salt concentration (200 mM to
150 mM) and raising the pH (7 to 7.4), were of arguable efficacy.

For RsaA \(\Delta 0–551\), no crystal formation was observed in any crystallization screening con-
dition or in any protein concentration used. Once successful crystallization conditions were
observed for RsaA \(\Delta 0–222\), those conditions were used with RsaA \(\Delta 0–551\) but no crystals were
observed. Figure 2.4 shows the DLS plot of a representative RsaA \(\Delta 0–551\) protein sample.

**2.3.2 Properties of the RsaA \(\Delta 0–222\) crystals**

Our first diffracting crystals of RsaA \(\Delta 0–222\) were ‘plate’ crystals having a rhomboid shape. Figure 2.5 shows examples of these rhomboid crystals. The proportional dimensions of the
crystals roughly match the relative dimensions of the crystal unit cell (see Table 2.1), although
the thickness seemed variable. The crystals grew often as clusters or stacks, which did not aid
in diffraction. Figure 2.1 shows one example of how many of these crystal clusters appeared.
The rhomboid ‘plate’ crystals were very strong, resisting crushing so well that we were initially
concerned that they were salt crystals. The ‘plates’ also resisted X-ray induced damage well—often diffracting acceptably well after two or three full data collections.

Later crystallization efforts for RsaA Δ0–222 started to yield crystals that were not rhombo-roid, rather they were hexagonal prisms, like pencils. Drops identical to and next to drops producing ‘plate’ crystals would often grow ‘pencil’ crystals. A few drops even produced both ‘plate’ crystals and ‘pencil’ crystals. These ‘pencil’ crystals looked like protein crystals under a polarizing filter on the microscope and they were confirmed to be protein by X-ray diffraction. The ‘pencil’ crystals were very fragile and often broke or cracked under gentle manipulation. The ‘pencils’ grew as solitary crystals often on contact with the lower surface of the crystallization drop. Figure 2.6 shows a few examples of the hexagonal prism crystals of RsaA Δ0–222.

Low pH solutions would readily crack and destroy any form of RsaA Δ0–222 crystal. Unfortunately, this pH limitation precluded us from using some soaking agents like uranium compounds and higher concentrations of tantalum bromide. The ‘plate’ crystals were generally more stable in soaking conditions with high salt concentrations than the ‘pencil’ crystals. The ‘pencil’ crystals would quickly crack in high concentrations of KI and KBr.

2.3.3 Crystallization of RsaA Δ0–222 GSCC723

RsaA Δ0–222 GSCC723 readily crystallized in the successful conditions used for RsaA Δ0–222. However, the crystals that formed had a different morphology. The crystals looked dendritic, like trees or snowflakes. Extensive optimizations were performed to improve the shape and size of the crystals but no improved conditions could be found. The final conditions used were 9% PEG, 100 mM tris pH 7.4, and 150 mM SrCl₂. The protein concentration was 3.5 mg ml⁻¹, 6 µl of protein solution was mixed with 3 µl of mother liquor.

A few of the dendritic crystal clusters had branches that were clean, individual crystals. In those cases, the crystal branches were gently broken off of the main cluster and used for X-ray diffraction.
Figure 2.5: A panel of ‘plate’ crystals of RsaA Δ0–222. The left figure show an individual crystal with the common rhomboid shape and relatively thin width. The center figure shows a wide-view of a drop containing multiple ‘plate’ crystals. The right figure shows three crystals of RsaA Δ0–222 that were grown in the presence of 5 mM EuNO₃. The color in the first two images is imparted by a polarizing filter on the microscope, the crystals are naturally colourless. The colour of the right figure is unadjusted and representative of the natural colour of the crystals.

2.3.4 Properties of the RsaA Δ0–222 GCC723 crystals

The only crystals that formed when using the RsaA Δ0–222 GCC723 protein were ‘dendritic’ crystals that resembled trees or snowflakes. Figure 2.7 shows a few examples of these crystals. Most of these crystals were too clustered to be useful for diffraction but a few of the clusters had branches that could be removed and used for diffraction. Figure 2.8 shows a few of the dendritic clusters that were broken apart and diffracted.

The RsaA Δ0–222 GCC723 crystals had similar properties to the RsaA Δ0–222 ‘plate’ crystals; the GCC723 crystals were very strong and stable in most soaking conditions. Unfortunately due to the poor diffraction, the GCC crystals’ potential ability to bind more heavy ions for phasing was never realized.
Figure 2.6: A panel of ‘pencil crystals of RsaA Δ0–222. The left figure shows how some of the ‘pencils’ grew in the presence of ‘plate’ crystals, note the ‘plates’ in that panel were not suitable for X-ray diffraction. The right figure highlights a particularly large ‘pencil’ crystal. These crystals were very fragile. The color in the images is imparted by a polarizing filter on the microscope, the crystals are naturally colourless.

Figure 2.7: Examples of ‘dendritic’ crystals formed when the RsaA Δ0–222 GSCC723 protein construct was crystallized. The color is imparted by a polarizing filter on the microscope, the crystals are naturally colourless.

2.3.5 X-ray diffraction and phasing

Crystal variability impeded the determination of the structure of RsaA. The majority of the crystals of RsaA that were grown had very poor X-ray diffraction. There was not an effective way to predict which crystal would diffract well. Our best results came from a large crystal that was relatively thin—two features that seemed to be important because small crystals did not diffract well and overly thick crystals were plagued with smeared, messy reflections. Figure 2.9 shows one frame of data from our best dataset. An interesting aspect of the data, specifically
Figure 2.8: A panel of better ‘dendritic’ crystals of RsaA Δ0–222 GSCC723. These are examples of some of the GSCC723 crystals having branches that were suitable for X-ray diffraction. A few of the side crystals were broken off and used for diffraction. The color in the images is imparted by a polarizing filter on the microscope, the crystals are naturally colourless. The colour of the right figure is unadjusted and representative of the natural colour of the crystals.

from ‘plate’ and ‘dendrite’ crystals, is the anisotopy we observed in our diffraction. Due to the often poor quality of the data, we would always collect our data through 360° which would lead to long collection strategies and increased radiation damage to the crystal. Radiation damage was never too much of an issue though, as the ‘plate’ crystals were remarkably resistant to damage. Table 2.1 gives an overview of the X-ray diffraction data.

Once we had crystals that diffracted to a suitable extent the final step in generating an electron density map would be phasing. Molecular replacement strategies were not possible due to a lack of homologous structures in the Protein Data Bank. The RTX motifs in RsaA provide a tempting region for structural homology, such as the structure in fig. 1.6. Neither the structures of solved RTX motifs as molecular replacement nor the generated Phyre structure for RsaA’s RTX motifs were effective molecular replacement search models. To try and solve the phase problem we turned to the X-ray fluorescence techniques of multi-wavelength anomalous dispersion (MAD) and single-wavelength anomalous dispersion (SAD).

Our main two candidates for MAD/SAD phasing elements were the halides, iodine and bromine. These halides were introduced into the crystals by soaking the crystals in KI and KBr. Iodine does not have an accessible K or L edge but provides a SAD suitable anomalous signal at achievable wavelengths. Refer back to fig. 2.2 for the theoretical profiles of anomalous signal. Iodide-soaked ‘plate’ crystals, among the first conditions ever tested, diffracted well and an anomalous signal could be detected in their datasets but not enough to allow for densities to be calculated. A great deal of effort went into trying to improve our initial datasets with more of an anomalous component from iodide. Very few subsequent iodide-soaked crystals diffracted at all and those that did were of low resolution. Bromine has an accessible $K_a$ edge at 0.9202 Å and so it is useful for MAD. Unlike iodide, bromide-soaked crystals did not have the high rate of crystals with no diffraction but like iodide, few bromide-soaked crystals diffracted beyond 4 Å.

Many other phasing compounds were tested. The lanthanides all provide a strong anomalous signal and have been successfully used in the past for SAD. Lanthanum, lutetium, and europium were all used as soaking agents but none produced any detectable anomalous dispersion in the resulting datasets. Europium was also successfully used in the mother-liquor of the crystals to be co-crystallized. Unfortunately, the co-crystallized crystals had no detectable anomalous signal, despite good diffraction. Europium’s $L_1$ edge is accessible with synchrotron radiation (see fig. 2.2), and MAD fluorescence scans were done to confirm that there was no detectable Europium in the crystals. Figure 2.5 features a few of the europium co-crystallized crystals. In the vein of co-crystallization, strontium and barium were effective components of the mother liquor, and possibly potent phasing elements, but neither provided an effective solution to the phase problem when tried. The more exotic compounds Tantulum bromide ($Ta_6Br_{12}$) and triiodide were also tried without resulting in the necessary data.

There were only ever a few ‘dendritic’ crystals that were suitable for X-ray diffraction and none of those few diffracted to an extent or in a manner that was conducive to solving the

---

Figure 2.9: This figure shows a frame from a X-ray diffraction data collection. This crystal was a large ‘plate’ crystal that was our best diffracting crystal to date. The crystal was soaked in 0.5 M KI for 30 sec. The wavelength of the incident X-ray beam was set to 1.541 Å (the copper K$_\alpha$ edge). The resolution from this data was 2.50 Å. The crystal used here was soaked in an iodide solution. Notice the anisotropy in diffraction, i.e. the spots reach further from the center to the upper right and lower left.

structure of RsaA Δ0–222 GSCC723. One ‘dendritic’ crystal did diffract reliably to 3.5 Å but exhibited multiple overlapping signals, probably due to separate crystals being fused together.

The ‘pencil’ crystals never diffracted to any significant extent, at most to 6–10 Å. Initial processing of one of the ‘pencil’ diffraction patterns suggested that its space group was P6. A P6
space group has higher symmetry than the P2$_1$ space group we observed in our ‘plate’ crystals. We had hopes that the higher symmetry would provide better quality data, unfortunately no well performing ‘pencil’ crystals were ever isolated. The only pencil that we did try to perform a complete data collection on quickly lost diffraction due to radiation damage.

2.4 Discussion

Our protein construct, RsaA Δ0–222, was expressed our in C. crescentus protein expression system. All attempts to produce protein in E. coli resulted in highly insoluble inclusion bodies that resisted refolding (data not shown). Expression levels of RsaA in its native host is notably high but the system did not initially produce protein in an appropriate form. The problem of aggregation persisted in C. crescentus secreted RsaA; the problem was assuaged by the N-terminal deletion construct, RsaA Δ0–222, but was only completely solved when we started to culture the cells without agitation in wide, shallow flasks. Our shallow culture expression system proved successful at producing 130–150 mg of purified, concentrated protein per litre of culture supernate. RsaA is the natural product of C. crescentus but this system could be useful in producing exogenous protein for crystallization by fusion to the required type 1 section signal.

The reasons for an N-terminal truncation of RsaA are two-fold: (1) deletion of the N-terminus produces protein that no longer anchors itself to outer membrane, resulting in free-secreted supernatant protein and (2) the N-terminus has been identified as the center of three-fold symmetry. Therefore, its deletion results in soluble protein that no longer leads to wide-scale S-layer formation or aggregation and is freely secreted into the supernate. We can visualize the success in preventing aggregation and S-layer formation by measure the particle sizes of the RsaA constructs in situ. Figure 2.4 shows how the RsaA Δ0–222 protein behaves in solution compared to full-length RsaA. The latter exhibits a much higher particle size and rate of aggregation. The success of structurally characterizing any S-layer protein impinges upon

---

overcoming such issues and our strategies resulted in large well-diffracting crystals.

Our ‘plate’ crystals diffracted to 2.5 Å with unit-cell parameters of a=210.79 Å, b=80.74 Å, c=221.9 Å, and β=117.46 degrees. These crystal unit-cell sizes almost exactly match the 2D unit size of native S-layer repeating unit, a hexamer of RsaA monomers. These results suggest that the unit-cell of our crystals is six copies of RsaA Δ0–222, corresponding to a solvent content of about 60%. Our hexagonal columnar crystals have never diffracted to a suitable resolution, but they are promising because they have higher symmetry and that may result in improved datasets. As our protein has no close matches in the structural databases to use for the purposes of molecular replacement, solving the phase problem was attempted with iodine derivitization with SAD and bromide with MAD. Neither of these approaches gave us the required anomalous data that we would need for a solution. RsaA does not have many charged amino acids, even after we installed the GSCC cassette, and that is why the halides have been so strongly pursued as they do not require specific residues for activity. Patterson analysis of the dataset suggests issues with pseudotranslational symmetry confounding phasing attempts, with the largest Patterson peak at a height of about 30%. Twinning analyses indicate that no twinning is suspected.

As of this writing, the crystallography portion of this work is still ongoing. Only the phasing problem remains to be solved in this structural determination project; we hope that a solution exists and that we will find it soon. The fact remains that we have developed techniques to crystallize the S-layer protein and generally that is the most important step in a crystallization project. To conclude this chapter, below is a quote from Alexander McPherson’s 2004 review on macromolecular crystallography:

“Presently, and in the foreseeable future, the only technique that can yield atomic level structural images of biological macromolecules is X-ray diffraction analysis as applied to single crystals. While other methods may produce important structural and dynamic data, for the purposes described above, only X-ray crystallography is

---

adequate. As its name suggests, application of X-ray crystallography is absolutely
dependent on crystals of the macromolecule, and not simply crystals, but crystals
of sufficient size and quality to permit accurate data collection. The quality of the
final structural image is directly determined by the perfection, size, and physical
properties of the crystalline specimen, hence the crystal becomes the keystone
element of the entire process, and the ultimate determinant of its success.”
Table 2.1: A summary of the X-ray diffraction data from a planar crystal of RsaA Δ0–222

<table>
<thead>
<tr>
<th></th>
<th>RsaA Δ0–222</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>210.79, 80.74, 221.9</td>
</tr>
<tr>
<td>β (°)</td>
<td>117.46</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.00–2.50 (2.54–2.50)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.13 (0.54)</td>
</tr>
<tr>
<td>$I / \delta I$</td>
<td>13.7 (2.8)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (97.6)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.8</td>
</tr>
</tbody>
</table>
Chapter 3

The core and O-polysaccharide structure of the Caulobacter crescentus lipopolysaccharide

“And so, progressively, the veil behind which Nature has so carefully concealed her secrets is being lifted where the carbohydrates are concerned.”
— H. Emil Fischer, 1902 Nobel lecture

My great-great-great-great-great-great-grand adviser.

3.1 Introduction

Caulobacter crescentus is an aquatic alphaproteobacterium well known for a stalked, crescent cell morphology, asymmetric cell division, and a protein surface layer (S-layer). Caulobacter crescentus is a widely studied model organism but despite this, the structure of its lipopolysaccharide (LPS) has not previously been fully determined.

Interest in the LPS of Caulobacter crescentus is focused on its immunological profile and its structural role as an anchor for the self-assembled, para-crystalline S-layer. The LPS of Caulobacter crescentus possesses a much reduced immunogenic activity, most likely due to its lipid A structure, which is significantly different from that of LPS from enteric bacteria. The lipid A structure has been reported. It is a unique molecule containing a di-diaminoglucose backbone (instead of di-glucosamine) and two galacturonate moieties that replace the canonical phosphates that are

---

on each end of the disaccharide in most lipid A molecules. Figure 3.1 shows the structure of the *C. crescentus* lipid A.

![Diagram of lipid A structure](image)

**Figure 3.1:** The structure of the *C. crescentus* lipid A.

The *C. crescentus* S-layer non-covalently attaches to the O-specific polysaccharide (OPS)\(^{128}\). However, the OPS structure has not been resolved. Genetic analyses have pointed towards the unusual dideoxyamino sugar N-acetylperosamine being a major component\(^{147}\). A notable

feature of this OPS is that it exists completely hidden beneath the S-layer, inaccessible to the environment by at least such tools as specific antibodies. Carbohydrate structures from non-pathogenic bacterial LPS are rarely studied and an LPS that is sequestered beneath an S-layer is not represented in the literature.

In the present study our data has determined the core oligosaccharide structure from *C. crescentus* CB15 NA1000 (advancing an earlier report of core composition), as well as the central backbone and non-reducing ends of its OPS. Unexpectedly, we identified a previously unknown rhamnan polysaccharide. Along with previous reports on lipid A and extracellular polysaccharide (EPS), we believe that all the major carbohydrate structures in *C. crescentus* cell envelope have now been solved.

