Bacteriophage SfII Mediated Serotype Conversion in Shigella flexneri

by

Maria Mavris B.Sc.(Hons)(Adel)

Microbial Pathogenesis Unit
Department of Microbiology and Immunology
University of Adelaide
Adelaide 5005
Australia

A Thesis Submitted for the Degree of Doctor of Philosophy
May, 1998
Abstract

*Shigella flexneri* is the causative agent of bacillary dysentery and is responsible for approximately $10^5$ deaths per year. Organisms that belong to this species can be classified according to the O-antigenic component of their lipopolysaccharide (LPS). Determination of the serotype is based on agglutination of the organism with polyclonal antisera. *S. flexneri* exhibits at least 12 different serotypes, all of which are variations of the basic tetrasaccharide repeat:

$$\rightarrow 3)-\beta-D-GlcNAc-(1\rightarrow 2)-\alpha-L-Rha-(1\rightarrow 2)-\alpha-L-Rha-(1\rightarrow 3)-\alpha-L-Rha-(1\rightarrow$$

Variations resulting in different epitopes are due to the addition of either glucosyl or O-acetyl residues at specific positions along the repeat unit. Immunity to *S. flexneri* strains is serotype specific and therefore protection against each serotype is required.

The modifications have been found to be mediated by temperate bacteriophages. To date, two such phages have been identified and characterised; Sf6, responsible for the O-acetylation of rhamnose III and SfX which encodes genes which add a glucosyl residue to rhamnose I.

*S. flexneri* strains of serotype 2 have been associated with a highly virulent phenotype and are more prevalent in nature. Bacteriophage SfIII is responsible for the glucosylation of rhamnose III of the repeat unit and has been isolated from a strain of serotype 2b. Phage SfII possesses an hexagonal shaped head, a tail with a contractile sheath and tail fibres and has a genome of 43.5 kb in size.

A 4 kb *BamHI* fragment was isolated which was found to mediate serotype conversion. This fragment was found to contain 4 open reading frames, however, only two were required for serotype conversion. These genes were named *bgt*, which encodes a putative bactoprenol glucosyl transferase, and *gtrII* encoding the putative type II antigen determining glucosyl
transferase. These genes are adjacent to the integrase and excisionase genes and the attachment site (attP) which are highly homologous to those of Salmonella bacteriophage P22. In addition, a gene named ORF2 was identified of unknown function, which did not appear to be required for serotype conversion.

*In vitro* T7 polymerase/promoter analysis identified a protein of 34 kDa in size corresponding to Bgt. Subsequent subcellular fractionation localised this protein to the cytoplasmic and membrane fractions. Topology studies determined that the carboxy and amino terminal ends were cytoplasmically located. The active site was located to the cytoplasmic domain which is consistent with the role of Bgt. A protein was not observed for GtrII.

The ability of SfII phage and plasmid copies of serotype converting genes to modify LPS with only a single repeat unit was assessed. O-antigen polymerase mutants (rfc) of serotype X and Y, were transformed with various serotype converting plasmids and characterised by colony immunoblotting for reactivity to a variety of type and group antisera. In *S. flexneri*, it appears that a single O-antigen repeat unit is sufficient to be modified.
This thesis is dedicated to my parents,

and to my grandmother
Acknowledgments

I would, firstly, like to thank my supervisors Dr. Renato Morona and Professor Paul Manning for the opportunity to work in this area of research. To Renato, I, particularly, thank you for your encouragement and helpful ideas and discussions. I would also like to acknowledge Professor Paul Manning for reading this thesis and for his generosity in lending me the laptop on which this thesis was written.

To the members of the “Shigella group”, Luisa Van den Bosch, Craig Daniels and Elizabeth Parker, I extend my extreme gratitude and appreciation for the sharing of resources and many interesting discussions. To the remainder of the MPU, it has been a real honour to work with some fine scientists and to experience the assistance and humour of such a unique collection of individuals.

To the many departmental members, past and present, who have made a significant impact on my life during the course of this study, I extend my thanks. In particular, Debbie Charter, Elena Voss, Angelo and Evi Guidolin and Vincent Collins. To fellow (ex) students-in-arms, David Oggunniyi, Edward Bertram, Rebecca Pinyon and Jon Iredell, who have shared the ups and downs of the project as well as copious amounts of coffee and cake. Thankyou.

This study was funded with the support of the Faculty of Science scholarship. In addition, I would like to thank Alex Lech, (Microbiology and Immunology) for the opportunity to teach and Carole Gannon (Medicine) for the opportunity to learn.

I would especially like to thank my good friends Joanna Rentis and Paul Pehlivanides, who kept me sane with ‘Tuesday’ night sessions and for their interest and offers of help. Many thanks also to Sean Geary and Michael Maeorg for heartfelt and interesting conversations about everything.

Monica Ogierman, who has seen the best and worst of a student and a friend. Thankyou for your support, understanding and wonderful friendship. I will always remember you for being the sort of person who laughs so hard you have to sit on the floor, sings off-key really loudly and who enjoys herself in any situation.

To my parents, I thank you both for your endless love and encouragement, your unshakeable faith in me and for teaching me that everything is possible. I can not express my gratitude to you for everything you have done. Ἐυχαριστώ πολύ. To my family, Con, Paul and Jenny, thankyou all for the words of encouragement and advice, the ever-present laughter and your unique (sick) sense of humour.

To the Campbell family, in particular Mr. and Mrs. Campbell, thank you for the warmth and welcome you have shown me and for the interest you have expressed in my studies over the last year.

Finally, Matthew Campbell, my best friend............................................thankyou.
Abbreviations

A: adenine
A_{260}: absorbance at 260nm
aa: amino acid
ACP: acyl carrier protein
ACL: antigen carrier lipid (bactoprenol)
Ap: ampicillin
ATP: adenosine-5’-triphosphate
bp: base pair
BSA: bovine serum albumin
C: cytosine
CIP: calf intestinal phosphatase
Cm: chloramphenicol
CTP: cytosine-5’-triphosphate
DIG: digoxigenin
DNA: deoxyribonucleic acid
DNase: deoxyribonuclease
dNTP: deoxyribonucleoside triphosphate
ddNTP: dideoxyribonucleoside triphosphate
DO: deoxycholate
DTT: dithiothreitol
ECA: enterobacterial common antigen
EDTA: ethylene-diamine-tetra-acetic acid
EtBr: ethidium bromide
G: guanine
Gm: gentamycin
GTP: guanine-5’-triphosphate
HRP: horse radish peroxidase
RBS: ribosome binding site
RNA: ribonucleic acid
RNase: ribonuclease
rpm: revolutions per minute
RT: room temperature
s: sensitive
SD: Shine-Dalgarno
SDS: sodium dodecyl sulphate
S-LPS: smooth LPS;
SR-LPS: semi-rough LPS;
Sm: streptomycin
Sp: spectinomycin
sv: serovar
T: thymine
TBS: Tris-buffered saline
Tc: tetracycline
TEMED: N,N,N',N'-tetramethyl-ethylene-diamine
Tn: transposon
Tris: Tris (hydroxymethyl) aminomethane
TTBS: Tris-buffered saline with Tween-20 added
TTP: thymine-5'-triphosphate
ts: temperature sensitive
U: uracil
UV: ultraviolet
V: voltage
v/v: volume per volume
wc: whole cells
w/v: weight per volume
X-gal: 5-bromo-4-chloro-3-indolyl-β-galactopyranoside
# Table of Contents