### 3.2 Methods

#### 3.2.1 Bacterial strain construction and growth conditions

The strain used for the preparation of LPS was JS1025, a derivative of *C. crescentus* CB15 NA1000. The salient features are that it has an engineered amber mutation in rsaA leading to the loss of the S-layer and the gene CCNA_00471 has been inactivated by a partial deletion. CCNA_00471 encodes a putative GDP-L-fucose synthase. The knockout (Δ471) confers a deficiency in an EPS that was previously found to contain L-fucose. CCNA_00471 was disrupted in the same manner as previously in JS4038, except the starting strain used here was JS1023.

Cells were grown to mid-to-late log phase (OD$_{600}$ = 0.9) in M16HIGG defined medium at 30°C in 2.8 l Fernbach flasks containing 1250 ml of medium, with shaking at 100 rpm.

---

M16HIGG is a modification of M6HIGG medium containing 0.31% glucose, 0.09% glutamate, 1.25 mm sodium phosphate, 3.1 mm imidazole, 0.05% ammonium chloride, and 0.5% modified Hutner’s Mineral Base.

3.2.2 Lipopolysaccharide isolation

LPS was isolated from the cells via disrupting the outer membrane by chelation. The protocol we used was a modification of the procedure reported by Walker et al. Cells were centrifuged at 12,400 x g for 10 min. The pellets were suspended with distilled water and recentrifuged. These pellets were resuspended in 1/10 original culture volume in phosphate-buffered saline (PBS) amended with 35 mm ethylenediaminetetraacetic acid (EDTA), agitated at room temperature for 10 min and then centrifuged at 15,300 x g for 15 min. The supernatant was retrieved and re-centrifuged, as before, to ensure clarity and then dialyzed against 5 mm MgCl₂. DNase and RNase were added to final concentrations of 10 µg ml⁻¹ and 100 µg ml⁻¹, respectively, and incubated at 37°C for 2 h. Proteinase K was added to a final concentration of 0.3 mg ml⁻¹ and the preparation was incubated at 50°C overnight. The sample was then ultracentrifuged at 184,000 x g for 3 h. Glassy pellets formed which were suspended in distilled water to 1/100 original culture volume. A Bligh-Dyer extraction was performed to reduce contaminating lipids.

3.2.3 Bligh Dyer extraction

A Bligh Dyer extraction was performed on all LPS preparations to reduce the presence of contaminating lipids. The extraction was performed as first published. In short, to one volume of aqueous LPS sample, 3.75 volumes of chloroform/methanol (1:2 v/v) was added and the sample was vortexed for 30 sec. Chloroform, 1.25 volumes, was added and the sample was vortexed again for 30 sec. Water, 1.25 volumes, were added and the sample was vortexed a final time before centrifugation and extraction.

---

time for 30 sec. This mixture was then centrifuged at 15 300 x g for 10 min. After centrifugation, the mixture separated into two phases, a lower organic phase and an upper aqueous phase. The LPS partitions into the aqueous phase, while other lipids partition into the organic phase. The aqueous phase was retrieved. Samples were stored at 4°C for short term use or frozen at -20°C for long term storage. Occasionally, NaN₃ was added to a concentration of 3 mM to prevent microbial growth and it was never found to effect analysis.

3.2.4 Silver stain and gel electrophoresis

Discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed with a 13% separating gel. Detection of LPS was done by periodate oxidation and silver staining as described by Zhu et al. The silver stain protocol is a newer technique that uses vitamin C (ascorbic acid) in place of formaldehyde in traditional LPS stains. Table 3.1 give the details of the silver stain procedure. All samples were boiled for 5 min in sample buffer prior to loading them to a gel. Sample buffer is made as a 4X concentrate comprising 200 mM Tris pH 6.8, 8% sodium dodecyl sulfate, 40% glycerol, 4% β-mercaptoethanol, 0.25% bromophenol blue.

3.2.5 Preparation of RsaA Δ277–784 probe

The plasmid, p4BRsaAΔ277–784, was in our laboratory from the work of Bingle and Nomellini. In short, various versions of rsaA were engineered to have BamHI sites at locations throughout the gene. The BamHI–HinDIII fragment from a version of rsaA with a BamHI site at the 831 bp from the start codon was replaced with a BamHI–HinDIII fragment from a version of rsaA with a BamHI site at the 2352 bp position. In effect this made a rsaA gene missing 1521 bp from the center, corresponding to amino acids 277–784 within the resulting protein. This

Table 3.1: LPS silver stain procedure

<table>
<thead>
<tr>
<th>Step</th>
<th>Solution</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidation</td>
<td>30% EtOH, 10% HOAc, 0.7% periodic acid</td>
<td>10 min</td>
</tr>
<tr>
<td>Rinse</td>
<td>dH₂O</td>
<td>5 min x2</td>
</tr>
<tr>
<td>Silver</td>
<td>0.2% AgNO₃</td>
<td>5 min</td>
</tr>
<tr>
<td>Rinse</td>
<td>dH₂O</td>
<td>20 sec x2</td>
</tr>
<tr>
<td>Develop</td>
<td>3% NaCO₃, 0.04% Na₂S₂O₃, 0.02% ascorbic acid, 0.05% NaOH</td>
<td>8 min</td>
</tr>
<tr>
<td>Stop</td>
<td>10% HOAc</td>
<td>1 min</td>
</tr>
</tbody>
</table>

‘internally truncated’ gene was then moved into the context of our expression vector ‘p4B’ as a EcoRI–HindIII fragment. p4BRsaAΔ277–784 was electroporated into C. crescentus JS1030. JS1030 is the same strain used for protein production in chapter 2; its construction is detailed in section 2.2.2. Isolation and concentration of the RsaA Δ277–784 probe followed the same protocol used for RsaA Δ0–222, as found in section 2.2.2, but unpurified and unconcentrated cell culture supernate, containing secreted RsaA Δ277–784, was also to be an effective probe for far-western blotting.

For the protein probe used in the glycan array screening, the RsaA Δ277–784 was labeled with fluorescein isothiocyanate (Life Technologies); manufacturer’s instructions were followed for the labeling procedure.

3.2.6 Western blotting and far-western blotting

Western blotting of LPS followed standard laboratory blotting techniques. SDS-PAGE gels of LPS were electro-blotted on to a nitrocellulose membrane that was subsequently agitated in Tris-buffered milk blocking buffer for one hour. Occasionally, spot-blot immunoblots were performed. Spot-blots are like western blots that forgo the SDS-PAGE; the sample is spotted

---

directly on to the membrane and allowed to dry.

The primary probe was rabbit serum raised against a preparation of \textit{C. crescentus} holdfast. The holdfast is a complex carbohydrate ‘glue’ that is present on the stalk of \textit{C. crescentus}. The holdfast acts to anchor the cells to surfaces. This so-called anti-holdfast serum was fortuitously found to have specific activity against the \textit{C. crescentus} LPS.\textsuperscript{[28]} AlexaFluor680 goat anti-rabbit IgG secondary antibody (Invitrogen) was used at 50,000$^{-1}$ and was detected and quantified by the Odyssey Infrared Imaging System using Odyssey 2.0 software (Licor Biosciences, Lincoln, NE), similar to previously published methods.\textsuperscript{[31]} Two 5 min washes were performed between each step.

Far-western blotting was originally developed to screen protein expression libraries with $^{32}$P-labeled glutathione-S-transferase-fusion proteins,\textsuperscript{[39]} but the idea has developed as a quick method to test for \textit{in vitro} protein interactions on the surface of a western blot membrane. The idea to use far-western blotting to check for RsaA–LPS interaction was spurred by the realization that the denaturing conditions in standard \textit{SDS-PAGE} would have no detrimental effect on the primary structure of LPS. Our far-western blot protocol started as a standard western blot, except the primary probe was replaced with RsaA$\Delta277–784$ in distilled water. Primary probe concentration was not a critical variable in the range tested, 0.1–4 mg ml$^{-1}$. Anti-RsaA$\Delta277–784$ rabbit serum was used as a secondary probe, followed by AlexaFluor680 goat-anti rabbit IgG as a tertiary probe. Controls checking non-specific binding between the LPS and the secondary probes and tertiary probes showed were performed (data not shown).

### 3.2.7 Matrix assisted laser desorption/ionization-time of flight mass spectroscopy

MALDI-TOF analysis was done with the UBC Proteomics Core Facility. For our initial analysis we used LPS purified from \textit{SDS-PAGE} gels, but the resulting acrylamide/polyacrylamide contamination proved insurmountable (data not shown).

The MALDI-TOF procedure used our purified LPS mixed with a sinapinic acid matrix and analyzed on a Applied Biosystems Voyager System 4311. Table 3.2 highlights the machine setting used for the MALDI-TOF.

**Table 3.2:** The settings of the Applied Biosystems Voyager 4311 that were used to do MALDI-TOF analysis of the *C. crescentus* LPS

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of operation:</td>
<td>Linear</td>
</tr>
<tr>
<td>Extraction mode:</td>
<td>Delayed</td>
</tr>
<tr>
<td>Polarity:</td>
<td>Negative</td>
</tr>
<tr>
<td>Accelerating Voltage:</td>
<td>25 kV</td>
</tr>
<tr>
<td>Extraction Delay Time:</td>
<td>670 nsec</td>
</tr>
<tr>
<td>Number of Laser Shots:</td>
<td>100 per spectrum</td>
</tr>
<tr>
<td>Laser Intensity:</td>
<td>2933</td>
</tr>
<tr>
<td>Laser Rate:</td>
<td>20 Hz</td>
</tr>
<tr>
<td>Calibration Matrix:</td>
<td>Sinapinic Acid</td>
</tr>
<tr>
<td>Low Mass Gate:</td>
<td>2000 Da</td>
</tr>
</tbody>
</table>

### 3.2.8 Periodic acid-Schiff stain

Periodic acid-Schiff stain was performed on LPS samples that were run on SDS-PAGE.

The staining procedure starts by soaking the gel on a slow shaker in 0.7% periodic acid in distilled water for 15 min. The gel is then rinsed in distilled water and placed in Schiff reagent and shaken slowly for 15 min. The gel is then washed in distilled water for 5 min, at which time pink bands should appear that are corresponding to polysaccharides/LPS in the gel.

The protocol to prepare Schiff reagent started by dissolving 5 g of basic fuchsin in 900 ml of boiling distilled water. Once dissolved the fuchsin solution was removed from heat and allowed to cool for a few minutes, then 100 ml of 1 M HCl was slowly added. After the solution was allowed to cool to room temperature, 10 g of K$_2$S$_2$O$_5$ was added. This solution was thoroughly
mixed, then incubated in the fume hood overnight. The next day, a quantity of activated charcoal were added to the Schiff reagent to decolourize the solution. The solution was stirred and then filtered through Watman paper #1. The now clear Schiff stain solution was ready for use. A test for Schiff reagent activity is to add a few drops of Schiff reagent to 5 ml of formaldehyde in a test tube. The mixture should instantly turn a bright pink colour, indicating activity.

3.2.9 NMR spectroscopy

Nuclear magnetic resonance spectroscopy (NMR) experiments were carried out on a Varian INOVA 600 MHz ($^1$H) spectrometer with 5 mm gradient probe at 25–50°C with acetone internal reference (2.225 ppm for $^1$H and 31.45 ppm for $^{13}$C), using standard pulse sequences gradient correlation spectroscopy (gCOSY), total correlation spectroscopy (TCOSY) (mixing time 120 ms), rotating frame nuclear Overhauser effect spectroscopy (ROESY) (mixing time 300 ms), gradient heteronuclear single quantum coherence (gHSQC), and gradient heteronuclear multiple bond coherence (gHMBC) (100 ms long range transfer delay), heteronuclear multiple-quantum correlation spectroscopy (HMQC) for $^1$H-$^{31}$P correlation, $J_{HX}$ set to 10 Hz. Acquisition time was kept at 0.8–1 sec for H-H correlations and 0.25 sec for heteronuclear single quantum coherence (HSQC). 256 increments were acquired for $t_1$ in all 2D spectra, except 512 for gCOSY.

3.2.10 LPS chromatography

Gel chromatography was performed on a Sephadex G-15 column (1.5 x 60 cm) or a Bio-gel P6 column (2.5 x 60 cm) in pyridine-acetic acid buffer (4 ml:10 ml:1 l water), and monitored by refractive index detector (Gilson). Anion exchange chromatography was done on an Hitrap Q column (2x5 ml size, Amersham), with ultraviolet Light (UV) monitoring at 220 nm in a linear gradient of NaCl (0–1 M, 1 h) at the 3 ml min$^{-1}$. Fractions of 1 min were collected and additionally tested for carbohydrates, by spotting on an SiO$_2$ thin-layer chromatography (TLC) plate, dipping them in 5% H$_2$SO$_4$ in ethanol and heating with a heat-gun. All fractions of
interest were dried in a Savant drying centrifuge and $^1$H spectra were recorded for each fraction without desalting. For 2D NMR, desalting was performed on a Sephadex G15 column.

### 3.2.11 Monosaccharide analysis

Samples with added inositol standard were hydrolysed with 3 M trifluoroacetic acid (TFA) at 120°C. Monosaccharides were converted to alditol acetates by conventional methods and identified by gas chromatography-mass spectroscopy (GC-MS) on a Varian Saturn 2000 instrument on a DB17 capillary column (30 m x 0.25 mm ID x 0.25 µm film) with helium carrier gas, using a temperature gradient 170°C (3 min)–250°C at 5°C min$^{-1}$.

### 3.2.12 Determination of absolute configurations of monosaccharides

To the polysaccharide sample (0.2 mg) (R)-2-BuOH (0.2 ml) and acetyl chloride (0.02 ml) were added at room temperature, heated at 90°C for 2 h, dried by air stream, acetylated, analysed by GC-MS as described above. Standards were prepared from monosaccharides of known configuration with (R)- and (S)-2-BuOH.

### 3.2.13 Methylation analysis

For the methylation analysis core sample (2 mg) was dephosphorylated with 50 µl of 48% HF for 20 h at +10°C, diluted with 2 ml of ethanol, the precipitate collected by centrifugation, washed with 2 ml of ethanol, and dried.

Methylation was performed by Ciucanu-Kerek procedure$^{160}$ Half a milligram of the sample was dissolved in 0.5 ml of dry DMSO with heating at 100°C for 5-10 min until complete dissolution. Powdered NaOH (about 50 mg) was added and the mixture was stirred for 30 min. 0.2 ml of MeI was added and the mixture was stirred for a subsequent 30 min. The sample was then flushed with air to remove the MeI and diluted to 10 ml with water. The sample was

---

passed through a C18 Seppak cartridge, washed with 10 ml of water, and then the methylated compound was eluted with 5 ml of methanol. The methylated product was hydrolysed with 3 M TFA (120°C, 3h), dried, reduced with NaBD₄, and the reagent destroyed with 0.5 ml of 4 M HCl. The solution was dried under a stream of air and dried twice more with the addition of MeOH (1 ml). The sample was acetylated with 0.4 ml Ac₂O and 0.4 ml pyridine for 30 min at 100°C. It was then dried and analysed by GC-MS.

3.2.14 Periodate oxidation

Polysaccharide (PS) (10 mg) was dissolved in water (2 ml). NaIO₄ (20 mg) was added and the solution was incubated at room temperature for 24 h. Ethylene glycol (0.2 ml) and an excess NaBD₄ were added. The solution was then kept for 1 h before being treated with 0.2 ml of AcOH and desalted on a Sephadex G-15 column. The product was hydrolysed with 2% AcOH, 2 h at 100°C, and separated on a Sephadex G-50 column to give oligosaccharide (OS)₁.

3.2.15 Cell culture

The HEK-Blue htloll-like receptor 4 (TLR4)™ (InVivoGen) cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) (Gibco) with 10% heat-inactivated fetal bovine serum (FBS) (Thermo Scientific), 2 mM GlutaMAX, 25 mM HEPES buffer, and 50 U ml⁻¹ penicillin and 50 µg ml⁻¹ streptomycin (GIBCO, Invitrogen). The THP-1 cells were grown in Roswell Park Memorial Institute medium (RPMI) 1640 (Gibco) with 10% heat-inactivated FBS (Thermo Scientific). Cells were incubated at 37°C in humid air with 5% CO₂. The FBS was heat inactivated by incubating it at 56°C for 30 minutes.

3.2.16 HEK-Blue hTLR4 assay

To assay the stimulatory effect C. crescentus LPS has via TLR4, we used a cell line that expresses a secreted phosphatase reporter under the regulation of nuclear factor kappa-light-chain-
enhancer of activated B cells (NF-κB). The NF-κB assay was done in collaboration with Gyles Ifill from the Fernandez lab at UBC. HEK-Blue hTLR4™ were cultured and maintained according to the supplier's instructions. Cells were grown in standard cell culture flasks then seeded into 96-well plates with 5×10^4 cells per well. When the cells were established in the 96-well plates, their growth medium were replaced with fresh media containing LPS at concentrations of 0, 1, 10, and 100 µg ml⁻¹. After 24 hr the supernatant medium from all the wells were assayed for phosphatase activity. All experimental conditions were mirrored with cells not expressing TLR4, Null2 cells™. Fernandez lab at UBC. HEK-Blue hTLR4™ cells were cultured and maintained according to the supplier’s instructions. Cells were grown in standard cell culture flasks then seeded into 96-well plates with 5×10^4 cells per well. When the cells were established in the 96-well plates, their growth medium were replaced with fresh media containing LPS at concentrations of 0, 1, 10, and 100 µg ml⁻¹. After 24 hr the supernatant medium from all the wells were assayed for phosphatase activity. All experimental conditions were mirrored with cells not expressing TLR4, Null2 cells™.

3.2.17 THP-1 derived macrophage assay

To measure the cytokine stimulating ability of *C. crescentus* LPS, we treated macrophage cells, derived from a THP-1 cell line, with LPS from *C. crescentus* or *Salmonella enterica* var. Typhimurium.

The *C. crescentus* LPS was isolated from *C. crescentus* JS1013, a derivative of NA1000. The LPS isolation protocol can be found in section 3.2.2. The *S. enterica* LPS was *S. enterica* serotype minnesota Re 595 and it was purchased from Sigma-Aldrich.

The initial starting cultures of THP-1 cells were generous gifts from the Fernandez lab and the Finlay lab. For cell culture, we followed previously published protocols. We maintained two separate cell lines from the start from the two sources to ensure neither was non-viable or contaminated; both were found to have identical growth rates and both were clean of con-

\[^{161}\text{N. Marr et al. Infection and Immun, 78: 2060–2069, 2010.}\]
tamination. Upon the recommendation of Dr. Joris Van Der Heiden, the initial passage of the cells was in medium containing 20 µg ml\(^{-1}\) of kanamycin. The cells with kanamycin seemed to achieve maximal growth rate earlier than cells that did not have any antibiotic added. The Fernandez lab cell line was expanded and used for all subsequent assays. No antibiotics were added the medium other than during that first initial passage.

When the THP-1 cells were in their third passage and at a density of 1.4x10\(^6\) live cells ml\(^{-1}\) they were pelleted by centrifugation and resuspended in fresh medium containing 50 ng ml\(^{-1}\) phorbol 12-myristate 13-acetate (PMA). One hundred thousand THP-1 cells, in PMA, were placed into the wells of 24-well plates and allowed to differentiate. After 48 hr, all the cells were found to have become adherent to the well surfaces, the PMA-containing medium was removed, the cells were washed once with fresh medium, and then fresh medium was applied to all the wells and the cells allowed to rest. After a subsequent 48 hr, all the growth medium was replaced with medium containing LPS at concentrations of 0, 1, 10, 100 ng ml\(^{-1}\) of either \textit{C. crescentus} or \textit{S. enterica} origin. At the time points of 4 hr and 24 hr, culture supernates were removed and frozen at -80°C for cytokine assay. Biological replicates were done in quadruplicate. When the THP-1 cells were in their third passage and at a density of 1.4x10\(^6\) live cell per millilitre they were pelleted by centrifugation and resuspended in fresh medium containing 50 ng ml\(^{-1}\) PMA. 1x10\(^5\) THP-1 cells, in PMA, were placed into the wells of 24-well plates and allowed to differentiate. After 48 hr, all the cells were found to have become adherent to the well surfaces, the PMA-containing medium was removed, the cells were washed once with fresh medium, and then fresh medium was applied to all the wells and the cells allowed to rest. After a subsequent 48 hr, all the growth medium was replaced with medium containing LPS and concentrations of 0, 1, 10, 100 ng ml\(^{-1}\) of either \textit{C. crescentus} or \textit{S. enterica} origin. At the time points of 4 hr and 24 hr, well solutions were removed and frozen at -80°C for cytokine assay. Biological replicates were done in quadruplicate.

Kits for assaying various cytokines were purchased from Meso Scale Discovery: tumor necrosis factor \(\alpha\) (TNF-\(\alpha\), a proinflammatory cytokine released by macrophages in response
to LPS\textsuperscript{162}, interleukin (IL) 6, an important cytokine for generating adaptive immunity\textsuperscript{163}, IL 10, an anti inflammatory signal, “Master Regulator of Immunity to Infection”\textsuperscript{164}, and monocyte chemotactic protein 1 (MCP-1), a key chemokine for macrophage trafficking\textsuperscript{165}. The assays were done following the supplier’s instructions and the data readouts were produced on a QuickPlex SQ 120 imager (courtesy of the Timothy Kieffer’s lab, UBC).

### 3.2.18 Statistical Analysis

Statistical analyses were done with Microsoft Excel. An unpaired \textit{t}-test was used to compare the means between treatment groups at all time points or dosage levels. A \textit{P} value of <0.05 was taken as a statistically significant difference.

### 3.3 Results

#### 3.3.1 Characterization of whole LPS

The protocol we developed for isolating the LPS from \textit{C. crescentus} was unconventional compared to the majority of existing literature. Our EDTA chelation method was found to be a great fit for our needs and effective at producing LPS. The negative charges in Gram-negative bacterial LPS are stabilized and bridged by available divalent cations (\textit{i.e.} Ca\textsuperscript{2+} and Mg\textsuperscript{2+}). Lieve first noted in 1965\textsuperscript{77} that one can strip the counter ions from the outer membrane to disrupt it and cause it to shed into the medium. We chose a chelation strategy because we found it easy to scale up, it did not require noxious organic solvents, and because our lab’s first attempt to isolate and analyze \textit{C. crescentus} LPS\textsuperscript{148} using the Darveau extraction method\textsuperscript{166} only resulted in the isolation of core+lipidA (rough LPS).

Whole *C. crescentus* LPS was examined using a number of different techniques. Electrophoresis and staining was used as routine checks for sample quality. Multiple methods of staining were used to visualize the LPS. The traditional silver stain method of Tsai and Frasch\(^ {156}\) is very susceptible to errors and it is not convenient for day-to-day sample checks, so the periodic acid-Schiff stain was co opted from common histology protocols to visualize our carbohydrates. The Schiff reagent based stain was a very quick and reliable method that was used until the improved silver stain technique of Zhu et al. was published in 2012.\(^ {155}\) Zhu et al.’s protocol was superior to the traditional silver stain methods in our hands in both in terms of reliability and difficulty. Despite their established use, LPS silver stains may not be acting on the PS portion of the molecule as often though\(^ {167}\) but none the less silver stains were the primary tool used for checking sample quality and accuracy. Zinc-imidazole stain\(^ {168}\) was marginally successful at visualizing our LPS (data not shown) but the results were weak and poorly recordable. Zinc-imidazole stains were only used during attempts to purify LPS by electrophoresis. Figure 3.2 demonstrates how the LPS appears on an SDS-PAGE.

*C. crescentus* LPS runs as two separate bands; a high molecular weight smooth-LPS band and a low molecular weight rough-LPS (see fig. 3.2 E.). The *C. crescentus* smooth-LPS band runs as a single band, which is in contrast to the canonical laddering pattern of most other well studied smooth-LPS. The laddering pattern of LPS is from the heterogeneity in length that arises when OPS subunits are polymerized. Figure 3.3 shows *C. crescentus* LPS next to the LPS from *Escherichia coli* (*E. coli*) O122.

Figure 3.2 also demonstrates how our LPS samples appeared when analyzed by western blotting and far-western blotting using RsaA as the primary probe. Both western blotting and far-western blotting primarily identified the smooth-LPS band on the blot, indicating that the OPS is probably the antigenic portion of the LPS and the primary anchor molecule for RsaA and the *C. crescentus* S-layer. For the far-westerns, controls were run to test for LPS cross reactivity from the anti-RsaA antibodies and the anti-rabbit antibodies. None were detected.

Figure 3.2: Four different visualisations of *C. crescentus* LPS. A. Western blot with rabbit anti-OPS serum used as the primary probe. B. Far-western blot with RsaA Δ277–784 used as the primary probe and rabbit anti-RsaA serum used as the secondary probe. C. Schiff stained SDS-PAGE. D. Silver stained SDS-PAGE. The smooth-LPS (i.e. containing a complete OPS) runs as a single band at roughly the equivalent rate of a 50 kDa protein and is visible in all lanes. The fainter band visible at 30 kDa in the western blot (A) is of unknown source. The faint 30 kDa band is not seen in by far western (B), Schiff stain (C), or silver stain (D). E. A differently resolved silver stained SDS-PAGE. The high molecular weight smooth-LPS is present at 50 kDa and less well resolved band at around 10 kDa is present which is the low molecular weight rough-LPS.

We attempted to isolate the core oligosaccharide-OPS fraction from *C. crescentus* LPS prior to pursuing NMR analysis. Hydrolysis of the ketosidic bond between Kdo and lipid A (see fig. 1.8) was carried out with one of the classical protocols, 0.1 M HCl at 100°C for one hour. The reaction caused the expected cleavage and precipitation of lipid A, but the supernate had no detectable OPS. We used standard silver stain SDS-PAGE, western blot, spot-blot immunoblot, and far-western blot methods to detect the presence of free-OPS. None of the methods gave a positive result. Milder hydrolysis conditions, acetic acid or SDS accelerated, also produced no detectable free-OPS. Addressing the concern that OPS might not bind to nitrocellulose membrane, we tried to chemically link the polysaccharide to amine-coated surfaces using

Figure 3.3: Comparison of silver stained *E. coli* O122 LPS and *C. crescentus* LPS. LPS samples were run on a SDS-PAGE and silver stained. Lanes 1 and 3 contain equal amounts of LPS from *E. coli* O122. Lanes 2 and 4 contain our isolated LPS from *C. crescentus*. Lane 4 has twice the loaded amount compared to lane 2 to demonstrate that there are no minor, hidden bands. Notice the canonical ‘laddering’ pattern of *E. coli* LPS and the distinctly singular band of *C. crescentus* LPS.

carbodiimide chemistry\textsuperscript{170}, this method also failed to immobilize any detectable free-OPS.

We had no evidence from the literature to indicate that the OPS could be destroyed by the mild hydrolysis conditions. At the time the only clue about the structure of the OPS was the presence of N-acetylperosamine; perosamine containing OPSs have no known inherent instability.\textsuperscript{171,172} We were left to conclude that either the carbohydrates were in fact being degraded during hydrolysis or that the carbohydrates we were interested in were unable to migrate in gel electrophoresis, unable to bind to nitrocellulose, and unsuitable for chemical linkage. It was this curious situation, where procedures that should be routine consistently gave negative results, that led us to pursue a collaboration with Dr. Evgeny Vinogradov enabling a broader and hopefully more comprehensive analysis of the core and OPS regions of the LPS by NMR.

Attempts to accurately measure the molecular weight of our intact smooth-LPS were partly successful. Getting accurate MALDI-TOF results for large molecular weight polysaccharides is at best difficult and the technology is most often used for the analysis of smaller oligosaccharides (as below) and lipid A molecules. Figure 3.4 shows our best results. Our mass spectrum shows a tight clustering of m/z peaks at or near 10.8 kDa (assuming a single negative charge). Like our electrophoresis data, our MALDI-TOF data demonstrate that *C. crescentus* LPS is essentially homogenous in length. The small differences in m/z between the peaks in fig. 3.4 are around

180 Da, the size range of a single hexose molecule, not an entire OPS subunit.

Figure 3.4: MALDI-TOF analysis of intact, whole LPS from *C. crescentus*. The insert highlights the peaks attributed to LPS. The major LPS peak is labelled 10834.90 m/z

3.3.2 Identification of similar glycan structures by RsaA binding

The successful demonstration that RsaA Δ277–784 could bind *C. crescentus* OPS *in vitro*, led us to hypothesize that we could probe *in vitro* libraries of carbohydrates to identify the structure of our OPS. A service available at Consortium of Functional Glycomics (CFG) Core H at Emory
University is glycan array screening\textsuperscript{173,174}. Fluorescently labeled lectins can be screened over an microarray of 465 known glycan structures commonly found in mammals. The array can be used to identify targets of known carbohydrate-binding proteins, such as RsaA. We hoped that RsaA may bind a glycan on the array and that the structure of that carbohydrate may inform us on the structure of the OPS from \textit{C. crescentus}. Due to our initial component analysis, knew that our OPS differed significantly from any of the glycans in the array, at least compositionally (see below). RsaA (specifically RsaA $\Delta$277–784) was found to not measurably bind any of the glycan structures found on the array.

\subsection{Initial assessment and component analysis}

The PS was released from the LPS by hydrolysis with acetic acid. $^1$H NMR spectrum of the PS (see fig. \ref{fig:3.5}) contained a large number of partially overlapping signals of various intensities in the anomeric region. It was obviously not a regular polymer with well-defined repeating units. It was nevertheless clear that an intact polymer was released by hydrolysis. The free-OPS was detectable by NMR but not by the other methods that typically do allow monitoring of purification (see above, section \ref{sec:3.3.1}). Attempts to separate this material by anion-exchange chromatography led to the isolation of a number of fractions from neutral to slightly retained, but all of them had virtually identical NMR spectra. Methylation of the polysaccharide led to the identification of 3- and 3,4-substituted mannopyranose, terminal glucopyranose (derived from side-chain 3-O-MeGlc), terminal, 3-, 4-, and 2,4-substituted rhamnopyranose, 3-substituted PerNAc, and an unidentified derivative resembling methylated PerN that eluted between dimethylhexose derivatives and 3-substituted PerNAc. To identify the position of the methyl groups in naturally methylated monosaccharides, methylation was conducted with CD$_3$I. This confirmed the identification of tetramethylglucitol as originating from 3-O-MeGlc, but did not identify any other naturally methylated monosaccharides, visible in NMR spectra. An unknown derivative

received two deuterated methyl groups.

### 3.3.4 O-specific polysaccharide structure determination (PS1)

A set of 2D spectra \( \text{gCOSY, TCOSY, nuclear Overhauser enhancement spectroscopy (NOESY),} \) \(^1\text{H}-^{13}\text{C gHSQC, gHMBC} \) was obtained for the PS. There were many (more than 20) lines of correlations from the anomeric signals. Later, after the analysis of PS degradation products, most of them could be assigned to particular structures (figs. 3.7, 3.9, and 3.10). Polysaccharide heterogeneity was not caused by random acetylation, as was first hypothesized. Monosaccharide analysis revealed L-Rha, D-Man, D-PerN (perosamine, 4-aminodeoxyrhamnose), and 3-O-MeGlc. Other methylated monosaccharides were not identified by GC-MS as alditol acetates, possibly due to low content or degradation during hydrolysis. The GC-MS data is presented in fig. 3.6.

In an attempt to simplify the structure, PS was oxidized with NaIO₄, reduced with NaBD₄, hydrolysed with 2% AcOH, and the products were separated on a Biogel P6 column to give a polymer and an oligosaccharide, OS1. Analysis of OS1 will be described below. For some reason not all of the rhamnan was oxidized, and some of its signals persisted in the spectra of the remaining polymer (without side-chain Rha F). To remove the rest of it, the oxidation was repeated to produce PS1. Spectra still contained some signals of minor components that were analyzed later. Assignment of the spectra of the non-oxidizable polymer PS1 was difficult due to complete or partial overlap of the H-2,3,4,5 signals of PerNAc. To improve signal spread, PS1 was deacylated with 4 M NaOH. At this point the major polymer became positively charged and an attempt was made to separate it from the minor components using cation-exchange chromatography. However, all material was eluted together at high salt concentration, thus indicating that all components were chemically bound together. Assignment of the spectra (fig. 3.8, table 3.3) became possible at this stage due to better signal spread (H-4 signals of PerN moved to high field due to deacylation) and the sequence shown on fig. 3.7 was proposed. Spectra contained the signals of two \( \beta \)-mannopyranose, \( \alpha \)-3-O-MeGlc, and two -Per₄N. The
Figure 3.5: $^1$H NMR spectra of the intact C. crescentus OPS (bottom trace), double oxidized polysaccharide (middle trace) and N-deacylated double oxidized polysaccharide (upper trace).
Figure 3.6: Monosaccharide analysis of *C. crescentus* LPS. Alditol acetates derivatives of the LPS component monosaccharides were separated and identified by GC-MS. The protocol used is described on page 62.