1. Introduction ..................................................................................... 1
   1.1 Introduction ............................................................................... 1
   1.2 Discovery of Shigella Species ...................................................... 1
      1.2.1 Early serotyping of Shigella flexneri species ......................... 2
      1.2.2 Work of J.S. Boyd................................................................. 3
   1.3 Shigella Species .......................................................................... 3
      1.3.1 Epidemiology ....................................................................... 4
   1.4 Lifestyle of S. flexneri ................................................................ 5
   1.5 Genetics of Shigella Virulence .................................................... 5
      1.5.1 Chromosomal loci ................................................................. 6
         1.5.1.1 $virR$, a central regulatory locus .................................... 6
         1.5.1.2 Virulence associated chromosomal (vac) genes ................. 7
         1.5.1.3 Outer membrane protein C (OmpC) ............................... 8
   1.6 Plasmid Mediated Functions ..................................................... 8
      1.6.1 Adhesion of bacteria to cell surface ..................................... 9
      1.6.2 Entry by induced phagocytosis .......................................... 9
      1.6.3 Escape from the phagocytic vacuole ..................................... 10
      1.6.4 Intracellular multiplication and spread ............................... 10
      1.6.5 Lysis of double membrane surrounding protrusion ............. 11
      1.6.6 Biology of IcsA .................................................................. 11
   1.7 Lipopolysaccharide (LPS) ............................................................ 13
      1.7.1 Biosynthesis of LPS ............................................................... 14
      1.7.2 Lipid A- Genetics and biosynthesis .................................... 15
      1.7.3 Core region ......................................................................... 16
         1.7.3.1 Inner core: 3-deoxy-D-manno-2-octulosonic acid (KDO) .... 17
         1.7.3.2 Outer core .................................................................... 17
      1.7.4 O-antigen .......................................................................... 18
         1.7.4.1 Genetics of O antigen .................................................... 19
         1.7.4.2 O-polysaccharide biosynthesis ......................................... 19
            1.7.4.2.1 Enterobacterial Common Antigen (ECA) ................. 21
         1.7.4.3 Other genes required for O antigen synthesis ................. 21
         1.7.4.4 RfbX ............................................................................. 22
         1.7.4.5 RfaL, O-antigen ligase .................................................. 22
         1.7.4.6 Rfc, O-antigen polymerase .......................................... 22
         1.7.4.7 Rol, regulator of O antigen chain length ......................... 23
   1.8 Lipopolysaccharide and the Distribution of IcsA ......................... 24
   1.9 Chemical Understanding of O-Antigen of S. flexneri ................. 26
   1.10 Variation in O Antigen ............................................................... 27
      1.10.1 Vibrio cholerae and serotype conversion ............................ 27
      1.10.1.1 V. cholerae O139 ......................................................... 28
      1.10.2 Salmonella enterica serovar Typhimurium and serotypic variation................................................ 29
   1.11 Bacteriophages Which Have Their Receptor in the O-Antigen .... 30
      1.11.1 Lysogenisation of group E Salmonellae .............................. 31
      1.11.2 Serotype converting bacteriophages of S. flexneri .............. 32
         1.11.2.1 Identification of S. flexneri bacteriophages ................... 33
1.11.2.2 Identification of bacteriophages fII, f7,8 and fV .......................................................... 33
1.11.2.3 Bacteriophage Sf6 .......................................................... 34
1.11.3 Molecular basis for serotype conversion .......................................................... 35
1.12 AIMS OF THIS STUDY: .......................................................... 35

2. MATERIALS AND METHODS ........................................................................... 36

2.1 BACTERIAL STRAINS, BACTERIOPHAGES AND PLASMIDS ......................... 36
2.2 MAINTENANCE OF BACTERIAL AND BACTERIOPHAGE STRAINS .............. 36
2.3 GROWTH MEDIA ........................................................................ 37
2.4 CHEMICALS AND REAGENTS .......................................................... 38
2.5 ENZYMES AND IMMUNOCONJUGATES ................................................. 39
2.6 TRANSFORMATION PROCEDURE ....................................................... 39
2.6.1 Super competent cells ............................................................... 40
2.7 ELECTROPORATION OF S. FLEXNERI AND E. COLI .............................. 40
2.8 BACTERIAL CONJUGATION ............................................................ 41
2.9 PLASMID DNA EXTRACTION PROCEDURES ..................................... 41
2.10 PREPARATION OF S. FLEXNERI OR E. COLI GENOMIC DNA ................. 44
2.10.1 Quick method for genomic DNA extraction .................................... 44
2.11 ANALYSIS AND MANIPULATION OF DNA ........................................... 45
2.11.1 DNA quantitation ................................................................. 45
2.11.2 Restriction endonuclease digestion of DNA .................................... 45
2.11.3 Analytical and preparative separation of restriction fragments .......... 45
2.11.4 Calculation of restriction fragment size ...................................... 46
2.12 DNA CLONING PROCEDURES .......................................................... 47
2.12.1 Dephosphorylation of DNA using alkaline phosphatase .................. 47
2.12.2 Ligation of DNA .................................................................. 47
2.12.3 In vitro cloning ................................................................... 47
2.12.4 Cosmid Cloning ................................................................... 48
2.13 SEQUENCING METHODS ................................................................. 48
2.13.1 Sequencing using dye labelled primers .................................... 48
2.13.2 Sequencing with dye-labelled terminators ................................. 49
2.13.3 Analysis of DNA sequences .................................................. 49
2.14 POLYMERASE CHAIN REACTION PROTOCOL (PCR) ......................... 50
2.14.1 Synthesis of oligodeoxynucleotides ....................................... 50
2.15 SOUTHERN TECHNIQUES ............................................................... 51
2.15.1 Preparation of DIG-labelled DNA probes ................................ 51
2.15.1.1 DIG labelled PCR ....................................................... 51
2.15.2 Southern transfer and hybridisation .......................................... 52
2.16 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) AND WESTERN TRANSFER ................................. 52
2.16.1 SDS-PAGE ........................................................................ 52
2.16.2 Western transfer and detection ................................................ 53
2.17 LIPOPOLYSACCHARIDE (LPS) PREPARATION .................................... 54
2.17.1 LPS-specific silver staining ..................................................... 54
2.17.2 Colony immunoblotting of LPS ................................................ 54
2.17.3 Bacterial agglutination assays .................................................. 55
2.17.4 O antigen hydrolysis assay ...................................................... 55
2.17.4.1 Preparation of formaldehyde (formalin)-fixed cells ................ 55
2.17.4.2 O antigen hydrolysis assay ............................................. 55
2.18 T7 POLYMERASE / PROMOTER EXPRESSION SYSTEM AND L-[35S]-METHIONINE LABELLING 56
2.18.1 Cell Fractionation ......................................................... 57
2.18.2 Autoradiography .......................................................... 57
2.19 ALKALINE PHOSPHATASE AND B-GALACTOSIDE ASSAYS .......... 58
  2.19.1 Alkaline phosphatase assays ........................................ 58
  2.19.2 β-galactosidase assays ................................................ 58
2.20 BACTERIOPHAGE TECHNIQUES ........................................... 59
  2.20.1 Construction of lysogens .............................................. 59
  2.20.2 Phage assays ........................................................... 59
  2.20.3 Preparation of phage stocks ......................................... 60
    2.20.3.1 Low titre phage stocks ........................................ 60
    2.20.3.2 High titre phage stocks by liquid infection ..................... 60
    2.20.3.3 CsCl block density gradient for preparation of high titre phage stocks ........................................... 61
    2.20.3.4 Phenol extraction of bacteriophage DNA ......................... 61
    2.20.3.5 Bacteriophage sensitivity tests ................................... 62
2.21 ELECTRON MICROSCOPY (EM) ........................................... 62

3. ISOLATION AND CHARACTERISATION OF S. FLEXNERI SEROTYPE CONVERTING BACTERIOPHAGE SFII .................................................. 63
  3.1 INTRODUCTION .................................................................. 63
  3.2 RESULTS ......................................................................... 64
    3.2.1 Attempts to isolate bacteriophage SFII ................................ 64
    3.2.2 Cosmid cloning of serotype conversion genes from strain PE655 ......................................................... 64
    3.2.3 Isolation of bacteriophage SFII from NCTC4 .......................... 65
    3.2.4 Purification and characterisation of bacteriophage SFII ................ 65
      3.2.4.1 Proteins of bacteriophage SFII ..................................... 66
      3.2.4.2 O-antigen hydrolysis assays ....................................... 66
      3.2.4.3 SFII lysogenisation of rfc mutant RMM109 of serotype Y .................. 67
      3.2.5 Characterisation of lysogens of bacteriophage SFII ................ 67
      3.2.6 Isolation and cloning of PstI fragments of SFII genome ................ 68
      3.2.7 Characterisation of PstI fragments of SFII genome ................ 69
      3.2.8 Identification of a region of bacteriophage SFII with serotype-converting ability .................. 70
  3.3 SUMMARY .................................................................. 70

4. SEQUENCE OF PRMM264 AND IDENTIFICATION OF THE ATTACHMENT REGION OF BACTERIOPHAGE SFII ........................................... 71
  4.1 INTRODUCTION ................................................................. 71
  4.2 RESULTS ......................................................................... 72
    4.2.1 DNA sequencing of the BamHI-BamHI insert in PRMM264 ............ 72
    4.2.2 Open reading frames in 4 kb BamHI fragment of PRMM264 ............. 72
    4.2.3 Open reading frames in 5.2 kb sequence .................................. 73
    4.2.4 Integrase (int) and Excisionase (xis) ...................................... 73
    4.2.5 The attachment site, attP ................................................ 73
  4.3 SUMMARY .................................................................. 75