Following interresidual nuclear Overhauser enhancement (NOE) and heteronuclear multiple bond coherence (HMBC) correlations were used to determine the sequence: R1:L3, L1:Z3, Z1:Q3, Q1:W3, W1:X3, A1:X4. PS1 had trisaccharide repeating units composed of β-mannose and two α-PerNAc residues, and every second repeating unit carried a side branch of 3-O-MeGlc. It seems that side-chains were present quite regularly at each second trisaccharide repeat of the main chain, because NOE correlations were observed between the repeating units with and without 3-O-MeGlc, and not between units of the same structure. Thus altogether, the repeating unit contained seven monosaccharides.
Figure 3.7: The structure of the *C. crescentus* OPS. A. The intact repeating unit, PS1. B. The deacylated product of PS1.
Figure 3.8: Overlap of COSY (green), TCOSY (red) and ROESY (black) correlations from anomeric protons of double oxidized deacylated *C. crescentus* PS1.

### 3.3.5 Minor component determination

PS and PS1 spectra contained signals of minor components, which could not be removed by chromatography, as described above. They probably represented the non-reducing ends of the major chain, PS1 (fig. 3.9). The minor components contained methylated Rha (2-O-Me-Rha residue J and 2,3-Me₂-PerN residue B). The position of the methyl groups were found from HMBC correlations between protons of methyl groups and carbon atoms bearing OMe groups, which all were well visible and did not overlap with other signals due to their low field position. Thus, two independent structural fragments, 1 and 2, were found and are shown in fig. 3.9. Mannose residues Z’ and Z” at the non-reducing ends of these fragments were further linked to PerN residues, indistinguishable from the PerN of the main chain. PerN residue D had upfield shifted C-2 and downfield shifted C-3 signals (table 3.4), which have not been explained. It appears that its O-3 was phosphorylated, producing typical phosphorylation signal shifts and
**Table 3.3:** NMR data for *C. crescentus* PS1 (40°C) and deacylated PS1 (50°C). Me at 3.62/61.3 ppm.

<table>
<thead>
<tr>
<th></th>
<th>PS1</th>
<th><strong>Deacylated PS1</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>PerNAc R</td>
<td>H</td>
<td>4.97</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>103.3</td>
</tr>
<tr>
<td>PerNAc Q</td>
<td>H</td>
<td>4.98</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>103.3</td>
</tr>
<tr>
<td>PerNAc L</td>
<td>H</td>
<td>5.05</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>103.9</td>
</tr>
<tr>
<td>PerNAc W</td>
<td>H</td>
<td>5.02</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>104.0</td>
</tr>
<tr>
<td>(\beta)-Man X</td>
<td>H</td>
<td>4.74</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>97.8</td>
</tr>
<tr>
<td>(\beta)-Man Z</td>
<td>H</td>
<td>4.72</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>98.2</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.0</td>
</tr>
</tbody>
</table>

77
broadening of the H-3 signal, but \(^1\)H-\(^{31}\)P HMQC NMR spectrum showed no signals, possibly due to the low abundance of this residue. Possibly Rha residues inserted in the structure resembling PS1 represented the attachment point of the rhamnan (PS2) to PS1, if they were linked together.

![Diagram A](image1.png)

\(\alpha\)-PerNAc-p-3-\(\alpha\)-Man-p-3-\(\alpha\)-Man-p-3-3-\(\alpha\)-Rha2Mep-p-3-\(\beta\)-Rhap-p-3-\(\beta\)-Man-p-PS

![Diagram B](image2.png)

\(\alpha\)-PerNAc2,3Mep-p-3-\(\beta\)-Man-p-PS

**Figure 3.9:** The structure of minor component, the end caps of the OPS. **A.** Fragment 1. **B.** Fragment 2.

### 3.3.6 Rhamnan polysaccharide determination (PS2)

Periodate oxidation of the PS produced an OS1, which was analyzed by NMR and its structure, as shown on fig. 3.10 was determined using standard 2D NMR methods. Signal assignment is shown in the table 3.5. It contained three rhamnopyranose units and 4-deoxy-1-deutero-erythritol, produced by the oxidation-reduction of 4-substituted rhamnose. Formation of this oligosaccharide could be explained by oxidation of the side chain Rha F and 4-substituted Rha G in the PS1 (letter labels for monosaccharides were given using anomeric signals in the whole PS spectra starting from low-field). The unoxidized 4-substituted residue T in the oligosaccharide originally carried side-chain Rha F at position 2. Knowing the OS1 structure the signals of a corresponding polymer (PS2) were identified in the spectra of the whole PS and are given in the
Table 3.4: NMR data for the minor components of the double oxidized non-deacylated PS (50°C). Methyl group signals: B2: 3.48/59.5; B3: 3.42/57.9; J2: 3.45/59.6 ppm (H/C)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>PerN2,3Me B</td>
<td>H</td>
<td>5.25</td>
<td>3.96</td>
<td>3.71</td>
<td>3.81</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.2</td>
<td>76.2</td>
<td>78.2</td>
<td>53.2</td>
<td>69.5</td>
</tr>
<tr>
<td>α-Rha2Me J</td>
<td>H</td>
<td>5.25</td>
<td>3.95</td>
<td>4.02</td>
<td>3.45</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.2</td>
<td>75.9</td>
<td>75.4</td>
<td>71.8</td>
<td>69.5</td>
</tr>
<tr>
<td>α-PerN D</td>
<td>H</td>
<td>5.14</td>
<td>4.27</td>
<td>4.20</td>
<td>3.95</td>
<td>3.97</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.1</td>
<td>67.4</td>
<td>75.2</td>
<td>52.2</td>
<td>69.4</td>
</tr>
<tr>
<td>α-Man E</td>
<td>H</td>
<td>5.13</td>
<td>4.17</td>
<td>3.99</td>
<td>3.75</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.1</td>
<td>70.9</td>
<td>79.8</td>
<td>67.1</td>
<td>74.4</td>
</tr>
<tr>
<td>α-Man P</td>
<td>H</td>
<td>5.05</td>
<td>4.14</td>
<td>4.00</td>
<td>3.85</td>
<td>3.96</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>97.6</td>
<td>71.0</td>
<td>79.5</td>
<td>66.9</td>
<td>74.0</td>
</tr>
<tr>
<td>β-Rha U</td>
<td>H</td>
<td>4.78</td>
<td>4.16</td>
<td>3.69</td>
<td>3.48</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>100.9</td>
<td>71.8</td>
<td>82.0</td>
<td>72.4</td>
<td>73.1</td>
</tr>
<tr>
<td>β-Man Z''</td>
<td>H</td>
<td>4.72</td>
<td>4.06</td>
<td>3.75</td>
<td>3.75</td>
<td>3.46</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>98.2</td>
<td>71.9</td>
<td>86.0</td>
<td>67.0</td>
<td>77.4</td>
</tr>
<tr>
<td>β-Man Z'</td>
<td>H</td>
<td>4.72</td>
<td>4.06</td>
<td>3.73</td>
<td>3.75</td>
<td>3.46</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>98.2</td>
<td>71.9</td>
<td>82.3</td>
<td>67.0</td>
<td>77.4</td>
</tr>
</tbody>
</table>

3.3.7 Core oligosaccharide determination

The core oligosaccharide of the *C. crescentus* LPS isolated after AcOH hydrolysis contained one non-degraded Kdo, two LD-Hep, one DD-Hep, mannose, galactose, and glucuronic acid in pyranose form. 2D NMR analysis led to the structure shown on fig. 3.11 (NMR assignments are...
Figure 3.10: The structure of the *C. crescentus* rhamnan. A. the intact rhamnan, PS2. B. the oxidized rhamnan product, OS1.

In table 3.6, the HSQC spectrum is in fig. 3.12. The sequence followed from the observed NOE: E1:C5,C7,F5; F1:E2; G1:F3; H1:C7,E2; K1:C4; L1:K4. Correlation E1:C7 is always observed in the α-Hep-5-Kdo fragment. E1:F5 was due to the α-Man-2-Hep linkage. H1E2 indicates spatial proximity of the residues E and H, linked to the same Kdo C. All expected transglycoside correlations were observed in the HMBC spectrum, together with intra-ring correlations H-1:C-3 and H-1:C-5 for all α-pyranoses. Methylation analysis revealed terminal DD-Hep, terminal and 2-substituted LD-Hep, 3-substituted Man and terminal Gal. The structure agreed with mass spectral data, electrospray ionization (ESI) mass spectrometry (MS) in negative ion mode, [M-H]⁻ = 1314.9, [M-2H]/2⁻ = 656.7, calculated exact mass Hex x 2 + Hep x 3 + HexA x 1 + Kdo x 1 = 1314.4 Da.
Table 3.5: NMR data for *C. crescentus* PS2 and its NaIO$_4$ oxidation product OS1 (40°C).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)-Rha N, OS1</td>
<td>H 5.04</td>
<td>4.07</td>
<td>3.85</td>
<td>3.47</td>
<td>3.85</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>C 103.1</td>
<td>71.1</td>
<td>71.1</td>
<td>72.9</td>
<td>69.9</td>
<td>17.6</td>
</tr>
<tr>
<td>(\alpha)-Rha N, PS</td>
<td>H 5.02</td>
<td>4.02</td>
<td>3.79</td>
<td>3.45</td>
<td>3.72</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>C 103.0</td>
<td>71.5</td>
<td>78.8</td>
<td>73.3</td>
<td>70.4</td>
<td>17.5</td>
</tr>
<tr>
<td>(\beta)-Rha T, OS1</td>
<td>H 4.78</td>
<td>4.13</td>
<td>3.71</td>
<td>3.54</td>
<td>3.53</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>C 100.7</td>
<td>71.1</td>
<td>72.3</td>
<td>83.4</td>
<td>71.7</td>
<td>17.5</td>
</tr>
<tr>
<td>(\beta)-Rha T, PS</td>
<td>H 4.79</td>
<td>4.22</td>
<td>3.82</td>
<td>3.62</td>
<td>3.61</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>C 101.7</td>
<td>76.7</td>
<td>73.5</td>
<td>83.4</td>
<td>73.2</td>
<td>17.5</td>
</tr>
<tr>
<td>(\beta)-Rha V, OS1</td>
<td>H 4.75</td>
<td>4.12</td>
<td>3.67</td>
<td>3.49</td>
<td>3.49</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>C 101.1</td>
<td>71.4</td>
<td>81.3</td>
<td>71.9</td>
<td>72.9</td>
<td>17.5</td>
</tr>
<tr>
<td>(\beta)-Rha V, PS</td>
<td>H 4.77</td>
<td>4.13</td>
<td>3.66</td>
<td>3.66</td>
<td>3.50</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>C 101.4</td>
<td>71.2</td>
<td>81.7</td>
<td>67.2</td>
<td>73.3</td>
<td>17.5</td>
</tr>
<tr>
<td>X (ox. G), OS1</td>
<td>H 3.69; 3.71</td>
<td>3.75</td>
<td>3.99</td>
<td>1.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C 61.8</td>
<td>84.8</td>
<td>67.7</td>
<td>18.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\alpha)-Rha G, PS</td>
<td>H 5.10</td>
<td>4.13</td>
<td>3.94</td>
<td>3.58</td>
<td>3.90</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td>C 103.1</td>
<td>71.5</td>
<td>70.3</td>
<td>84.5</td>
<td>17.5</td>
<td></td>
</tr>
<tr>
<td>(\alpha)-Rha F, PS</td>
<td>H 5.11</td>
<td>4.09</td>
<td>3.87</td>
<td>3.45</td>
<td>4.07</td>
<td>1.27</td>
</tr>
<tr>
<td></td>
<td>C 102.4</td>
<td>72.1</td>
<td>71.3</td>
<td>73.3</td>
<td>17.5</td>
<td></td>
</tr>
</tbody>
</table>

3.3.8 Cellular response to *C. crescentus* LPS

We analyzed the immune/cellular response against *C. crescentus* LPS by two methods, assaying NF-\(\kappa\)B activation in HEK293 cells and measuring cytokine production levels from THP-1 derived macrophages. Figure 3.13 presents our data from the NF-\(\kappa\)B assays. When compared to *S. enterica* LPS, *C. crescentus* LPS is much less stimulatory. We found a significantly lower
level of NF-κB activation by *C. crescentus* LPS at 1, 10, and 100 ng ml$^{-1}$. 

The cytokine production assays were a more in depth analysis of the profile of clear-cut as the NF-κB data. We hypothesized in light of the NF-κB experiment and past work with lipid A, that the *C. crescentus* LPS would significantly stimulate less cytokine production than enteric LPS. Figure 3.14 compiles the data from all the cytokine assays. At 4 hr after treating the cell with only 1 ng ml$^{-1}$ of LPS added, all the cells that were treated with *C. crescentus* LPS had notably lower production levels of all cytokines tested than the cells that were treated with *S. enterica* LPS. At 24 hr, and also with 1 ng ml$^{-1}$ of LPS added, *C. crescentus* LPS stimulated less cytokine production than *S. enterica* LPS for all cytokines other than MCP-1. At the 10
Table 3.6: NMR data for the core oligosaccharide (25°C).

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>C</th>
<th>H</th>
<th>C</th>
<th>H</th>
<th>C</th>
<th>H</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdo C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.05; 2.35</td>
<td>4.23</td>
<td>4.28</td>
<td>4.09</td>
<td>3.91</td>
<td>3.81; 3.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>98.2</td>
<td>34.7</td>
<td>74.9</td>
<td>74.2</td>
<td>71.0</td>
<td>79.2</td>
<td>61.5</td>
<td></td>
</tr>
<tr>
<td>DLHep E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.24</td>
<td>3.85</td>
<td>4.12</td>
<td>3.89</td>
<td>3.88</td>
<td>4.04</td>
<td>3.73; 3.73</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>101.1</td>
<td>80.2</td>
<td>71.1</td>
<td>67.3</td>
<td>74.1</td>
<td>70.0</td>
<td>63.9</td>
<td></td>
</tr>
<tr>
<td>Man F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.03</td>
<td>4.30</td>
<td>4.00</td>
<td>3.79</td>
<td>3.80</td>
<td>3.80; 3.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>103.7</td>
<td>70.6</td>
<td>78.1</td>
<td>67.5</td>
<td>74.5</td>
<td>62.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDHep G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.15</td>
<td>4.05</td>
<td>3.86</td>
<td>3.77</td>
<td>3.84</td>
<td>4.06</td>
<td>3.76; 3.84</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>103.1</td>
<td>71.0</td>
<td>71.8</td>
<td>68.7</td>
<td>74.5</td>
<td>72.9</td>
<td>62.9</td>
<td></td>
</tr>
<tr>
<td>DLHep H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.07</td>
<td>4.14</td>
<td>3.89</td>
<td>3.89</td>
<td>3.80</td>
<td>4.10</td>
<td>3.79; 3.81</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>102.7</td>
<td>73.7</td>
<td>71.8</td>
<td>67.3</td>
<td>73.2</td>
<td>70.2</td>
<td>64.9</td>
<td></td>
</tr>
<tr>
<td>GlcA K</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.22</td>
<td>3.66</td>
<td>4.00</td>
<td>3.99</td>
<td>4.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>98.8</td>
<td>72.4</td>
<td>74.9</td>
<td>73.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gal L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.49</td>
<td>3.80</td>
<td>3.85</td>
<td>3.99</td>
<td>3.94</td>
<td>3.69; 3.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>99.9</td>
<td>69.7</td>
<td>70.3</td>
<td>70.0</td>
<td>71.8</td>
<td>61.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and 100 ng ml⁻¹ LPS treatment levels, the differences between C. crescentus and S. enterica become less clear. At the 100 ng ml⁻¹ LPS treatment conditions C. crescentus LPS actually stimulated more cytokine production in almost every test (except MCP-1, 4 hr). clear-cut as the NF-κB data. We hypothesized in light of the NF-κB experiment and past work with lipid A, that the C. crescentus LPS would significantly stimulate less cytokine production than enteric LPS. Figure 3.14 compiles the data from all the cytokine assays. At four hr after treating the cell with only 1 ng ml⁻¹ of LPS added, all the cells that were treated with C. crescentus LPS had notably lower production levels of all cytokines tested than the cells that were treated with S. enterica LPS. At 24 hr, and also with 1 ng ml⁻¹ of LPS added, C. crescentus LPS stimulated less cytokine production than S. enterica LPS for all cytokines other than MCP-1. At the 10 and 100
Figure 3.12: Fragment of $^1$H-$^{13}$C HSQC spectrum of the core.

ng ml$^{-1}$ LPS treatment levels, the differences between C. crescentus and S. enterica become less clear. At the 100 ng ml$^{-1}$ LPS treatment conditions C. crescentus LPS actually stimulated more cytokine production in almost every test (except MCP-1, 4 hr).

3.4 Discussion

The LPS of C. crescentus has an unusually complicated structure with two different polysaccharides, irregular substituents, resulting in NMR spectra that were especially difficult to analyze. Presented data show structures of the core, two polymers, and putative terminal structures. The polysaccharides could not be separated by size exclusion or anion-exchange chromatography
Figure 3.13: The activation of the NF-κB promoter as measured by the HEK-Blue hTLR4™ cell line (InVivoGen) treated with LPS from either S. enterica or C. crescentus. Error bars signify ± standard deviation, n=3 for all groups, p-values less than 0.05 are noted.

and so are probably linked together through the same core. The core of the C. crescentus LPS has been studied previously and an initial assessment of its composition was made, but the complete structure had not been determined. The structure of the OPS had not been studied before. In our view the polysaccharide structure of the C. crescentus LPS represents one of the most complicated bacterial LPS polysaccharide structures identified so far.

An early goal in working with the C. crescentus LPS was to isolate the S-layer anchoring portion of the molecule and use it to assemble the S-layer onto an abiotic surface. The hypothesis we were following was that a simple mild-acid hydrolysis of the LPS would produce a free-OPS that could be chemically linked to aminated surfaces by virtue of a free carboxyl on the Kdo residue exposed by hydrolysis. We were unable to demonstrate the presence of a free-OPS after hydrolysis using biochemical methods, such as SDS-PAGE and western immunoblotting. Fortunately, NMR analysis following mild acid hydrolysis did demonstrate an intact OPS structure.
Figure 3.14: Cytokine production levels produced by THP-1 derived macrophages in response to treatment with LPS from *C. crescentus* or *S. enterica*. All data points are present, n=4 for all groups, *p*-values less than 0.05 are marked with a star.
None of the determined carbohydrate structures give clues as to why we were unable to detect free-OPS by spot-blot immunoblot (see section [3.3.1]). We do not have much information about which particular structures in the C. crescentus LPS are bound by RsaA or the antibodies used in western blotting. Genetic knock outs point towards the requirement of N-acetylperosamine, but now that the complete structure has been determined it is not clear why the N-acetylperosamine containing structures (OPS and end caps) were not previously detectable.

The Kdo present in the LPS core structure (see fig. [3.11]) has the typical substitutions at C-4 and C-5 of a manno-configured sugar and a negatively charged sugar, respectively.\[102\]\[175\] It also has a rarely observed third substitution at C-7 with a heptose moiety. The Kdo C-7 position is known to be occupied by a galactose moiety in the core of Rhizobium leguminosarum bv. Viciae VF39,\[176\] and the secondary Kdo in the core oligosaccharide from Acinetobacter baumannii ATCC 19606 has an C-7 substituted with a glucosamine.\[177\] This arrangement of a triple substituted Kdo is extremely rare in known structures. The core oligosaccharide of Vibrio cholerae O139 has triple substituted Kdo with the core oligosaccharide elongating off the C-5 position, a phosphate on the C-4 position, and a phosphoethanolamine on the C-7 position.\[107\] Our core structure is, as far we are aware, the only core oligosaccharide with a triple substituted Kdo with sugars on all three sites.

In the traditional model LPS occupies the outer leaflet of the outer membrane of a Gram-negative bacterium, and so (excepting the presence of cell associated EPS) is the outermost layer of the cell. For C. crescentus, however, LPS is the penultimate barrier below the protein S-layer. The C. crescentus OPS serves as the anchor for the S-layer and is likely not accessible to the environment.\[128\] The carbohydrates found in the OPS are particularly hydrophobic, marked by an abundance of deoxy-sugars, acetyl groups, and methyl groups. This hydrophobicity is possibly a result of particular sugars needed for S-layer anchoring, as these carbohydrate structures likely

---

evolved as the cognate ligands for the S-layer protein, RsaA. The distance between the S-layer and the outer membrane is about 17–19 nm.\textsuperscript{178} It is possible the hydrophobicity aids in packing the polysaccharides between the S-layer and the LPS. Further determination of RsaA’s structure should help illuminate the interaction between the S-layer and OPS.

Knowledge of the structure of \textit{C. crescentus} OPS and LPS will facilitate the determination and characterization of their biosynthetic enzymes and mutant variants. Already, the enzymes Lpx\textsuperscript{179} and GDP-L-perosamine acetylase\textsuperscript{180} from \textit{C. crescentus} have been characterized. One uncharacterized enzyme, WbqL, is necessary for proper OPS synthesis. Disruption of \textit{wbqL} leads to the accumulation of aberrant OPS precursors in the inner membrane which inhibits Crescentin-mediated cell curvature\textsuperscript{181} Many genes, such as \textit{wbqL}, have been identified as essential for OPS synthesis\textsuperscript{147} but most if not all have not yet been characterized. Other genes, that must be essential for OPS synthesis, have yet to be identified or characterized, such as the OPS polymerase and ligase.

The subunit-based repeating nature and branching structure of the \textit{C. crescentus} OPS suggests that a Wzy-dependent pathway synthesizes the polymer.\textsuperscript{69} The previous study that aimed to identify genes essential for OPS did not identify many of the canonical genes in the Wzy-dependent pathway\textsuperscript{147} such as the O-unit transporter, Wzx, OPS polymerase, Wzy; the chain-length determinate protein, Wzz; and the OPS ligase, WaaL. Genes that have been annotated as putative OPS synthesis genes do appear in the sequenced genomes for \textit{C. crescentus} CB15, but they have not been experimentally confirmed.

An additional aspect to this LPS is the fact that its OPS is of homogeneous length. While other LPSS vary in size due to the number of OPS repeat groups, appearing as a laddering of bands by SDS–PAGE, the LPS from \textit{C. crescentus} appears as a single band.\textsuperscript{128} MALDI-TOF analysis of the entire LPS indicates a size of about 10.8 kDa (see fig. 3.4). After accounting for the

solved structures for the lipid A and core regions, this suggests the LPS contains approximately 5 repeats of the proposed heptameric OPS structure. There is not currently a known mechanism for the regulation and synthesis of a strictly homogeneous length OPS. Another example of a homogenous length OPS is *Aeromonas hydrophila* AH-1.\(^{182}\) The *Aeromonas* OPS is formed by an ATP binding cassette (ABC) transporter-dependent pathway. The homogenous length and unique non-reducing end caps suggest that the *C. crescentus* OPS is synthesized via an ABC-transporter-dependent pathway\(^{69}\), but there may also be another heretofore undiscovered mechanism.

Previously, our lab showed that alone the *C. crescentus* lipid A fraction of the LPS stimulates less than a hundredth of the TNF-α production than *E. coli* D31m4 rough-LPS\(^{86}\). We hypothesized that intact *C. crescentus* LPS would have a similar effect because lipid A is generally considered the innate immune stimulating portion of the LPS molecule. Our first immune cell-based assay used a cell line constructed with a TLR4–NF-κB reporter system. The NF-κB results clearly support the previous reports that *C. crescentus* LPS is less inflammatory than enteric LPS. Expanding our investigation into cytokine production, using THP-1 derived macrophages, produced some results that did not entirely fit with our hypotheses. Figure 3.14 shows that in some circumstances *C. crescentus* LPS causes the production of less cytokines than *S. enterica* LPS (for IL 6, IL 10, and TNF-α) but at highest doses it stimulates more, at least in this assay. Perhaps the data show that the stimulatory effect of *C. crescentus* LPS is not always less than its enteric counterparts, but there are enough examples to the contrary that weaknesses in the experimental design may be at fault. Because TLR4 is the major sensor for LPS in both our assays and because NF-κB activation is a major contributor to many cytokine production levels,\(^{183}\) our two assays are not completely independent of one another. If they completely contradict each other, then one of them is wrong or there are more complexities involved than are obvious.

The conditions in our assays that contradict our hypotheses were most often at the high

---


LPS concentrations (100 ng ml$^{-1}$). The *C. crescentus* LPS, at higher concentrations, may be activating alternate receptors than just TLR4. If it is not the LPS stimulating in a non-TLR4 manner, then something else is. One difference in our assays is the cells used. The THP-1 cells express a full complement of toll-like receptors, in addition to TLR4. If our purified *C. crescentus* LPS contained any microbe-associated molecular pattern (MAMP) contaminants (e.g., flagellin, peptidoglycan) then they may have stimulated THP-1 cells in addition to the signal from the LPS. These possible contaminants would not have affected the results from the HEK-Blue hTLR4 cells because they only expressed the LPS sensor, TLR4. Possibly, the commercially sourced *S. enterica* LPS was depleted of any contaminants by the manufacturer. There were no contaminants in the *C. crescentus* LPS that were abundant enough to affect (or be detected by) the NMR or MS analysis, but it is possible that the cell assays are more sensitive and it is possible any contaminants present were not in a detectable form with those other methods. If contamination is the culprit of our surprising cytokine results, then repeating the experiments with a cell line only expressing TLR4 should show what we expect: *C. crescentus* LPS is less stimulatory than enteric LPS.
Chapter 4

OmpW of Caulobacter crescentus functions as an outer membrane channel

“Now in the 21st century, the boundaries separating chemistry, physics, and medicine have become blurred, and as happened during the Renaissance, scientists are following their curiosities even when they run beyond the formal limits of their training”
— Peter Agre, 2003 Nobel Prize Banquet Speech. Ion channel researcher

4.1 Introduction

Gram-negative bacterial cell envelopes consist of different layers or membranes. The inner, or cytoplasmic, membrane contains the respiration chain components, proteins for the transport of nutrients, and proteins involved in the synthesis and transport of phospholipids, peptidoglycan, and lipopolysaccharide (LPS)\textsuperscript{184,185} The periplasm contains the peptidoglycan and a large number of proteins. The outer membrane is composed of protein, lipid and LPS\textsuperscript{184} It typically contains many fewer proteins than the cytoplasmic membrane. Normally at least one of the constitutive outer membrane proteins is a porin, a general diffusion pore with a defined exclusion limit for hydrophilic solutes.\textsuperscript{115} In addition to constitutive porins, an outer membrane

may contain porins that are induced under special growth conditions. They often form solute-specific channels and contain binding sites for neutral substrates such as carbohydrates or nucleosides, and phosphate. Many of the specific porins are part of uptake and degradation systems, such as the maltose uptake system of *Escherichia coli*.

*Caulobacter crescentus*, an alphaproteobacterium found in oligotrophic aquatic environments, is an unusual Gram-negative bacterium in that genes coding for typical general diffusion porins of the OmpF/C type of enteric Gram-negative bacteria have not been identified in its genome. Similarly, genes coding for specific porins such as Tsx or LamB are also absent. To date, no definitive porin has been demonstrated in *C. crescentus*. Instead, the genome of *C. crescentus* contains a large number of genes that code for TonB-dependent receptors. More than 60 of these receptors have been identified, which probably means that most of the nutrients from dilute environments are taken up actively by these systems. Examples for this are the uptake of maltose and N-acetylglucosamine (GlcNAc) into the cells, as opposed to the passive mechanisms of porins.

In this study we investigated whether the outer membrane of *C. crescentus* also contained a porin-like channel. The results suggest that a porin-like channel could be detected in artificial membranes with a single-channel conductance of about 125 pSv in 1 M KCl. The protein responsible for channel formation has been identified to be a member of the large OmpW family of outer membrane proteins. OmpW analogue proteins are found in many Gram-negative bacteria. Two members of this family, OmpW of *E. coli* and OprG of *Pseudomonas aeruginosa* have been crystallized and their 3D-structures are known at high resolution (2.7 and 2.4 Å.)
respectively).\textsuperscript{195,197} OmpW of \textit{C. crescentus} functions as a channel for cations, which is in sharp contrast to OprG of \textit{P. aeruginosa} and OmpW of \textit{E. coli}, which are believed to be plugged completely or involved in the transport of small, yet unknown hydrophobic molecules across the cell wall of these bacteria.\textsuperscript{195,197} The 3D structure of OmpW of \textit{C. crescentus} was modeled on the basis of the known structures of OmpW of \textit{E. coli} and OprG of \textit{P. aeruginosa}.\textsuperscript{196,197} The results indicated that OmpW of \textit{C. crescentus} could have a larger diameter and more hydrophilic interior than the two crystallized members of the OmpW family.

4.2 Methods

4.2.1 Growth and maintenance of bacteria

\textit{C. crescentus} NA1000 353Φ (JS1013), a strain that carries an amber mutation in the gene rsaA resulting in S-layer deficiency,\textsuperscript{132} was used for initial identification and characterization of OmpW. Wildtype \textit{C. crescentus} NA1000 was used for generating an \textit{ompW}-knockout strain and as a comparison against the knockout strain. The bacteria were grown to mid log phase (optical density at 600 nm (OD\textsubscript{600}) ≈ 0.8) in peptone-yeast extract medium (PYE)\textsuperscript{198} at 30°C in 5 ml cultures, which were used to start large cultures in 2.8 l Fernbach flasks containing 1250 ml M16HIGG medium, shaken at 100 rpm. M16HIGG is a modification of M6HIGG medium,\textsuperscript{12} containing 0.31% glucose, 0.09% glutamate, 1.25 mM sodium phosphate, 3.1 mM imidazole, 0.05% ammonium chloride and 0.5% modified Hutner’s Mineral Base.

4.2.2 Generation of *C. crescentus* ∆*ompW*

The gene *ompW*, CCNA_01475 (CC_1409), which encodes for OmpW, was knocked out in wildtype *C. crescentus* NA1000 via a two step method, derived from previously published protocols\footnote{M. C. Toporowski et al. *J Bacteriol*, 186: 8000–8009, 2004.}, resulting in a dysfunctional copy of *ompW* with a large internal deletion. The amino acid sequence for OmpW can be found in the appendix on page\footnote{A. Gilchrist and J. Smit. *J Bacteriol*, 173: 921–925, 1991.} 135.

The suicide plasmid ‘pKMOBsacB-*ompW*-A/B’ was constructed by combining two fragments of DNA homologous to the 5′ and 3′ ends of the gene *ompW* into a suicide vector containing both positive and negative selection elements. The ‘A-fragment’ was polymerase chain reaction (PCR) amplified using the primers F-ompW-A (5′–TAC CGG AAT TCT CGG GCG CTG GGC CTG TCT GTT GAG −3′) and R-ompW-A (5′–GTC CCA AGC TTG CGG AAG ATC TAT TGG CGC CGG CGG CAG TCA GGA TG -3′), resulting in a 994 bp product with a 5′ *EcoRI* cleavage site and 3′ *BglII* and *HinDIII* cleavage sites. The ‘A-fragment’ PCR product was blunt ligated into pBSK-ESH\footnote{pBSK-ESH}, resulting in pBSK-*ompW*-A. The ‘B-fragment’ PCR product was amplified using the primers F-ompW-B (5′–CGG GAT CCA CGT CAA GAA GGT CTA TTT CAG CAC −3′) and R-ompW-B (5′–GTC CCA AGC TTG CGT CGA TGC TAG TGC GCT GCG ATG −3′), resulting in a 1090 bp product with a 5′ *BamHI* cleavage site and a 3′ *HinDIII* cleavage site. The ‘B-fragment’ PCR product was similarly blunt subcloned into pBSK-ESH, then excised by *BamHI* and *HinDIII* digestion, and ligated into pBSK-*ompW*-A, digested with *BglII* and *HinDIII*, resulting in the plasmid pBSK-*ompW*-A/B. The plasmids pBSK-*ompW*-A/B and pKMOBsacB were digested with *EcoRI* and *HinDIII*, the 2070 bp fragment from pBSK-*ompW*-A/B was ligated into the pKMOBsacB fragment creating the plasmid pKMOBsacB-*ompW*-A/B.