5. LOCALISATION OF THE SEROTYPE CONVERTING GENES OF BACTERIOPHAGE SFII ............................................................ 76
  5.1 INTRODUCTION ................................................................. 76
  5.2 RESULTS ......................................................................... 77
    5.2.1 ORF2 .................................................................. 77
5.2.2 ORF3 bactoprenol glucosyl transferase (bgt) ........................................ 77
5.2.3 ORF4 - glucosyl transferase gtrII ...................................................... 78
5.3 COMPLEMENTATION ANALYSES ............................................................... 79
5.3.1 Serotype conversion in S. flexneri with plasmid constructs .................... 79
5.3.2 Serotype conversion with gtrX and bgt ............................................ 79
5.3.3 Serotype conversion using E. coli K-12 DH5α ORF3 homologue, o306 .... 80
5.4 SEROTYPE CONVERSION IN S. FLEXNERI RFC MUTANTS ......................... 80
5.4.1 rfc mutants of S. flexneri strains of serotype Y and X .......................... 81
5.4.2 Serotype conversion of spontaneous rfe mutants ................................ 81
5.5 DISTRIBUTION OF BGT AND GTRII IN S. FLEXNERI ............................ 82
5.6 COMPARISON OF GENOME ORGANISATIONS IN PHAGE P22 AND E. COLI K-12 83
5.7 SUMMARY ............................................................................................... 83

6. CHARACTERISATION OF S. FLEXNERI TYPE II NEGATIVE MUTANTS AND BGT PROTEIN ANALYSES ............................................................... 85

6.1 INTRODUCTION ...................................................................................... 85

6.2 RESULTS ............................................................................................... 86
6.2.1 Isolation of type II antigen negative mutants of S. flexneri ..................... 86
6.2.2 Characterisation of type II antigen negative derivatives of RMM254 ...... 86
  6.2.2.1 Inverse PCR of TnphoA insertion in RMA901 .................................. 88
6.2.3 Complementation studies ................................................................. 89
  6.2.3.1 LPS and western immunoblot analyses of complemented mutants RMA901 and RMA903 ................................................................. 90
6.3 DETECTION OF BGT AND GTRII GENE PRODUCTS ................................. 91
  6.3.1 Cellular location of Bgt in E. coli ..................................................... 92
  6.3.2 Studies on the topology of Bgt using phoA and lacZ fusions ............... 92
  6.3.2.1 Tm-predict model of Bgt topology ............................................ 93
  6.3.2.2 Western immunoblotting of Bgt::PhoA and Bgt::LacZ fusions .... 94
6.4 SUMMARY AND CONCLUSIONS ............................................................. 95

7. DISCUSSION ............................................................................................ 96
7.1 INTRODUCTION ..................................................................................... 96
7.2 BACTERIOPHAGE SFII AND OTHER SEROTYPE DETERMINING BACTERIOPHAGES ................................................................. 97
7.3 LYSOGENS OF BACTERIOPHAGE SFII .................................................. 99
7.4 GENOME MAPPING OF BACTERIOPHAGE SFII .................................... 100
7.5 BACTERIOPHAGE SFII ATTACHMENT SITE ........................................ 101
7.6 A TWO STEP PATHWAY FOR SEROTYPE CONVERSION ....................... 102
7.7 DPM1 - DOLICHOL PHOSPHATE MANNOSYL SYNTHASE .................. 103
7.8 COMPLEMENTATION ANALYSES ........................................................... 104
7.9 COMPARISON OF ORGANISATION OF GENOMES ............................... 106
7.10 BGT AND GTRII PROTEINS ............................................................... 106
7.11 ADDITION OF MODIFICATION ............................................................ 107
7.12 CONCLUSIONS ..................................................................................... 108

8. BIBLIOGRAPHY ....................................................................................... 110
CHAPTER ONE

INTRODUCTION

1.1 Introduction

“Dysentery” was first used by Hippocrates to describe a condition characterised by the frequent passing of stools containing blood and mucus, accompanied by straining and painful defecations. Shigellosis is a localised, ulcerative infection of the colon whose symptoms include fever, abdominal pain and dysentery, and in its most severe form can be fatal, particularly in young children.

1.2 Discovery of *Shigella* species

Dysentery, prior to 1889, had only been attributed to an amoebic cause. An alternative pathogen was first identified when Shiga found he was able to agglutinate the bacillus isolated from the stools of a patient suffering from dysentery, with serum from the same patient. This dysenteric bacillus was initially named *Bacillus dysenteriae*.

Two years later, Flexner discovered local varieties of the dysenteric bacillus on the Philippine islands (Flexner, 1900), however, Shiga’s bacillus was found to be distinct from Flexner’s bacillus, initially, on the basis of sugar fermentation. Flexner’s bacillus fermented both mannose and maltose whilst Shiga’s bacillus was unable to ferment mannose and also appeared to be more pathogenic.
1.2.1 Early serotyping of *Shigella flexneri* species

"The opportunity presented by the recent war for a comprehensive study of the dysentery group has been unrivalled" (Andrewes and Inman, 1919).

Up until the first world war, laboratory strains of Shiga’s and Flexner’s bacilli were few and serological data based on these limited strains was not very definitive. During the war, researchers received strains of dysenteric bacilli from France, Italy, Malta, Singapore, England and the Philippines; it was the first time representative strains from all parts of the world could be compared. Most of the isolates were from the faeces of patients suffering from bacillary dysentery. Little or no cross-reactivity was seen between Flexner’s bacillus and Shiga’s bacillus, although a high rate of co-agglutination between the isolates of Flexner’s bacillus was observed.

In 1919, Andrewes and Inman published a comprehensive study of the serological classification of Flexner’s bacillus. The initial classification of this organism was carried out using 21 bacilli and 15 antisera. Agglutination patterns using these limited reagents led to the conclusion that at least 4 antigenic components were present, which could be divided into the groups (or races) V, W, X and Z. Two intermediates were also seen which appeared to be related to more than one of the antigenic groups, and were therefore designated VZ and WX. A potential fifth race, Y, which continued to agglutinate with its immune serum even after absorption with any of the other 4 races, also existed. Andrewes and Inman decided that Y was not to be granted a “race” status and was classified as the most primitive of all the races.

The conclusions drawn from this study were that four races existed: V, W, X and Z, and while each of these comprised a major antigen in addition, they also had minor antigens of each of the other races (Fig. 1.1).
Figure 1.1 Pie diagram of *Shigella flexneri* serotypes V-Z

Major antigens correspond to the patterns in the large area of the circle, group antigens comprise the smaller areas. Each serotype was thought to contain a proportion of antigen of every other serotype (race).

Diagram reproduced from Boyd (1940)
Antigenic Structure of Flexner Group Bacilli
(After Andrewes & Inman)

FLEXNER V.
FLEXNER W.
FLEXNER X.
FLEXNER Z.
FLEXNER Y.
1.2.2 Work of J.S. Boyd

Subsequent work by Major J.S. Boyd revealed a great deal more about the serotypes of *B. dysenteriae* Flexner strains. Boyd revised the work conducted by Andrewes and Inman and added two more groups to the 4 races they identified. Boyd concluded that, in fact, each race possessed a distinguishing “type” antigen and a “group” antigen which may be shared by a race of a different type (Boyd, 1936; 1938; 1940). This explained the co-agglutination initially observed by Andrewes and Inman. Boyd proposed that the V-Z nomenclature be replaced with the type antigen indicated by Roman numerals and the group antigens by Arabic numerals (Boyd, 1940). The W-Z series was now known as types I, II, III, IV, V and VI. (Table 1.1)

An antigen common to all groups, (Boyd, 1940), was named group antigen 1 (Wheeler, 1944a and b). Wheeler determined that the group antigens of the *B. dysenteriae* Flexner group could in fact be separated into 9 different antigens (1-9). This system forms the basis of the typing nomenclature used currently. By the end of the 1950’s the *Shigella* Commission (1950) and the International *Enterobacteriaceae* Subcommittee (1954 and 1958) had established a universally accepted nomenclature and classification for all the *Shigella* species based on their serology (Table 1.1).