The plasmid pKMOBsacB-*ompW*-A/B was electroporated into *C. crescentus* NA1000 cells\footnote{A. Gilchrist and J. Smit. *J Bacteriol*, 173: 921–925, 1991.}. The resulting transformants that were kanamycin resistant were re-passaged through kanamycin-free PYE medium three times and plated on kanamycin-free PYE agar plates that contain 3%...
sucrose. All colonies that were found to be sucrose resistant were screened for kanamycin sensitivity to confirm that the pKMOB sacB-ompW-A/B plasmid had crossed out of the genome. Colonies that were both sucrose resistant and kanamycin sensitive had their ompW genes PCR amplified using the primers F-ompW (5'-CGC ACT GGG CTT GCT GGC CTT TTT C -3') and R-ompW (5'-GGA GCC AGA GGA CGG ACG ACC GGG G -3'); intact ompW resulted in a roughly 800 bp product, and knocked-out ompW resulted in a roughly 500 bp product.

4.2.3 Outer membrane-enriched preparations

The protocol for preparing outer membrane extracts enriched in OmpW was derived from our LPS isolation protocol (on page 56). The crucial difference here is the omission of proteinase K. The lack of a nuclease step was primarily for convenience, but also it was deemed unecessary. Nucleic acids can be significant contaminants in nuclear magnetic resonance spectroscopy (NMR) analysis, even in trace quantities, but they are not a serious concern in the these bilayer experiments or sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analyses.

Cells were pelleted by centrifugation at 12,400 x g for 10 min. Cell pellets were washed by suspension in distilled water and repelleted. The pellets were resuspended in 1/10 original culture volume of phosphate-buffered saline (PBS) amended with 10 mM ethylenediaminetetraacetic acid (EDTA), agitated at room temperature for 5 min and then centrifuged at 15,300 x g for 15 min. The supernatant was retrieved and re-centrifuged to clarify. The supernatant was then ultracentrifuged at 184,000 x g for 2 h. Glassy pellets formed which were suspended in 1/100 original culture volume in 10 mM Tris pH 8.0. This treatment led preferentially to the disruption of the outer membrane and periplasmic contents were released without significantly releasing cytoplasmic contents.

4.2.4 Crude membrane preparations

For comparison to the PBS-EDTA membrane enrichment method, crude membrane preparations were prepared from 5 ml of mid logarithmic culture. The culture was sonicated at 50% intensity for 5 x 30 sec bursts. DNAse and RNAse were added to final concentrations of 0.06 mg ml\(^{-1}\) and 0.60 mg ml\(^{-1}\) respectively, and incubated at 37\(^{\circ}\)C for 1 h. The preparation was then ultracentrifuged for 2 h at 107 000 x g. A glassy pellet formed which was resuspended in 200 µl of distilled water.

4.2.5 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

Electrophoretic analyses of protein samples were done by the standard methods of discontinuous SDS-PAGE\(^{154}\) in a similar fashion to the methods used to analyze LPS in chapter 3 on page 57. Initially, samples were boiled in sample buffer for 5 min prior to running on the gel. When we wanted to probe our samples for heat-modifiabilitiy, samples were not boiled; to melt the sample buffer, it was heated to 37\(^{\circ}\)C prior to mixing with the protein samples. Gels were stained to visualize protein with Coomassie brilliant blue as previously described\(^{199}\).

4.2.6 Isolation and purification of the channel-forming protein from outer membrane-enriched preparations

The outer membrane-enriched samples were prepared by PBS-EDTA extraction were inspected for channel-forming activity by treatment with the detergent lauryldimethylamine-oxide (LDAO). The detergent extracts of the enriched OM showed rapid channel formation in the lipid bilayer assay. The protein responsible for channel formation was identified by preparative SDS-PAGE. Highest channel-forming activity was observed in the molecular mass range between 20 and 25 kDa.


Analytical and preparative SDS-PAGE was performed according to Laemmli. The gels were stained with Coomassie brilliant blue or Colloidal Coomassie blue.

4.2.7 Tryptic digestion and peptide sequencing

The pure 22 kDa protein eluted from preparative SDS-PAGE was subjected to amino acid sequence analysis. Direct sequencing was not possible presumably because of blocking of the N-terminus. The 22 kDa protein was then cleaved with trypsin as described. The peptides were separated by reverse phase HPLC on a Purospher RP18 encapped 5 µm column (Merck, Darmstadt, Germany) using a solvent gradient from 0 to 60% acetonitrile in 0.1% trifluoroacetic acid/water (v/v). The flow rate was 60 µl min⁻¹ and UV-detection was performed at 206 nm. The amino acid sequence analysis of the tryptic peptides was performed using an ABI 472A protein sequencer (Applied Biosystems, Langen, Germany).

4.2.8 Lipid bilayer experiments

The method used for the reconstitution experiments using black lipid bilayer membranes has been described previously. The membranes were formed from a 1% (w/v) solution of diphytanoyl-phosphatidylcholine (PC) (Avanti Polar Lipids, Alabaster, AL, U.S.A.) in n-decane. The membrane current was measured with a pair of calomel electrodes switched in series with a voltage source and an electrometer (Keithley model 617). For single-channel recordings the electrometer was replaced by a highly sensitive current amplifier (Keithley model 427). Zero-current membrane potentials were measured by establishing a salt gradient across membranes containing 100–1000 channels, as has been described earlier using a high impedance electrometer (Keithley model 617).

---

4.2.9 Modeling of the OmpW structure

The possible 3D-structure of OmpW of *C. crescentus* was derived using the homology modeling approach. Three dimensional model of *C. crescentus* OmpW was built using the program Modeller\textsuperscript{205} taking *E. coli* OmpW as a template structure.\textsuperscript{196}

4.3 Results

4.3.1 Protein composition of the outer membrane-enriched preparations of *C. crescentus*

Crude membranes from *C. crescentus* CB15A NA1000 353Φ cells that were disrupted by sonication and membrane material released by PBS-EDTA treatment were analyzed by SDS-PAGE. The membranes from the sonicated cells (presumably a mixture of the cytoplasmic and outer membranes) (fig. 4.1, lane 1) and the membranes from the PBS-EDTA extract (fig. 4.1, lane 2) showed numerous proteins present, whereas the PBS-EDTA extracted cell membranes (not boiled before analysis) showed an enrichment of two protein bands at about 20 and 22 kDa. When the PBS-EDTA extract preparations were boiled prior to SDS-PAGE, the two enriched bands resolved as a single band of about 22 kDa in size (fig. 4.1, lane 3). This ‘heat modifiability’ of the enriched protein (i.e. lower mobility when boiled in the presence of SDS) is occasionally observed for Gram-negative bacterial outer membrane proteins and is probably caused by their unfolding.\textsuperscript{206}

**Figure 4.1**: Coomassie stained SDS-PAGE of different protein preparations of *C. crescentus*. Lane 1, 7.5 µl of crude membrane preparation. Lane 2, 15 µl of PBS-EDTA extract, not boiled prior to loading. Lane 3, 15 µl of PBS-EDTA extract, boiled prior to loading. The star highlights the position of OmpW.

### 4.3.2 Identification of channel-forming activity in outer membrane-enriched preparations of *C. crescentus*

The outer membrane-enriched preparations were treated with 0.5% LDAO. When small amounts of this detergent extract of the outer membranes was added to one or both sides of a black PC/n-decane membrane a remarkable increase of membrane conductance was observed, indicating that the enriched outer membranes contained membrane-active components. After an initial rapid increase of conductance for 5–10 min, the increase slowed down and saturated at 30–60 min after addition of the detergent solution to the black membranes. When the detergent alone was added in the same concentration as with the PBS-EDTA extracted membranes, it had no
influence on the conductance of the lipid bilayer membranes. When the detergent extract of the outer membranes was added at much lower concentrations to the aqueous phase bathing the black lipid membrane, the current increased in a step-wise fashion (see fig. 4.3A, upper panel). A histogram of the channel distribution demonstrated that most of the conductance steps had a conductance of about 125 pSv in 1M KCl (see fig. 4.3A, lower panel). Besides the 125 pSv channel, some fluctuations with higher conductance (250 and 350 pSv) were also obtained which probably represented oligomers of the 125 pSv channel that could not be separated at the time scale of the experimental conditions. It is significant that the conductance of these steps was notably much lower than that of general-diffusion pores from enteric bacteria, which is at least 10 times higher at the same conditions (1–4 nSv in 1 M KCl\(^{207}\)). This result suggested that the PBS-EDTA extracted membranes of \textit{C. crescentus} contained a porin-like channel of small conductance.

### 4.3.3 Identification of the purified channel-forming protein from the enriched outer membranes

To identify the protein responsible for the channel-forming activity, the PBS-EDTA extracted membranes were dissolved in detergent and subjected to preparative SDS-PAGE. The gel was cut into thin slices, corresponding to defined molecular mass ranges, and each was eluted overnight with a buffer containing 1% Genapol. The eluted molecular mass fractions were examined for channel-forming activity in the lipid bilayer assay. Extremely high channel-forming activity was exclusively localized within the molecular mass range around 20 to 22 kDa, corresponding to the protein bands enriched by PBS-EDTA extraction (see fig. 4.1). The other bands from the gel, in particular the 66 kDa band, had no activity in the lipid bilayer assay. When excised and eluted, the 20–22 kDa band was again subjected to SDS-PAGE. Without heating, the gel showed a single protein band of about 20 kDa suggesting that the excised protein was essentially pure (fig. 4.2, lane 2). When the excised 20 kDa band was heated to 100°C prior to addition to

SDS-PAGE, most of the protein ran at an apparent molecular mass of about 22 kDa (fig. 4.2, lane 2), with the indication that some of the protein ran at 20 kDa (fig. 4.2, lane 2). This suggested again that the 20 kDa protein from *C. crescentus* is heat modifiable.

**Figure 4.2:** 15% SDS-PAGE showing OmpW of *C. crescentus* obtained by elution of the 20 kDa band from preparative SDS-PAGE. Lane 1: 5 µg of OmpW solubilized at room temperature for 10 min in 5 µl of sample buffer. Lane 2: 5 µg of OmpW solubilized at 100°C for 10 min in 5 µl of sample buffer. The gel was stained with Coomassie.

### 4.3.4 Partial sequencing of the 22 kDa protein of *C. crescentus*

The 20 kDa protein was subjected to sequencing using Edman-degradation. In a first run the protein could not be sequenced starting from the N-terminus, presumably because of N-terminal
blockage. Following trypsin treatment one stretch corresponding to the N-terminal end with a molecular mass of 1764.8 Da could be resolved by sequencing. The sequence of the partial peptide was QDFTPNAKGDLIVHAR, which suggested that the N-terminus was blocked by the formation of pyroglutamate. A basic local alignment search (BLAST) analysis of the sequenced peptide unambiguously demonstrated that the 20 kDa protein was OmpW of *C. crescentus* and was a member of the extensive OmpW-family of outer membrane proteins of Gram-negative bacteria. To ensure that the higher molecular mass band observed in boiled samples was also OmpW, this protein band (about 22 kDa) was also subjected to sequencing following trypsin treatment. Its N-terminal end was identical to that of the 20 kDa protein (OmpW), indicating again that OmpW existed in two different configurations where one was heat-modifiable.

**4.3.5 Analysis of the channels formed by OmpW channel-forming protein of *C. crescentus***

Figure 4.3B (upper panel) shows a single-channel recording of a PC membrane in the presence of the purified OmpW protein of *C. crescentus*, which was added to a black lipid membrane at a concentration of about 20 ng ml$^{-1}$. The single-channel recording demonstrates that the protein formed the same defined channels as found in PBS-EDTA extracted membranes of *C. crescentus*. The average single-channel conductance was about 125 pSv in 1 M KCl (almost 40% of all conductance steps). Only a minor fraction of channels with other conductance was observed (see the histogram in fig. 4.3B, lower panel) suggesting that conductance steps with more than one unit conductance were less frequent for the purified OmpW than for the crude outer membrane fraction. It is noteworthy that the channels formed by OmpW of *C. crescentus* had a long lifetime, similar to those that have been detected previously for porins of Gram-negative and Gram-positive bacteria. All these channel-forming proteins from the cell walls of

---

bacteria form channels in lipid bilayer membranes with long lifetimes at small transmembrane potential (mean lifetime at least 5 min). Furthermore, no voltage-dependence closure was observed in KCl solution voltages up ±120 mV (data not shown).

Single-channel experiments were also performed with salts other than KCl to obtain information on the size and selectivity of the channels formed by OmpW of *C. crescentus*. The results are summarized in Table 4.1. The conductance sequence of the different salts within the channel was KCl ≈ KOAc ≈ NH₄Cl > RbCl > NaCl > CsCl > LiCl. The influence of cations of different size and mobility on the conductance was substantial (see Table 4.1) suggesting indeed high cation-selectivity of the OmpW channel. For more bulky cations such as N(CH₃)₄⁺ or Tris⁺ we observed a very low conductance of much less than 10 pSv suggesting that the size of the OmpW channel was indeed very small.

An additional significant result of the single-channel measurements was the extreme high conductance of OmpW in CaCl₂ (see Table 4.1). In 1 M CaCl₂, the conductance was about 250 pSv; this was again close to saturation because in 0.5 M CaCl₂ the conductance was only a little lower than in 1 M solution. Interestingly, conductance traces in CaCl₂ solutions were very noisy suggesting a strong interaction between the divalent cations and the OmpW channels. Also, similarly, channel-forming activity in salt solutions containing divalent cations was much lower than that in monovalent cation solutions, indicating that OmpW could be a channel for divalent cations, such as Ca²⁺ or Mg²⁺. Table 4.1 also shows the average single-channel conductance, G, as a function of the KCl concentration in the aqueous phase. Similarly, as in the case of some channels of Gram-positive bacteria, the conductance was not a linear function of the KCl-concentration, which means that OmpW did not form a wide, water-filled channel. The saturation with increasing salt concentration could be caused by point net charges and/or a binding site for ions.

---

Figure 4.3: Single-channel measurements with detergent treated outer membrane of *C. crescentus* and purified 20 kDa protein.

**A** Upper panel: Single-channel recording of a PC/n-decane membrane in the presence of enriched outer membranes of *C. crescentus*. The aqueous phase contained 1 M KCl and 100 ng ml$^{-1}$ protein from enriched outer membranes treated with 0.5% LDAO. The applied membrane potential was 20 mV; $T = 20^\circ$C.

Lower panel: Histogram of the probability $P(G)$ for the occurrence of a given conductivity unit observed with membranes formed of PC/n-decane in the presence of enriched outer membranes of *C. crescentus*. $P(G)$ is the probability that a given conductance increment $G$ is observed in the single-channel experiments. It was calculated by dividing the number of fluctuations with a given conductance increment by the total number of conductance fluctuations. The aqueous phase contained 1 M KCl. The applied membrane potential was 20 mV; $T = 20^\circ$C. The average single-channel conductance was 125 pSv for 105 single-channel events (left-hand maximum).

**B** Upper panel: Single-channel recordings of a PC/n-decane membrane in the presence of purified OmpW of *C. crescentus*. The aqueous phase contained 1 M KCl and 20 ng ml$^{-1}$ OmpW dissolved in 1% Genapol. The applied membrane potential was 20 mV; $T = 20^\circ$C.

Lower panel: Histogram of the probability $P(G)$ for the occurrence of a given conductivity unit observed with membranes formed of PC/n-decane in the presence of OmpW of *C. crescentus*. The applied membrane potential was 20 mV; $T = 20^\circ$C. The average single-channel conductance was 125 pSv for 95 single-channel events.
Table 4.1: Average single-channel conductance of OmpW of *C. crescentus* in different salt solutions and radii, hydrated radii, and limiting molar conductivity of the cations. This data is also presented in a graphical figure, fig. 4.5

<table>
<thead>
<tr>
<th>Salt</th>
<th>Molarity</th>
<th>Single-channel conductance</th>
<th>Ion Radius</th>
<th>Hydrated Ion Radius</th>
<th>Limiting Molarity Conductivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiCl</td>
<td>1.0</td>
<td>15 ± 3.0</td>
<td>0.059</td>
<td>0.216</td>
<td>38.68</td>
</tr>
<tr>
<td>NaCl</td>
<td>1.0</td>
<td>40 ± 3.5</td>
<td>0.100</td>
<td>0.163</td>
<td>50.10</td>
</tr>
<tr>
<td>KCl</td>
<td>0.01</td>
<td>30 ± 3.2</td>
<td>0.137</td>
<td>0.110</td>
<td>73.50</td>
</tr>
<tr>
<td>KCl</td>
<td>0.03</td>
<td>40 ± 3.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCl</td>
<td>0.1</td>
<td>55 ± 4.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCl</td>
<td>0.3</td>
<td>80 ± 6.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCl</td>
<td>1.0</td>
<td>125 ± 8.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KCl</td>
<td>3.0</td>
<td>150 ± 8.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH₄Cl</td>
<td>1.0</td>
<td>125 ± 9.5</td>
<td>0.147</td>
<td>0.110</td>
<td>73.55</td>
</tr>
<tr>
<td>RbCl</td>
<td>1.0</td>
<td>100 ± 7.6</td>
<td>0.152</td>
<td>0.105</td>
<td>77.81</td>
</tr>
<tr>
<td>CsCl</td>
<td>1.0</td>
<td>30 ± 2.9</td>
<td>0.167</td>
<td>0.106</td>
<td>77.26</td>
</tr>
<tr>
<td>N(CH₃)₄Cl</td>
<td>1.0</td>
<td>&lt;10</td>
<td>0.347</td>
<td>0.182</td>
<td>44.92</td>
</tr>
<tr>
<td>KAcO⁻ (pH 7)</td>
<td>0.1</td>
<td>60 ± 8.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KAcO⁻ (pH 7)</td>
<td>1.0</td>
<td>125 ± 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CaCl₂</td>
<td>0.5</td>
<td>225 ± 22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CaCl₂</td>
<td>1.0</td>
<td>250 ± 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1.0</td>
<td>275 ± 27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.3.6 The OmpW channel of *C. crescentus* is highly cation-selective

Additional information about the structure of the channel formed by OmpW of *C. crescentus* was obtained from zero-current membrane potential measurements in presence of salt gradients. A fivefold KCl gradient (100 mM versus 500 mM), across a lipid bilayer membrane in which about 100–1000 OmpW channels were reconstituted, resulted in an asymmetry potential of 35 mV at the more dilute side (mean of 3 measurements). This result indicated preferential movement of K\(^+\) ions over Cl\(^-\) through the channel at neutral pH. The zero-current membrane potentials were analyzed using the Goldman-Hodgkin-Katz equation, see eq. (4.1). The ratio of the potassium permeability, \(P_K\), divided by the chloride permeability, \(P_{Cl}\), was about 15 (mean of 3 measurements), indicating high cation selectivity of the channel formed by OmpW (see also Discussion, page 107). This result was confirmed by measurements with LiCl and KOAc; we observed under the same conditions as for KCl, asymmetry potentials around 32 to 35 mV for fivefold gradients, meaning that \(P_a/P_c\) was also in these cases also higher than 10. More precise numbers cannot be expected because small errors of the asymmetry potential result in this range of big variations of the permeability ratio \(P_a/P_c\). Size and mobility of the cations did not influence the cation selectivity of OmpW in contrast to the situation observed previously for general diffusion porins.

\[
E_{m,\text{ion}} = \frac{RT}{F} \ln \left( \frac{P_{\text{ion}}[\text{ion}^+]_{\text{out}}}{P_{\text{ion}}[\text{ion}^+]_{\text{in}}} \right) \tag{4.1}
\]

4.3.7 Knocking out *ompW* removes channel-forming activity

To confirm that the gene product encoded by CCNA_01475 was indeed responsible for the channel forming activity, we removed the genomic copy of the gene with a two-step recombination protocol that is standard in *C. crescentus* genetics. The gene loss in the resulting strain, *C. crescentus* \(\Delta\text{ompW}\), was confirmed by PCR (data not shown) and SDS-PAGE. The SDS-PAGE analysis of the knockout compared to wild-type can be seen in fig. 4.4. The band that was
identified previously as comprising the channel-forming protein was found to be absent in the ompW knockout. Likewise, the channel-forming activity previously found in our extracts was not present in our ompW knockout extracts. These results support the peptide sequencing in that ompW, CCNA_01475, encodes for an outer membrane porin.

Interestingly, all attempts to scale-down our outer membrane extractions, from large flasks to test tubes, were unsuccessful at producing a sample with a prominent SDS-PAGE band near 20 kDa (the OmpW band). It is not clear what the important differences are between a flask-based extraction and a tube-based extraction. A significant amount of time was spent trying to reproduce our original results; culture volume was not considered a significant variable until very late. The reason for this phenomenon is unknown.

Attempts to determine a phenotype for the knockout strain were inconclusive. We hypothesized that the OmpW was involved with Ca$^{2+}$ uptake and that without OmpW the cells would have reduced fitness under Ca$^{2+}$ limitation. It has previously been shown that C. crescentus does require Ca$^{2+}$ in its growth medium, but that independence from Ca$^{2+}$ can be achieved and ‘calcium-free’ mutants have been isolated.\(^{39}\) All recent attempts to produce growth media without Ca$^{2+}$, in the manner previously published, were not successful. It seems that there may be a trace level of calcium contamination of some medium components. Nevertheless, no reduction in fitness was observed for our ompW knockout strain compared to wild-type in media replete with Ca$^{2+}$ or in media with significantly reduced Ca$^{2+}$. In fact, no in vivo phenotype of any kind was observed with our strain C. crescentus $\Delta$ompW.

### 4.4 Discussion

The outer membrane-enriched preparations of C. crescentus contain a cation-permeable channel, encoded by ompW. Here we demonstrated the presence of a channel in the outer membrane-enriched extractions of C. crescentus. Unlike enteric Gram-negatives, the genome does not con-

Figure 4.4: SDS-PAGE analysis of outer membrane-enriched preparations of *C. crescentus*. Lane 1: *C. crescentus* NA1000. Lane 2: *C. crescentus ∆ompW*. The star highlights the position of OmpW, the prominent band at 22 kDa.

Certain genes that code for the classical outer membrane porins such as OmpC, OmpF, LamB or Tsx.\textsuperscript{[20][21][193]} Instead the genome of *C. crescentus* contains 67 genes coding for TonB-dependent receptors.\textsuperscript{193} This suggested that most nutrients from the dilute environments where *C. crescentus* is typically found are actively taken up by TonB-dependent transport systems. Nevertheless, it is clear from this study that *C. crescentus* also contains at least one outer membrane channel. The 20 kDa band excised from preparative SDS-PAGE had a very high channel-forming activity. No other protein bands excised from gels showed channel-forming activity, suggesting that the outer membrane of *C. crescentus* may contain only one porin-like channel. The channel had a very low single-channel conductance of about 125 pSv in 1 M KCl. The channel-forming protein was subjected to partial sequencing and was identified as OmpW of *C. crescentus*. The gene found to encode for the OmpW, CCNA_01475, was knocked out and the extracts of the resulting strain showed no channel-forming activity.

The OmpW porin family is widespread amongst Gram-negative bacteria, but with no well established function. Only for OmpW of *E. coli* and OprG of *P. aeruginosa* have channel functions been postulated that may have to do with the uptake of hydrophobic compounds.\textsuperscript{[196][197]} In other studies it has been suggested that the channel function of OmpW of *E. coli* and OprG
of *P. aeruginosa* may be plugged by the tryptophans W155 and W170, respectively (see section 4.4).

Two big-dataset studies from the Shapiro lab at Stanford University have shed some light on the biology of *ompW* in *C. crescentus*: the survey of essential genes by Christen et al. (2011) and the high-throughput analyses of transcription and translation by Schrader et al. (2014). The gene *ompW* is non-essential for *C. crescentus*, at least in the rich growth medium PYE. The gene lies on the forward strand of the genome surrounded by genes of seemingly unrelated function. It is transcribed as a monocistronic mRNA, while the preceding and succeeding genes are both transcribed on separate polycistronic mRNAs. *ompW* is not highly transcribed, nor translated in either rich medium (PYE) nor limiting medium (M2G), but interestingly the gene is transcribed at a much higher relative rate in rich medium than in limiting medium while its translation rate remains relatively equal in both media. The differences in transcription rate and translational efficiency between medium types was relatively small within the entire study by Schrader et al., and so it was not focused on or discussed in the original study. The reduction in transcriptional activity in minimal medium suggests some transcriptional control, but the translation rate (as measured by ribosome footprinting RNA-seq) remained relatively unchanged (0.9 versus 0.8 ribosomes per kilobase per million-reads) suggesting that the changes in transcription maybe superseded by a translational control mechanism.

**Properties of OmpW in the *C. crescentus* outer membrane** OmpW of *C. crescentus* was found to be highly cation-selective. Its selectivity for K\(^+\) ions over Cl\(^-\) was at least ten-fold; the calculated value was about 12. However, it has to be kept in mind that the permeability ratio Pa/Pc reacts in this range and is very sensitive to small changes of the asymmetry potential. This means that we found little indication for the permeation of anions through OmpW because the single channel conductance in 1 M KCl was the same as in 1 M KOAc, despite the fact

---

that chloride in the aqueous phase has a much higher mobility than acetate. In addition, the single channel conductance was not a linear function of the bulk aqueous salt concentration, see table [4.1]. Instead the conductance showed strong saturation; at 3 M KCl conductance was only slightly higher (150 pSv) than at 1 M KCl (125 pSv). Similarly, at 0.3 M KCl, conductance was only little smaller (80 pSv) than at 1 M KCl. Experiments with different alkaline cations indicated clearly that OmpW does not form an aqueous pore, which would be typical for most Gram-negative and Gram-positive bacterial porins, even if they contained point charges. Instead OmpW of *C. crescentus* appeared to be an ion channel; its single channel conductance had a maximum for ammonium and potassium ions and became much smaller for bigger cations indicating that they likely lose part of their hydration shell while moving through the channel. This argues for a small selectivity filter inside the channel. Presumably, the ionic radii of the different cations and not the sizes of their hydration shells, play an important role for their permeation through OmpW as table [4.1] clearly indicates. Figure [4.5] shows the single channel conductance of OmpW as a function of the ionic radii of the ions. The conductance data for another small substrate-specific channel (LamB of *E. coli*) is given for comparison. The relationship for OmpW follows approximately the Eisenman series VI for carrier-mediated transport of cations or for the transport of monovalent cations through channels. This means that the field strength inside the channel is medium-sized. Ion transport trough general diffusion channels of enteric bacteria follow instead Eisenman series I or II indicating low field strength in the channels. The minimum diameter of OmpW is presumably close to that of Cs⁺ or smaller than that of N(CH₃)₄⁺ ions because larger organic cations have a very low permeability through OmpW, if any.

OmpW of *C. crescentus* clearly functions as a cation-permeable channel, having a prefer-

---

Figure 4.5: Single-channel conductance of OmpW in 1 M salt solution as a function of the ionic radii of monovalent cations. The data were taken from table 4.1. The single-channel conductance through LamB of *E. coli* is given for comparison (data taken from R. Benz et al. Pore formation by LamB of *Escherichia coli* in lipid bilayer membranes. *Journal of Bacteriology*, 165: 978–986, 1986.).

ence for divalent cations, since calcium and magnesium ions had a higher permeability through OmpW than potassium ions. This could indeed mean that it is a channel for the transport of calcium and/or magnesium ions. This function has not been found in other homologous versions of OmpW. The reason for this is that OmpW is a rather small channel with only eight beta strands and so has little possibility for the passage of solutes. Previous data from structural studies of OmpW of *E. coli* and OprG of *P. aeruginosa* suggested that members of the OmpW family could be involved in the transport of small hydrophobic molecules across the bacterial outer membrane. Lipid bilayer experiments

111
with OmpW of *E. coli* suggested that it forms small ion-permeable channels with a conductance of about 20 pSv in 1 M KCl[^196] which is considerably lower than that of *C. crescentus* OmpW. The 3D-structure of *P. aeruginosa* OprG, another member of the OmpW family, was resolved at 2.4 Å resolution[^197] Again the structure suggested that OprG forms a channel for the diffusion of small hydrophobic molecules, although lipid bilayer experiments proposed a single-channel conductance of about 500 pSv for OprG[^221] This means that OmpW of *C. crescentus* forms a channel that has, despite sequence homologies with OmpW of *E. coli* and OprG of *P. aeruginosa* (see below), a completely different function than other outer membrane proteins.

**Structure of OmpW of *C. crescentus***  
Sequence comparison, BLAST, of OmpW of *C. crescentus* with other members of the OmpW family suggested that it had highest homology with OmpW of *Caulobacter segnis* ATCC 21756, *Asticcacaulis excentricus* CB 48, *Brevundimonas* sp. BAL3, and *Phenylobacterium zucineum* HLK[^208][^209] These bacteria are closely related to the genus *Caulobacter*, belonging to the family *Caulobacteraceae*, order *Caulobacterales*. All these OmpW polypeptides are similar in length to *C. crescentus* OmpW (200–250 amino acids) and share many strategically positioned conserved residues. The homology of the *C. crescentus* OmpW to the OmpW proteins with known 3D structure (OmpW of *E. coli* and OprG of *P. aeruginosa*) is less pronounced, but still obvious (amino acid identity is approximately 32% in both cases and 19.4% for all three proteins (Hong et al., 2006; Touw et al., 2010). Forty one amino acids, including many glycines and several prolines, appear to be highly conserved among all three OmpW family proteins. This allows a meaningful comparison between the three different OmpW species (see fig. 4.6) and it is also possible to build the possible structural model of OmpW of *C. crescentus* using a homology modeling approach[^205] (see below).

Crystal structures of *E. coli* OmpW[^196] and *P. aeruginosa* OprG[^197] suggested the formation of hydrophobic channels for these proteins and they are believed to be responsible for the passage of hydrophobic solutes across the bacterial outer membrane. The small and hydrophobic

**Figure 4.6:** Amino acid sequence alignment of OprG of *P. aeruginosa*, OmpW of *E. coli*, and OmpW of *C. crescentus*. The alignment was performed using Pole Bioinformatique Lyonnaise Network Protein Sequence Analysis ([http://npsa-pbil.ibcp.fr](http://npsa-pbil.ibcp.fr)). Amino acids identical in all three proteins are highlighted in red, strongly similar amino acids (:) are given in green and weakly similar ones (.) in blue. The replacement of W155 of OmpW of *E. coli* and W170 of OprG of *P. aeruginosa* by K160 of *C. crescentus* is given in green color and is indicated by an arrow. The eight beta strands in OprG of *P. aeruginosa* and in OmpW of *E. coli* are numbered and indicated by blue bars. The yellow highlighted sequence was found by N-terminal sequencing of OmpW of *C. crescentus*.

*E. coli* OmpW channel has a conductance of approximately 20 pSv in 1 M KCl. In contrast, the *C. crescentus* OmpW formed a channel with a relatively high conductance of 125 pSv in 1 M KCl in our bilayer measurements. A comparison of the structures of these OmpW family proteins revealed factors that might lead to a larger conductance in *C. crescentus* OmpW. Crystal structures suggest that both *E. coli* OmpW and *P. aeruginosa* OprG form a hydrophobic gate.
(fig. 4.7A; b and c) in the central region of the channel. Both channels have an aromatic, bulky and hydrophobic tryptophan residue as a part of a hydrophobic gate, which may make the passage of ions through channels difficult. Similarly, W155 of OmpW of *E. coli* and W170 of OprG of *P. aeruginosa* which are both considered as plugs of the channels are replaced in OmpW of *C. crescentus* by K160. This may indicate that this channel has a function different from that of the homologs in *E. coli* and *P. aeruginosa*. In addition, *C. crescentus* OmpW has a hydrophilic and relatively less bulky lysine residue at the corresponding position (fig. 4.7A; a), which is unlikely to hinder passage of ions through the channel. Also, the *C. crescentus* OmpW has a relatively more hydrophilic interior environment of the channel (fig. 4.7A; a), as shown by green color surface within the black box) as compared to the *E. coli* OmpW (fig. 4.7; b) and the *P. aeruginosa* OprG (fig. 4.7; c). For *E. coli* OmpW and *P. aeruginosa* OprG the interior surface of the channel is hydrophobic (white color surface), which could make the permeation of ions through the channel energetically unfavorable. *C. crescentus* OmpW, on the other hand, due to its relative hydrophilic environment may not provide the permeation barrier to ion passage, resulting in a relatively higher conductance of the channel.