1.3 *Shigella* species

The *Shigella* species can be divided into 4 groups; A- *dysenteriae*; B- *flexneri*; C-*boydii* and D-*sonnei*. The subdivision is made on the basis of serological and biochemical tests (Edwards and Ewing, 1972; Ewing and Lindberg, 1984). *S. dysenteriae* is recognised as the major cause of epidemic dysentery, due to the toxin it produces which has enterotoxic, cytotoxic and neurotoxic activity (Keusch *et al.*, 1972; Brown *et al.*, 1982; Eiklid and Olsnes,
Table 1.1 Nomenclature of *Shigella flexneri*

Reproduced from Edwards and Ewing (1986)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S. flexneri</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>1a</td>
<td>1a</td>
<td>1</td>
<td>1:4</td>
<td>1</td>
<td>I</td>
<td>Y</td>
</tr>
<tr>
<td>1b</td>
<td>1b</td>
<td>1b</td>
<td>1</td>
<td>1:4</td>
<td>1</td>
<td>II</td>
<td>VZ</td>
</tr>
<tr>
<td>2a</td>
<td>2a</td>
<td>2a</td>
<td>2b</td>
<td>2b</td>
<td>II: 7, 8, 9</td>
<td>IIII</td>
<td>WX</td>
</tr>
<tr>
<td>3a</td>
<td>3a</td>
<td>3a</td>
<td>3</td>
<td>III: 6, 7</td>
<td>III6</td>
<td>III</td>
<td>Z</td>
</tr>
<tr>
<td>III: (2, 4), 6, 7, 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b</td>
<td>3b</td>
<td>3b</td>
<td>4a</td>
<td>IV</td>
<td>IV: 4</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>III: (2, 4), 6</td>
<td></td>
<td></td>
<td>4b</td>
<td>IV</td>
<td>IV: 6</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>4a</td>
<td>4a</td>
<td>IV</td>
<td></td>
<td></td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>4b</td>
<td>4b</td>
<td>4b</td>
<td>IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5a</td>
<td>(V, 3, 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5b</td>
<td>5b</td>
<td>5b</td>
<td>5</td>
<td>V</td>
<td>V: 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>VI: 7, 8, 9</td>
<td></td>
<td>VI: 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>VI: 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>VI: 6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ten serotypes of *S. dysenteriae* have been identified. *S. flexneri* has been associated with endemic shigellosis and 12 serotypes have been identified, whereas *S. boydii* has 15 recognised serotypes. *S. sonnei* demonstrates phase variation; phase I exhibits smooth colony form and phase II is intermediate between smooth and rough, the difference being due to the presence or absence of a plasmid encoding O-antigen (Kopecko *et al.*, 1980).

*Shigella* and *Escherichia* share a large proportion of sequence similarity and are closely related based on biochemical reactions, serological cross reactions and amino acid sequence similarity in proteins (Brenner *et al.*, 1973). *S. flexneri* and *E. coli* K-12 have genome sizes of 2.56 x 10⁹ daltons, whilst *S. dysenteriae* is about 5-10% larger, *S. boydii* is about 10% smaller. Certain *E. coli* strains cause clinical illness indistinguishable from shigellosis, with enteroinvasive *E. coli* (EIEC) cross-reacting with *Shigella* antisera (Levine and Edelman, 1984) and displaying similar mechanisms of pathogenesis. Enterohemorrhagic *E. coli* (EHEC) also produce cytotoxins that resemble the Shiga toxin of *S. dysenteriae* (Konowalchuk *et al.*, 1978; Pupo *et al.*, 1997). However, the two can be differentiated on the basis of serological and biochemical tests.

### 1.3.1 Epidemiology

Approximately 5 million people per year die of diarrhoeal diseases; 10% of these are caused by bacillary dysentery or shigellosis (Rohde, 1984). Of the four species, *S. dysenteriae* and *S. flexneri* are the predominant types seen in developing countries such as Central Africa, Southeast Asia and India. *S. dysenteriae* type 1 is responsible for most epidemic outbreaks that occur whilst *S. flexneri* is the predominant species in endemic areas (Bennish *et al.*, 1990). Shigellosis can also be a problem in developed countries, especially in environments such as day-care centres, hospitals and institutions. In this instance *S. sonnei* prevails as the dominant species. The mode of transmission is either by the faecal-oral route.
or by contaminated foods (Black et al., 1978). The incidence of dysentery was also found to be correlated with the density of houseflies. Flies were found to be responsible for the transmission of *Shigella* (and *E. coli*), resulting in diarrhoea (Cohen et al., 1991). Measures taken to control populations of flies resulted in a decrease in cases of dysentery (Levine and Levine, 1991).

1.4 **Lifestyle of *S. flexneri***

*S. flexneri* is an intracellular bacterial pathogen that has developed mechanisms to facilitate its phagocytosis by host cells, multiply and survive intracellularly, move within the cytoplasm and invade adjacent cells.

*Shigella* enters the intestinal epithelial cells via specialised structures known as M cells (Fig. 1.2) (Wassef et al., 1989; Zychlinksky et al., 1992; Perdomo et al., 1994). M cells possess an invaginated surface which is covered with less mucous than other structures present in the intestine. This provides a good environment for interaction between the pathogen and the host macrophages and polymorphonuclear cells (PMN) which inhabit the central pocket of the M cell. M cells are located within Peyers patches, found primarily within the small intestine, although equivalent structures exist in the colon (Bye et al., 1984). After entry into M cells, the organism is engulfed by a macrophage and encapsulated within a phagolysosome, from which the organism escapes shortly after. Once in the cytoplasm, multiplication occurs followed by spread to the neighbouring cells.

1.5 **Genetics of *Shigella* virulence**

The similarity between *Shigella* spp. and *E. coli* K-12 was utilised in the transfer of chromosomal regions by conjugation from one species to the other. Several chromosomal loci
Figure 1.2 Infection of intestinal cells by *Shigella*.

Entry occurs via the M cells where they are phagocytosed by macrophages.

The macrophages are lysed, the bacteria escape and move toward the basolateral surface of the enterocytes.

Inside the epithelial cell, lysis from the vacuole occurs, followed by intracellular multiplication and intercellular spread.

Intercellular spread occurs via the formation of actin tails which propel the bacterium toward the neighbouring cell.

Diagram reproduced from Siebers and Finlay (1995)
have been identified which are involved in the virulence of *Shigella* (Formal *et al.*, 1970; Sansonetti *et al.*, 1983). It was already known that invasion of epithelial cell was required to initiate infection (LaBrec, 1964).

The discovery that the invasive ability of *S. flexneri* strains was plasmid encoded (Sansonetti *et al.*, 1982) enabled research into the actual mechanism of pathogenesis of the organism to progress.

Strains containing the virulence plasmid were found to be capable of invading HeLa cell monolayers and evoking keratoconjunctivitis in guinea pigs (Sansonetti *et al.*, 1982; 1983). Specifically, 31 kb of the 220 kb plasmid (Sasakawa *et al.*, 1988) were found to encode three essential loci: *ipa* (invasion plasmid antigen), *mxi* (membrane expression of invasion plasmid antigen) and *spa* (surface presentation of invasion plasmid antigen) genes. The plasmid is unstable and its loss also correlates with loss of congo red absorption by the organisms (Maurelli *et al.*, 1984a; Schuch and Maurelli, 1997). Chromosomal elements also contribute to virulence either by regulating genes located on the plasmid or by encoding essential structural components.

1.5.1 Chromosomal loci

1.5.1.1 *virR*, a central regulatory locus

The invasive phenotype of *S. flexneri* has been shown to be temperature dependent (Maurelli *et al.*, 1984b), with organisms observed to be invasive at 37°C but not at 30°C. The chromosomally encoded *virR* is responsible for the temperature regulation and is located at 27 minutes between *galU* and *trp* (Maurelli and Sansonetti, 1988b). *virR* is homologous to the *E. coli* gene *hns* (Hromockyj and Maurelli, 1989a; Hromockyj *et al.*, 1992) which regulates the virulence genes by binding to *virF*, inducing supercoiling (Maurelli *et al.*, 1985; Maurelli and Sansonetti, 1988b; Dorman *et al.*, 1990; Maurelli, 1990), and thus repressing transcription
(Tobe et al., 1993). *virF* is a positive regulator of *icsA (virG)* and *virB* which activates transcription of the *ipa* genes (Adler et al., 1989; Tobe et al., 1991). Recently, it has been reported that expression of the *virF* and *virB* genes may account for the plasmid instability observed (Schuch and Maurelli, 1997).

1.5.1.2 Virulence associated chromosomal (*vac*) genes

*vacB* is a chromosomal gene located near *purA* on *NotI*-B (Okada et al., 1991 a and b) (Fig. 1.3) which affects the production of the Ipa antigens, IpaB, IpaC and IpaD as well as IcsA (VirG) (Tobe et al., 1992). Although, in *vacB* mutants, the production of these antigens has been affected, the levels of mRNA of *ipaB, ipaC* and *ipaD* and *virF, virB* and *virG (icsA)* were normal. It is proposed that VacB acts at the post-transcriptional level, since it likely that it is cytoplasmically located. The mechanism of action of VacB has not yet been determined, however, it may act to stabilise the Vir (VirF and VirB) proteins or be involved in their assembly (Tobe et al., 1992).