We also examined the contribution of the size of the to the higher conductance of *C. crescentus* OmpW. We observed a slightly larger radius of *C. crescentus* OmpW channel as compared to the other two channels in several regions, particularly from the center of the channel towards the periplasmic side (fig. 4.8). Another important feature of the OmpW family channels from *E. coli* and *P. aeruginosa* is the presence of a lateral opening in the barrel wall, which has been suggested to allow diffusion of small hydrophobic solutes across the outer membrane by a lateral diffusion mechanism. In our modeled structure of *C. crescentus* OmpW, we do not observe such an opening in the channel, which further supports the view that the *C. crescentus* channel may have a different function from that of *E. coli* OmpW and *P. aeruginosa* OprG. The solved structures OmpW from *E. coli* and OprG from *P. aeruginosa* provide examples of successful crystallization efforts for this class of protein and they are exciting starting points for future work towards improving the structural knowledge of OmpW from *C. crescentus*. 
**Figure 4.7:** Comparison of structural features between a) OmpW *C. crescentus* b) OmpW *E. coli* c) OprG *P. aeruginosa*. Residues, which are a part of a putative hydrophobic gate in *E. coli* (W155 and L56) and *P. aeruginosa* (W170 and V65) channels are shown as spheres. Corresponding residues in OmpW *C. crescentus* (K159 and I65) are also shown. Additionally all the channels are shown as a surface representation and color coded according to residue type (Green: hydrophilic, white: hydrophobic, red: acidic, blue: basic). Channels are cut from the front to enable visualization of interior channel surface. Black colored box indicates a putative ion transport pathway across the channel. EC and PC denote extracellular and periplasmic sides of the channel respectively.

*C. crescentus* has a requirement for calcium, although it is not clear where the requirement lies. Ca$^{2+}$ ions are needed for assembly of the crystalline protein surface layer (S-layer) RsaA. Since RsaA is secreted by a type I secretion mechanism, it is likely that calcium is also needed for either secretion or folding of the protein following secretion, in a manner analogous to all other type I secreted proteins. However, the S-layer is a completely dispensable structure. Moreover, the type 1 secretion apparatus spans both the cytoplasmic and outer membranes; hence for S-layer there is no apparent need for a channel that enables Ca$^{2+}$ ion movement to the periplasm. There are, however, additional still undefined roles for Ca$^{2+}$ ions. Mutants no longer requiring calcium ions for growth can be isolated.\(^{128}\) One consequence of all these mutants is the loss of the O-specific polysaccharide (OPS) of LPS that is used for S-layer attachment. As a consequence, these so called ‘calcium-independent’ mutants all shed RsaA into the culture.

---


Figure 4.8: Comparison of channel radii between *C. crescentus* OmpW (blue), *E. coli* OmpW (red) and OprG of *P. aeruginosa* (green) along the channel axis. The channel radii were calculated using the program HOLE (O. S. Smart et al. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics*, 14: 354–360, 1996.). The extracellular side is indicted along the right side and the periplasmic side is indicated along the left side.

medium. The causal relationship has not been deciphered, but since OPS biosynthesis involves synthesis activities within the periplasmic space, it may be that a calcium selective OmpW porin plays a specific role in this process. Here we have provided clear evidence that OmpW could be responsible for the uptake of cations, especially divalent cations, like Ca$^{2+}$ in *C. crescentus*. 

Chapter 5

Conclusions

“We must never let ourselves fall in thinking ‘ignorabimus’ (‘We shall never know’), but must have every confidence that the day will dawn when even those processes of life which are still a puzzle today will cease to be inaccessible to us natural scientists.”
— Eduard Buchner, 1907 Nobel lecture

5.1 Relevance and contribution to the field

The Caulobacter field is dominated by research groups teasing apart the fine details of the Caulobacter crescentus lifecycle and the panoply of regulatory networks that control every aspect of bacterial biology. The protein surface layer (S-layer) and the outer membrane often are neglected by the rest of the field. The focus on laboratory cultures, and not the bacteria in the environment, is one reason for this. A pure culture living in rich medium does not have as complex a relationship with its environment as a wild specimen competing in a hostile setting. It is the cell envelope that acts as the barrier between the world and the cell. The cell envelope of C. crescentus was particularly suitable for inspection because it had a protein surface layer, various polysaccharides including an O-specific polysaccharide (OPS) of unknown structure, and apparently no porins. There is the added aspect that C. crescentus, and its S-layer, are actively being pursued for biotechnology applications. This means that new findings can often be quickly be applied towards biological products. This work presented in
chapters 2, 3, and 4 present our efforts to improve the structure and composition knowledge of the S-layer, lipopolysaccharide (LPS), and outer membrane proteins in \textit{C. crescentus}.

Chapter 2 detailed effort to solve the structure of RsaA. Protein crystallography is the only approach applicable to extend our current resolution of the \textit{C. crescentus} S-layer protein. S-layer proteins are notoriously resistant to crystallography. The work of Pavkov et al.\cite{Pavkov2008} and Baranova et al.\cite{Baranova2012} demonstrated that at least portions of S-layer proteins are possible to crystallize as long as the ability for the proteins to form S-layers is blocked. Our strategy to impede S-layer formation was a truncation approach, removing the N-terminal 222 amino acid (AA), which mirrored the successful approach used to solve SbsC\cite{Pavkov2008}. Novel protein expression and purification techniques had to be developed to produce well-behaved protein suitable for crystallography. Once we started to grow crystals, RsaA was set to be the third bacterial S-layer protein to be structurally solved, the second with the symmetry determining domain intact, the first from a Gram-negative, and the largest. But so far, finding a phasing solution has been the issue that is insurmountable. No phasing approach was successful, not the halide soaks that we had predicted would work, nor the lanthanide soaks that had worked for other groups,\cite{Baranova2012} nor any of a large range of compounds tried. The C-terminal 804 AAAs segment of RsaA has been crystallized, but the structure remains elusive. Solving this structure would provide the field with unique window into a self-assembling protein lattice from a commonly studied Gram-negative model organism.

LPS structures solutions are not as rare as S-layer structures and yet the complete structure of the \textit{C. crescentus} LPS had never been determined when we started the work covered in chapter 3. The \textit{C. crescentus} lipid A structure was solved by Smit et al. in 2008,\cite{Smit2008} revealing a novel structure lacking phosphates and centered on a di-diaminoglucose backbone. The structures of the oligosaccharide (OS) and OPS were identified for analysis because their fine structures had never been determined and because they comprise the anchor for the \textit{C. crescentus}.

An initial component analysis of the core OS was undergone in 1992. The OPS had never been studied directly but it was predicted that it was composed of perosamine. Through a very successful collaboration with Dr. Evgeny Vinogradov the structures of the C. crescentus OS and OPS are now determined. The core OS has a three armed configuration (see fig. 3.11 on page 82). The final determined components of the core OS are different than were first reported, especially in the finding that there is no phosphate in the core OS. The first arm is a disaccharide of glucuronic acid and galactose. The two negative charges in the galacturonic acids in the lipid A, the one in the central Kdo, and the one in the glucuronic acid are the only charges present in the C. crescentus LPS. The second arm is a trisaccharide of DD-heptose, mannose, and DL-heptose. The third arm is a single DL-heptose, linked to the C-7 position of the Kdo moiety. This triple substitution on the Kdo appears to be unique in the literature. There are examples of sugar substitutions at all of the same positions (C-4, C-5, and C-7) in Kdo but none for all three at once. The OPS has a repeating heptasaccharide structure (see fig. 3.7 on page 75). Seven sugar OPS subunits are amongst the largest sizes that subunits are known to exist in. For example, there is just one known example of a seven sugar OPS in Escherichia coli, and none larger. The C. crescentus OPS terminates in one or the other of two end groups (see fig. 3.9 on page 78). A complete structural view of the C. crescentus LPS is shown in fig. 5.1. Both the OPS and end groups contain an abundance of hydrophobic sugars, i.e. N-acetylperosamine, rhamnose, methylglucose, and dimethyl-N-acetylperosamine. There may be a role for this hydrophobicity to play in supporting an S-layer. The OPS composition of Aeromonas hydrophila AH-1, another alphaproteobacterium with an S-layer, supports this hypothesis; one in two monosaccharides in its OPS is the deoxysugar acetylrhamnose. In the process of studying the C. crescentus LPS we discovered a previously unknown rhamnose based polysaccharide (see fig. 3.10 on page 80). Just the presence of the polysaccharide alone is of

interest especially in light of recent investigations into the extracellular polysaccharides (EPSs) from *C. crescentus* by the Viollier lab at the University of Geneva.\textsuperscript{224}

Chapter 4 reports our characterization of OmpW, a porin in the outer membrane of *C. crescentus*. Previous to this work, no porin had been reported in *C. crescentus*. The wisdom on the subject was that *C. crescentus* primarily, if not solely, used active transport systems to take up nutrients from the environment.\textsuperscript{121} We found that OmpW acts as a passive protein channel in the outer membrane (see fig. 4.3). This is particularly interesting because the homologous proteins in *E. coli* and *Pseudomonas aeruginosa* do not have any pore forming activity\textsuperscript{196,197} and no clear function at all. The activity we observed is possibly due to a wider pore and specifically a lack of a plug seen in the other bacteria.

### 5.2 Future directions

Each aspect of this work, the protein crystallography, the LPS analyses, and characterization of OmpW, opened more questions than they answered. One goal of this work has always been to use all the gains made with basic science to improve the *C. crescentus* S-layer display technology in the Smit lab. The rewards in biotechnology pursuits are not yet obvious, though a complete knowledge of the LPS will undoubtedly guide efforts to assemble S-layers on abiotic surfaces.

The RsaA crystallography is the obvious loose end in this dissertation. We do not have the structure of RsaA solved to atomic resolution; this means the phase problem must be solved. It is possible that a particular compound exists or a certain soaking protocol exists that will provide us our solution. It is possible that the next RsaA Δ0–222 crystal grown will have the best diffraction yet and will give a nice clean signal. It is possible that we continue on our path and we will figure this out. We will continue trying the same approaches but doing only that

Figure 5.1: The fully assembled structure of the *C. crescentus* LPS. Included is the lipid A, the core oligosaccharide, five repeats of the OPS, and the longer of the two possible end groups. The red question mark denotes the one currently undetermined linkage, the bond between the OPS and the core.
cannot be the only way forward. The halides, specifically iodine, were the best candidates to consider at the beginning and we indeed got some very enticing early data from them. It seemed that small optimizations might be satisfactory to complete the project. We got drawn in by these successes and it delayed us trying new strategies.

Our troubles with phasing are at least in part due to a dearth of useful residues in the protein to couple with heavy atoms. For instance, there are too few methionines for seleomethionine incorporation (if that was possible in *C. crescentus*) and there are no cysteines which limits the use of many heavy metals. We can modify RsaA to contain any residues we want, at least at selected positions within RsaA; this is the strategy we have already begun with our GSCC cassette insert (see section 2.2.1). Cysteine-rich and methionine-rich inserts at positions known to tolerate heterologous insertions (e.g. AA 644, 690, 723, 784, 805, and 944) are obvious approaches to be tried. As of this writing, cloning has begun to insert various peptides into RsaA for the purposes of improving our phasing. We do have to consider the possibility that the insert-strategy is inherently flawed. We saw that our GSCC inserted RsaA did not crystallize well. The cause for that is not clear and it is possible that there is a structural instability introduced by exogenous peptides. Less common methods, like molecular replacement from a low resolution cryo-electron microscopy structure, are also possibilities that should be considered.

Once we have figured out our difficulties with the RsaA structural analysis, we can start to assess the current peptide display strategies in the context of a protein structure. Rational design choices can then be used to improve the display products. Currently display peptides are inserted into *rsaA* at sites that were first found by trial-and-error and evaluated by many years of use. Despite their consistent efficacy, we do not know how these display sites fit into the structure of RsaA. Locating sites on RsaA that are less buried in the protein may improve peptide accessibility to the environment. Choosing display sites with increased flexibility may allow for more complex tertiary structures for the display peptides to adopt.

---

Designing the insert peptides with respect to how they fit into the larger three-dimensional structure of the S-layer may open up new directions of innovation. For example, we can currently co-display two-separately engineered versions of RsaA on the surface of a single *C. crescentus* cell. But with more fine-detail information about RsaA’s structure, we could imagine displaying two types of RsaA each containing a specifically arranged half of a functional dimer that assembles in the S-layer. Another example that would be of immediate interest to our laboratory would be to improve the targeting of HIV. *C. crescentus* and RsaA constructs have been created in our laboratory that inhibit HIV entry into T-cells.\textsuperscript{59,60} The three copies of RsaA in the center of three-fold symmetry in the S-layer (see fig. 1.5) could each display a GP120-binding peptide, rationally spaced to match the trimer structure of GP120.

The LPS of *C. crescentus* is presented in chapter 3 but the study of Caulobacter’s outer membrane polysaccharides is not complete. Specifically for the OPS, there is one glycosidic bond that was not abundantly clear in our data. The linkage between the OPS and the core oligosaccharide is still unknown (see fig. 5.1). This single bond is particularly difficult to determine because it is present in nuclear magnetic resonance spectroscopy (NMR) spectra alongside the highly amplified data coming from the repeating OPS structure. A project is underway to double our efforts to identify that one missing bond by NMR. That bond could possibly be determined through genetic knockouts. There are six sugars in the core oligosaccharide that could be the base of the OPS, but only three sugars if we assume it is on the C-5 branch from the Kdo (as is common in LPS). If we knockout the glycosyltransferase responsible for adding the base sugar that supports the OPS, then we would see a loss of smooth-LPS. But that points out another gap in our understanding: the genetics of the *C. crescentus* LPS are still an open question. The work of Awram et al. (2001) did the first survey of OPS biosynthetic genes\textsuperscript{147} but now we could go back and start identifying these genes’ functions based on changes in polysaccharide structure. Only one true LPS synthesis protein from *C. crescentus* has been characterized, LpxI\textsuperscript{179}

---

(this study misidentified the *C. crescentus* lipid A precursors to be glucosamines instead of diaminogluceses); there are many more to be studied.

Even before that last bond is determined, progress can continue to use the LPS to improve the *Caulobacter* peptide-display technology and move S-layers onto abiotic surfaces. NMR analysis demonstrated that we can liberate free-OPS by mild-acid hydrolysis and that there are carboxyl groups present on the Kdo that can be chemically linked to an amine. We should redouble our efforts to chemically produce OPS functionalized surfaces, regardless of our previous negative results. Additionally, working with the Cullis Lab in the Department of Biochemistry and Molecular Biology at UBC, lipid nanoparticles have been created that are in part composed of *C. crescentus* LPS. The ultimate goal is to produce nanoparticles that display therapeutic targeting molecules. The work is preliminary but they have already been produced in various sizes and have been loaded with siRNA or fluorescent dyes. Efforts to demonstrate successful formation of S-layers of recombinant RsaA onto the nanoparticles are underway. Thus nanoparticles may be potent platforms to support an S-layer peptide-display technology.

A further area of future study is the new rhamnan polysaccharide. I am particularly interested in what role this rhamnan polysaccharide plays in the cells’ lifestyle. There are questions to be asked related to how it affects sedimentation rates, surface adhesion, and phage resistance. Discovering a new polysaccharide, rhamnan polysaccharide, like our discovery that OmpW acts as a outer membrane channel, begs the question of what else lies in the outer membrane and cell envelope waiting to be discovered next?

---

Bibliography


[16] K. MASUDA and T. KAWATA. Reassembly of the regularly arranged subunits in the cell wall of *Lactobacillus brevis* and their reattachment to cell walls. *Microbiology and Immunology, 24*: 299–308, 1980. (see p. 5)


[47] R. MEIER, T. DREPPER, V. SVENSSON, K.-E. JAEGER, and U. BAUMANN. A calcium-gated lid and a large β-roll sandwich are revealed by the crystal structure of


charide in *Francisella tularensis* subsp. novicida. *Biochemistry*, **45**: 14427–14440, 2006. (see p. 20)


Appendix

Figure A.1: The amino acid sequence of the S-layer protein, RsaA from *C. crescentus* NA1000. GeneID: CCNA_01059. Ensembl Accession number: ACL94524. The encoding gene has the coordinates of 1 159 693 bp–1 162 773 bp on the forward strand of the *C. crescentus* chromosome. The underlined sequence corresponds to the amino acids 1–222, the section of the protein that was removed to generate a crystallizable C-terminal fragment. The bolded sequences are the RTX motifs in RsaA. For the crystallization of RsaA, see chapter 2 page 31.
Figure A.2: The amino acid sequence of the outer membrane channel, OmpW from *C. crescentus* NA1000. GeneID: CCNA_01475. Ensembl Accession number: ACL94940. The encoding gene has the coordinates of 1 582 599 bp–1 583 243 bp on the forward strand of the *C. crescentus* chromosome. The bolded ‘K’ is the location of the putative hydrophobic gate in *E. coli* and *P. aeruginosa*, where that residue encodes for a tryptophan; in *C. crescentus* that residue is a lysine. For our investigations into OmpW, see chapter 4 on page 91.