*vacJ* is also a chromosomal gene, that maps to the J segment of the *NotI* map of *S. flexneri* (Okada et al., 1991a) (Fig. 1.3) and appears to be essential for lysis of the double membrane once the organism has invaded the adjacent cell. *vacJ* mutants are capable of forming membrane protrusions containing bacteria, but are unable to move from the protrusions into the adjacent cells (Suzuki et al., 1994). IcsB has been identified that is responsible for the lysis of this double membrane (Allaoui et al., 1992), and as with VacJ, a mechanism has not yet been determined. VacJ was found to be a lipoprotein which is expressed on the bacterial surface and may either be directly involved in lysis of the membrane or may act to mediate the transport of another protein which lyses the membranes (Suzuki et al., 1994).
Figure 1.3  

_NorI_ restriction map of the virulence plasmid of _Shigella flexneri_ strain YSH6000.

The circularised assignment of the eight virulence- associated loci determined by Okada _et al._, 1991b (outside map). The other virulence- associated loci reported in other studies are shown inside the map; _virR_ (Maurelli and Sansonetti, 1988a), _osmZ_ (Dorman _et al._, 1990), _rfb_ (Sansonetti _et al._, 1983), _ompR_ (Bernardini _et al._, 1990), _sodB_ (Franzon _et al._, 1990) and _iuc_ (Nassif _et al._, 1987). Figure is modified from Okada _et al._ (1991b) and (Rajakumar _et al._, 1994). Two other virulence loci, VacB and VacJ, have been mapped to _NorI_ fragments B and J, respectively (Tobe _et al._, 1992, Suzuki _et al._, 1994).
S. flexneri 2a
(YSH6000)
4592 kb
1.5.1.3 Outer membrane protein C (OmpC)

*E. coli* responds to changes in the osmotic strength of the environment by changing the ratio of outer membrane proteins OmpF and OmpC (Lugtenberg *et al*., 1976; van Alphen *et al*., 1977). The two component regulatory system, *envZ lompR*, contributes to the transcriptional regulation of *ompC* and *ompF* (Csonka, 1989; Bernardini *et al*., 1990) by the action of EnvZ, a transmembrane osmolarity sensor, which can phosphorylate OmpR, which acts directly on these genes.

Larger pores are formed by OmpF and predominate at low osmotic strength whereas the smaller pores produced by OmpC predominate at high osmotic strength. Mutations in OmpF did not affect the virulence of the organism, however an *ompC* mutation resulted in reduction of virulence. OmpC deficient mutants were able to lyse the phagocytic vacuole, multiply and move intracellularly by polymerising actin, however, they were unable to form extracellular protrusions. The cytoskeletal rearrangements that are required for bacteria to produce these protrusions, and hence invade adjacent cells, did not occur. OmpC may therefore be involved in the rearrangements of vinculin and fimbrin that are essential for intercellular spread (Bernardini *et al*., 1993).

1.6 Plasmid mediated functions

During *Shigella* infections, antibodies produced by the humoral immune response are predominantly directed to the virulence plasmid Ipa antigens (Oaks *et al*., 1986). The *ipa* region of the virulence plasmid contains 9 genes; *icsB, ipgA, ipgB, ipgC, ipgD, ipaB, ipaC, ipaD* and *ipaA*. The three *ipg* and the four *ipa* genes belong to the same transcriptional unit (Fig. 1.4) (Allaoui *et al*., 1992; Baudry *et al*., 1988; Sasakiwa *et al*., 1989). Mutations in any of the three *ipa* genes, *ipaB, ipaC* or *ipaD*, result in loss of invasive ability, including escape
Figure 1.4  Schematic map of the 31 kb virulence region on the *S. flexneri* 230 kb plasmid and unlinked virulence genes *virF* and *icsA*.

The orientation of transcription is indicated by horizontal arrows, and the locations of promoters (P) are marked. The vertical arrows represent the points of input for the positive regulatory proteins VirF and VirB.

Diagram has been modified from Porter and Dorman (1997)
from the phagosome, and lack of even low levels of actin polymerisation (Ménard et al., 1993).

1.6.1 Adhesion of bacteria to cell surface

Release of the Ipa proteins, via the Mxi/Spa type III secretion system, into the external medium has been demonstrated by many groups (Mills et al., 1988; Hromockyj and Maurelli, 1989b; Andrews et al., 1991; Andrews and Maurelli, 1992; Venkatesan et al., 1992; Allaoui et al., 1993; Sasakawa et al., 1993; Watarai et al., 1995). The IpaB, IpaC and IpaD proteins, have been shown to be released upon contact of the pathogen with the host cell (Watarai et al., 1995), in particular, the $\alpha_5\beta_1$ integrin molecules on the epithelial cell surface (Watarai et al., 1996). This interaction, in turn, results in the phosphorylation of an integrin regulated focal adhesion kinase (125FAK) which may lead to the cytoskeletal rearrangements required to endocytose the pathogen (Watarai et al., 1996). Interaction of bacteria with components of the extracellular matrix, such as fibronectin, laminin or collagen type IV also stimulates release of the Ipa proteins (Watarai et al., 1995).

Focal adhesion kinase (FAK) is a non-receptor protein-tyrosine kinase (PTK) that associates with integrin receptors and participates in extracellular matrix-mediated signal transduction events. (Schlaepfer et al., 1994).

1.6.2 Entry by induced phagocytosis

Released Ipa proteins, IpaB and IpaC form complexes with each other (Ménard et al., 1994a, b; Ménard et al., 1996). The IpaB/IpaC complex is responsible for initiating the appropriate cytoskeletal rearrangements required to facilitate bacterial uptake. IpgA acts as a chaperone in the cytoplasm to prevent the premature association of IpaB with IpaC (Ménard et
This phenomenon has recently been shown to be growth phase dependent, with bacteria in early exponential phase being six times more invasive than those in stationary phase (Mounier et al., 1997).

Apoptosis mediated by *S. flexneri*, in particular by IpaB, has been documented (Zychlinksky et al., 1992; 1994). However, recently, it has been found that IpaB binds to interleukin-1 beta-converting enzyme (ICE) or a highly homologous protease (Thirumalai et al., 1997). Members of the ICE family proteases are important for cell death by apoptosis (Nicholson et al., 1995; Henkart, 1996). This family of proteases converts the inactive pro-form to an active protease by specific processing at aspartic acid residues (Thornberry et al., 1994). This is possibly the mechanism of action of IpaB (Thirumalai et al., 1997).

### 1.6.3 Escape from the phagocytic vacuole

Escape from the phagocytic vacuole occurs quickly (Sansonetti et al., 1986) in comparison with other intracellular organisms such as *Salmonella* and *Listeria*. IpaB was found to be responsible for lysis of the phagocytic vacuole and therefore possesses cytotoxic functions in addition to being involved in bacterial entry into macrophages (High et al., 1992).

### 1.6.4 Intracellular multiplication and spread

*Shigella* can multiply within the cytoplasm of cells (Sansonetti et al., 1986) after lysis of the phagocytic vacuole has occurred. Once in the cytoplasm, bacteria associate with stress fibres and move toward the cap of the nucleus by a process called organelle-like movement (Olm) (Vasselon et al., 1991).

Host actin filaments are then recruited and polymerised into a unipolar tail (Bernardini et al., 1989; Kadurugamuwa et al., 1991; Prevost et al., 1992). IcsA (VirG) is required to
form the actin tail which is necessary for the ordinarily non-motile shigellae to propel themselves (Bernardini et al., 1989; Lett et al., 1989; Makino et al., 1986; Vasselon et al., 1991; 1992).

Two plasmid encoded genes, virA and virK, have been identified which affect the ability of the organism to spread intercellularly. virA mutants have a decreased invasive capacity and expression of virG (icsA), however, production of the Ipa proteins is not affected. VirA was found to be secreted by the Mxi/Spa type III secretion system and is under the control of virB (Uchiya et al., 1995). virK mutants, also reduced in their ability to spread intercellularly, have a greatly reduced level of VirG (IcsA). virK is highly conserved across Shigellae and enteroinvasive E. coli and is thought to play a role in the post-transcriptional regulation of virG (Nakata et al., 1992).

1.6.5 Lysis of double membrane surrounding protrusion

The IcsA mediated propulsion moves the organisms through the cytoplasm toward the cell membrane. The membrane- bound bacteria move toward the basolateral surface of the adjacent cell (Mounier et al., 1992) where the filopods are engulfed by the neighbouring cells, resulting in the organisms becoming enclosed within a double membrane. Lysis of this double membrane occurs via the action of the gene product of icsB (Allaoui et al., 1992) and the chromosomally encoded vacJ (Suzuki et al., 1994). Once the organism is released into the cytoplasm of the new host cell, the entire process begins again.

1.6.6 Biology of IcsA

IcsA has been extensively studied and has been found to be expressed in highest concentration during exponential phase (Goldberg et al., 1994). This molecule localises to
one bacterial pole, and is found as a 120 kDa outer membrane protein (Goldberg et al., 1993). IcsA is also found as a 95 kDa secreted protein and it was previously thought that this form was associated with the actin tail (Goldberg et al., 1993). This was later found to be a 70 kDa protein also present in uninfected host cells identified by a monoclonal antibody recognising glycine-rich boxes (D’Hauteville et al., 1996).

Recently, a protease was discovered responsible for the cleavage of IcsA by recognising a specific sequence SSRRASS, located at nucleotides 754-760, in the 120 kDa form. This protease, SopA (Shigella outer membrane protease), cleaves between the two arginine residues (R) and exhibits homology to the outer membrane proteins, OmpT and OmpP (Egile et al., 1997). These results were confirmed by Goldberg and coworkers (Shere et al., 1997) who also identified a protease, IcsP, which functions in the same way and is identical to SopA.

Previously, it was shown that when OmpT was introduced in Shigella it led to lack of ability of that strain to spread intercellularly (Pal et al., 1989; Nakata et al., 1993), due to the total degradation of IcsA. It has been hypothesised that SopA functions in the same manner as OmpT, which also cleaves between arginine residues (Grodberg and Dunn, 1988; Sugimura and Nishihara, 1988). SopA may, therefore, be produced at a lower level than OmpT or at a lower specific activity, resulting in cleavage of non-polarly located IcsA from the surface of the bacterium. It had also been noted that IcsA in the 120 kDa form was sufficient to elicit actin polymerisation (Fukuda et al., 1995; Goldberg and Theriot, 1995). Lack of cleavage was shown to result in circumferential localisation of IcsA, which leads to aberrant movement (D’Hauteville et al., 1996). The distribution of IcsA has also been found to be affected by the lipopolysaccharide (Sansonetti et al., 1983; Okada et al., 1991a; Sandlin et al., 1995; 1996; Van den Bosch et al., 1997).
1.7 Lipopolysaccharide (LPS)

Lipopolysaccharide is a major component of the outer membrane of Gram-negative organisms and consists of three structural regions: lipid A, core oligosaccharide and O-antigen (Fig. 1.5). The lipid A component is responsible for anchoring the LPS to the bacterial cell membrane and also activates the immune system (Galanos et al., 1984; 1985, Kotani et al., 1984), whereas the O antigen imparts serological specificity to the organism. The biosynthesis of LPS has been extensively studied in S. enterica, E. coli and Shigella (Nikaido et al., 1961; Johnston et al., 1968; Mäkelä and Stocker, 1984; Raetz, 1993, Reeves, 1993, Schnaitmann and Klena, 1993) and will be described in the next section.

Three LPS phenotypes have been described; rough (R), semi-rough (SR) and smooth (S). Each displays specific bacteriophage sensitivity patterns. These phenotypes can also be differentiated by silver staining of LPS samples electrophoresed on SDS polyacrylamide gel electrophoresis (SDS-PAGE) gels. R-LPS arises from mutations in the core biosynthetic genes, which results in lipid A-core with no O-antigen attached, whereas SR-LPS has one O antigen unit attached and migrates slightly slower than R-LPS. SR-LPS arises from a mutation in the O antigen polymerase gene (Naide et al., 1965; Morona et al., 1994). The S-LPS, or wild type, phenotype is associated with long O antigen chains attached to the lipid A-core oligosaccharide. A S-LPS ladder is observed on SDS-PAGE gels, indicating that each individual O antigen chain linked to core is of different chain length (Palva and Mäkelä, 1980). A smooth lipopolysaccharide (S-LPS) is required for the virulence phenotype of Shigella (Formal et al., 1970; Gemski et al., 1972; Sansonetti et al., 1983; Hale et al., 1984; Okada et al., 1991a; Sandlin et al., 1995; 1996; Van den Bosch et al., 1997) and for protecting the organism from damage induced by the non-specific defense mechanisms of the host (Hong and Payne, 1997).
**Figure 1.5** Schematic structure of lipopolysaccharide (LPS).

LPS is comprised of three components; lipid A, core oligosaccharide and O-antigen.

Diagram has been modified from Lindberg *et al.*, (1991)
1.7.1 Biosynthesis of LPS

The genetics of LPS biosynthesis have been extensively studied and the location of the genes involved have been determined. Genes involved in lipid A biosynthesis are scattered throughout the chromosome, as many also have a role in other housekeeping functions (Crowell et al., 1987). The old nomenclature for the LPS biosynthetic genes is to be used here (Reeves et al., 1996).

The genes determining core oligosaccharide biosynthesis are found mostly in the rfa region (Formal et al., 1970; Petrovskaya and Licheva, 1982; Schnaitman and Klena, 1993), whereas the rfb locus encodes for enzymes uniquely involved in the biosynthetic pathway of the nucleotide sugars and the various transferases required for assembly of the O antigen unit (Reeves, 1993).

LPS biosynthesis proceeds via two separate pathways, with lipid A and core synthesis combined, and O antigen synthesis occurring independently (Fig. 1.6). The biosynthetic pathway can be summarised: 1) the core oligosaccharide is synthesised on lipid A; 2) individual O-antigen units are synthesised on a lipid carrier (bactoprenol or undecaprenol phosphate); 3) individual ACL-linked O-antigen units are polymerised into long chains; 4) the O-antigen chain is ligated to lipid A–core oligosaccharide (Fig. 1.6) (Jann and Jann, 1984).

Lipid A synthesis begins in the cytoplasm and acts as the lipid carrier for assembly of the core sugars. The individual O units are synthesised onto the carrier, bactoprenol, at the cytoplasmic face of the cytoplasmic membrane (Mulford and Osborn, 1983; Marino et al., 1991; Wang and Reeves, 1994). The lipid A- core and the O units are transported across the inner membrane where the bactoprenol linked O units are polymerised into long chains in the periplasm (McGrath and Osborn, 1991). Ligation of the O antigen chain to the completed core oligosaccharide also occurs at the periplasmic face of the cytoplasmic membrane (Mulford and Osborn, 1983; McGrath and Osborn, 1991) and the completed LPS molecules
Figure 1.6 Pathway of LPS biosynthesis.

Lipid A-core are synthesised in a common pathway, where the lipid A acts as the carrier for the sequential addition of the core sugars. O-units is synthesised in a separate pathway until the first O unit is ligated onto the lipid A-core structure.

Diagram has been modified from Jann and Jann (1984)
antigen carrier lipid (ACL)

- O-polysaccharide synthesis on ACL, directed by rfb (and rfc) genes

- Translocation of O-polysaccharide to core-lipid A with recycling of ACL, directed by rfaL

- Synthesis of lipid A by unknown genes

- Core synthesis on lipid A, directed by rfa genes

- LIPID A

- CORE-LIPID A

- O-POLYSACCHARIDE-CORE-LIPID A (LIPOPOLYSACCHARIDE, O-ANTIGEN)
are transported to the outer membrane, although the mechanism for this final process is still unknown (Raetz, 1990).

1.7.2 Lipid A- genetics and biosynthesis

Lipid A has been found to be responsible for the endotoxic effects of LPS (Kadis et al., 1971) and provides structural integrity to the outer membrane of Gram negative organisms (Rietschel et al., 1983; Galanos et al., 1977). Lipid A is a conserved structure composed of a β-1-6 linked di-glucosamine disaccharide (Fig. 1.7). Fatty acids are attached as O- and N-acyl substituents to the glucosamine residues. Whilst the backbone structure of lipid A does not vary greatly within Enterobacteriaeeae, the fatty acids attached to the Lipid A disaccharide differ from species to species. Phosphates are usually attached to 4' and 1 position and may serve as linkage points for phosphoethanolamine, D-glucosamine and 4-amino- 4-deoxy- L-arabinose, ethanolamine or phosphate (Schnaitman and Klena, 1993).

The biosynthetic pathway of lipid A was completed upon the discovery of E. coli K-12 mutants deficient in phophatidylglycerol, which led to the elucidation of the entire pathway (Ray et al., 1984; Nishijima et al., 1981a and b). The precursors of lipid A biosynthesis are UDP-GlcNAc and R-3-Hydroxymyristoyl-acyl carrier protein which are also incorporated in peptidoglycan and glycerophospholipids, respectively. Initially, an acyl group is transferred from R-3-Hydroxymyristoyl-ACP to UDP-GlcNAc, mediated by LpxA, resulting in an O-acylated UDP-GlcNAc (Anderson and Raetz, 1987; Coleman and Raetz, 1988). The next step involves the N deacetylation of O-acylated UDP-GlcNAc as well as the transfer of a second acyl group from R-3-Hydroxymyristoyl-ACP to form UDP-2,3-diacylglucosamine, carried out by LpxC and LpxD (Crowell et al., 1987). UDP is then removed by one of two pyrophosphatases resulting in the formation of 2,3-diacylglucosamine-1-phosphate also referred to as lipid X (Raetz, 1987, 1990) (Fig. 1.7).
Figure 1.7  Biochemical pathway for the biosynthesis of lipid A and mature lipopolysaccharide.

Key: UDP- GlcNAc, uridine diphosphate N-acetyl-D- glucosamine; ACP, acyl carrier protein (ACL); GlcN, D-glucosamine; CMP-KDO, cystidine monophosphate 2-keto-3-deoxy-D-manno-octulosonic acid (Raeiz, 1990)
UDP-GlcNAc

\[ \text{UDP-GlcNAc} \rightarrow \text{LpxA} \rightarrow \text{ACP} \]

\[ \text{ACP} \rightarrow \text{H}_2\text{O} \rightarrow \text{ACP} \]

\[ \text{ACP} \rightarrow \text{UDP} \rightarrow \text{LpxB} \rightarrow \text{UDP} \]

\[ \text{(UDP 2,3 Diacyl-GlcN)} \rightarrow \text{LpxB} \rightarrow \text{UDP} \]

\[ \text{ATP} \rightarrow \text{ADP} \rightarrow \text{(Lipid IVa)} \]
One molecule of lipid X and one molecule of UDP-2,3-diacylglucosamine are condensed by LpxB to form β1,6 linked glucosamine disaccharide (Ray et al., 1984; Crowell et al., 1987). Each sugar is substituted with two 3-hydroxymyristoyl groups and which carries a single phosphate residue. A 4' monophosphate is added generating a symmetrical disaccharide with a phosphate at each end. This molecule is known as lipid IV_A (Raetz., 1987).

The extent of substitution of lipid A by phosphates is determined by pmrA (Mäkelä et al., 1978; Vaara et al., 1979; Vaara, 1981; Helander et al., 1994). pmrA/pmrb share sequence homology with members of the two component regulatory system OmpR and PhoP (Roland et al., 1993; Helander et al., 1994). PhoP/PhoQ activate the transcription of pmrA,B, resulting in the addition of aminoarabinose to lipid A. The modified lipid A was found to alter the response in comparison with wildtype lipid A in vitro. The PhoP/Q system is used by Salmonellae for environmental sensing and regulation of virulence genes. The lipid A may be modified in different environments in vivo, depending upon the expression of the pmr A,B genes (Guo et al., 1997)

1.7.3 Core region

The structure of the core component has been highly conserved across species (Lüderitz et al., 1971; Jansson et al., 1981) (Fig. 1.8). In E. coli, five core types have been described (R1-R5), S. flexneri, with the exception, of Type 6 produces R3 core (Lindberg et al., 1991). Elucidation of the structure of the core oligosaccharide has been made possible by the existence of LPS mutants defective in biosynthesis. Core can be divided into two regions; inner and outer core (Fig. 1.8).
Figure 1.8  Schematic diagram of the smooth LPS of *Salmonella enterica* sv. Typhimurium

Definition of chemotype variants and genetic loci involved in the biosynthesis of the *S. enterica* sv Typhimurium LPS molecule.

Key: Glc, D-glucose; Gal, D-galactose; GlcN, D-glucosamine; GlcNAc, N-acetyl-D-glucosamine; Hep, L-glycero-D-mannose-heptose; KDO, 2-keto-3-deoxy-D-manno-octulosonic acid; AraN, 4-amino-L-arabinose; P, phosphate; EtN, ethanolamine;

Ra to Re indicate incomplete R form LPS. The arrows with the *rfa, rfb, lpx, kds* and *gal* genes show the location of the enzymatic steps encoded (or putatively encoded) by these genes. Diagram adapted from Lüderitz (1971); Hitchcock *et al.*, (1986) and Raetz (1990).
O-ANTIGEN

CORE OLIGOSACCHARIDE

LIPID A

--- Rₐ LPS ---

--- Rₐ₂ LPS ---

--- Rₜ LPS ---

--- Rₜ LPS ---

--- Rₜ LPS ---

--- Rₜ LPS ---

--- Rₜ LPS ---

--- S (wild type) LPS ---
1.7.3.1 Inner core: 3-deoxy-D-mannoctulosonic acid (KDO)

The inner core is composed of 2-3 residues of octulose, which is a sugar unique to this region (KDO) whose function is to link the lipid A component, via an α 2-6' linkage, to the core. KDO is usually linked to two L-glycero-D-manno heptose residues.

Attachment of the first two residues of KDO to lipid IV_\text{A} occurs before the final acylation of lipid A (Brozek and Raetz, 1990; Raetz, 1987; Walenga and Osborn, 1980). *kdtA* encodes the gene for a bifunctional enzyme (Belunis and Raetz, 1992) which sequentially transfers two molecules of CMP-KDO to lipid IV_\text{A}. KDO-8-phosphate is synthesised by the gene product of the *kdsA* gene (Rick and Young, 1982; Woisetschläger *et al.*, 1988) which is then used in the synthesis of CMP-KDO by KdsB (Goldman *et al.*, 1986). The addition of an α 2,6' linked CMP-KDO to lipid A generates KDO I and the first α-2,4 linked branch KDO residue. KdtA may also be responsible for the addition of subsequent 2,4 linked KDO branches, however, these are added at the stage of the completion of the core region (Rick, 1987).

1.7.3.2 Outer core

The outer core is composed of hexoses, such as glucose, galactose, rhamnose and galactosamine, linked to the heptose of the inner core (Fig. 1.8). Further variation in the outer core is determined by the extent of substitution by phosphate, pyrophosphate and phosphoryl ethanolamine attached to the hexose sugars.

The genes required for the biosynthesis of the core region are referred to as *rfa* genes (Table 1.2). Core is synthesised by sequential addition of sugars to the lipid A -KDO structure. *rfaE* encodes an ADP-heptose synthase and *rfaD* encodes an epimerase which
Table 1.2 *rfa* genes involved in core oligosaccharide biosynthesis in *E. coli* and *S. enterica*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein product/ function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rfaD</td>
<td>ADP-1-glycero-D-mannoheptose-6-epimerase</td>
<td>Coleman, 1983</td>
</tr>
<tr>
<td>rfaC</td>
<td>heptosyl transferase I</td>
<td>Sanderson et al., 1974; Sirisena et al., 1992</td>
</tr>
<tr>
<td>rfaF</td>
<td>heptosyl transferase II</td>
<td>Mäkelä and Stocker, 1984</td>
</tr>
<tr>
<td>rfaG</td>
<td>UDP-glucosyl transferase I</td>
<td>Creeger and Rothfield, 1979; Parker et al., 1992a, 1992b</td>
</tr>
<tr>
<td>rfaI</td>
<td>galactosyl transferase I$_6$/glucosyl transferase II$_6$</td>
<td>Creeger and Rothfield, 1979; Carstenius et al., 1990; Pradel et al., 1992</td>
</tr>
<tr>
<td>rfaJ</td>
<td>glucosyl transferase I$_6$/glucosyl transferase III$_6$</td>
<td>Creeger and Rothfield, 1979; Carstenius et al., 1990; Pradel et al., 1992</td>
</tr>
<tr>
<td>rfaB</td>
<td>galactosyl transferase I$_6$/galactosyl transferase I$_6$</td>
<td>Wellin et al., 1983; Pradel et al., 1992</td>
</tr>
<tr>
<td>rfaK</td>
<td>N-acetyl glucosamine transferase</td>
<td>Hellerqvist and Lindberg, 1971; Klena et al., 1992</td>
</tr>
<tr>
<td>rfaS</td>
<td>? rhamnosyl transferase</td>
<td>Schnaitman and Klena, 1993</td>
</tr>
<tr>
<td>rfaP</td>
<td>phosphate transferase</td>
<td>Mühlradt et al., 1968; Parker et al., 1992a, 1992b</td>
</tr>
<tr>
<td>rfaL</td>
<td>O-antigen ligase</td>
<td>Cynkin and Osborn, 1968; Klena et al., 1992</td>
</tr>
<tr>
<td>rfaQ</td>
<td>putative transferase</td>
<td>Klena et al., 1992</td>
</tr>
<tr>
<td>rfaZ</td>
<td>? core completion?</td>
<td>Maclachlan et al., 1991; Klena et al., 1992</td>
</tr>
<tr>
<td>rfaY</td>
<td>? core completion, possibly associated with rfaJ</td>
<td>Pradel et al., 1992</td>
</tr>
<tr>
<td>rfaI5</td>
<td>putative regulator</td>
<td>Roncero and Casadaban, 1992</td>
</tr>
<tr>
<td>kdtA</td>
<td>KDO-transferase</td>
<td>Clementz and Raetz, 1991; Belunis et al., 1992</td>
</tr>
<tr>
<td>rfaH</td>
<td>? antiterminator or transcriptional activator</td>
<td>Lindberg and Hellerqvist, 1980; Stocker et al., 1980; Pradel and Schnaitman, 1991</td>
</tr>
<tr>
<td>rfaE</td>
<td>heptosyl transferase I</td>
<td>Kuo and Stocker, 1972; Sirisena et al., 1992</td>
</tr>
<tr>
<td>galE</td>
<td>UDP-galactose epimerase</td>
<td>Nikaido, 1962; Stocker et al., 1980</td>
</tr>
<tr>
<td>galU</td>
<td>UDP-glucose pyrophosphorylase</td>
<td>Nakae and Nikaido, 1971</td>
</tr>
</tbody>
</table>

1 St, *Salmonella enterica* sv Typhimurium; Ec, *Escherichia coli* K-12
converts ADP-D-glycero-D-manno-heptose to ADP-L-glycero-D-manno-heptose (Kuo and Stocker, 1972; Coleman and Leive, 1979; Coleman, 1983; Mäkelä and Stocker, 1984; Pegues et al., 1990; Sirisena et al., 1992). \(rfaC\) (Sanderson et al., 1974; Sirisena et al., 1992) and \(rfaF\) (Sanderson and Saeed, 1972; Sanderson et al., 1974; Mäkelä and Stocker, 1984) both encode heptosyl transferases whose gene products add \(\alpha 1,5\) linked backbone heptose residues (Hep1) to KDO1 and HepII respectively (Wilkinson et al., 1972). Glucose and galactose are sequentially attached to the inner core to provide the structural framework for the outer, or hexose (Fig. 1.8), region of the core by the products of genes \(rfaGBIJ\) (Creeger and Rothfield, 1979; Carstenius et al., 1990; Pradel et al., 1992; Wollin et al., 1983) whilst \(rfaK\) (Hellerqvist and Lindberg, 1971; Klena et al., 1992) encodes a protein which is involved in the addition of GlcNAc to the terminal core glucose in \(S.\ typhimurium\) and a substituent to a different core site in \(E.\ coli\) K-12.

### 1.7.4 O-antigen

The great diversity of LPS is due to its O-antigen component. It is composed of variable numbers of repeating units of oligosaccharides, the content of which differs greatly within \(Enterobacteriaceae\). Much of the understanding of the biosynthesis of O-antigen has been achieved from the study of \(S.\ enterica\) sv. Typhimurium (Mäkelä and Stocker, 1984).

\(S.\ flexneri\) can be divided into approximately 12 serotypes (Fig. 1.9) based on modifications of a tetrasaccharide repeat unit; GlcNAc-Rha-Rha-Rha- (Kenne et al., 1978). These modifications are additions of either glucosyl and/ or O-acetyl groups which are mediated by lysogenisation with temperate bacteriophages.
**Figure 1.9** Schematic diagram of the O-antigens of *Shigella flexneri* serotypes.

*S. flexneri* serotypes 1a, 1b, 2a, 2b, 3a, 3b, 4a, 4b, 5a, 5b and variants X and Y;

Key: Rha, rhamnose; GlcNAc, N-acetyl-D-glucosamine;

1.7.4.1 Genetics of O antigen

The rfb gene cluster in S. enterica sv. Typhimurium is located at 40 min between the his operon and the cps genes (Jiang et al., 1991). There are 16 genes within the rfb operon; rfbBCAD, involved in the synthesis of the nucleotide sugars dTDP-rhamnose (Klena and Schnaitman, 1993; Reeves, 1993; Stevenson et al., 1994), dCDP-abequose (rfbFGJHI) (Kessler et al., 1993; Liu et al., 1993; Reeves, 1993) and dGDP-mannose (rfbMK) (Stevenson et al., 1991; Xiang et al., 1993; Sugiyama et al., 1994), four sugar transferases (rfbNUVP) (Nikaido, 1973; Mäkelä and Stocker, 1984) and rfbX, a predicted integral membrane protein (Mäkelä and Stocker, 1984; Jiang et al., 1991; Reeves, 1993; Macpherson et al., 1995).

The rfb region of S. flexneri has been sequenced and characterised (Macpherson et al., 1991; 1994). The organisation of the rfb region is similar to that of S. enterica (Reeves, 1993), S. dysenteriae (Klena and Schnaitman, 1993) and E. coli K-12 (Stevenson et al., 1994). The dTDP-rhamnose biosynthetic genes (rfbB,C,A,D) are located upstream of the genes involved in assembly (rfbX, F,G and rfc) (Fig. 1.10). A high degree of homology is seen between S. flexneri and S. enterica with rfbB, rfbC and rfbA (85%, 81%, and 89% at the amino acid level) and 72% for rfbD (Macpherson et al., 1994).

1.7.4.2 O-polysaccharide biosynthesis

Traditionally two pathways have been recognised for O-antigen biosynthesis and these have been classified according to whether or not the O antigen is polymerised by the product of the rfc gene (Jann and Jann, 1984; Klena and Schnaitman, 1993; Mäkelä and Stocker, 1984; Rick, 1987). The biosynthesis of O antigen was referred to as rfc-dependent or rfc-independent.
Figure 1.10  Map of the rfb region of *S. flexneri* from the *CIA* site within *orf0x0* through to the *his* operon. The genes and the ORFs are shown in boxes, the dTDP-rhamnose biosynthetic genes indicated.

Diagram reproduced from Macpherson *et al.*, (1994)
Rhamnosyl transferases
TDP: Rhamnose
Transport?
O-antigen biosynthesis
Rhamnosyl transferases
O-antigen polymerase
TDP-Rhamnose
0x0 galF rfb B C A D X F G rfc 10X1 10X5 11X6 12X8 gnd ugd rol
1 3 5 7 9 11 13 15 17 19 kb
EcoRI
SacI
PstI
HindIII
HindIII
EcoRI
PstI
SacI
BglII
PstI
SacI
PstI
EcoRI
PstI
BglII
SphI
SacI
pPM2213
The *rfc* dependent pathway begins with synthesis of the component sugars of the O polysaccharide in the cytoplasm, followed by their assembly into repeat units. In order for polymerisation of these repeat units, into long chains, to occur, they must be translocated across the plasma membrane to the periplasmic face. In this location, they come into contact with the O antigen polymerase, Rfc, which polymerises the repeat units into long chains which are then ligated onto the lipid A- core structure.

Synthesis via the *rfc*- independent pathway is initiated at the cytoplasmic face of the plasma membrane by the *rfe* gene product. Rfe is a N-acetylglucosamine-1-phosphate transferase which transfers a GlcNAc onto the undecaprenol-phosphate carrier to form und-P-P-GlcNAc. This structure acts as a precursor for the biosynthesis of O-antigen and ECA, the enterobacterial common antigen (discussed later). Polymerisation, in this system, occurs without Rfc; sugars of the O antigen are sequentially transferred to the non-reducing end of the polymer. Translocation across the membrane to the periplasmic face is carried out by the ATP-binding cassette transporter (ABC) (Whitfield, 1995) followed by ligation to the lipid A- core occurs.

A third pathway has recently been discovered (Keenleyside and Whitfield, 1995) which is Rfc -independent and only requires the und-P-P-GlcNAc primer synthesised by Rfe. The remaining polymerisation occurs under the direction of the *rfbA* and *rfbB* genes. RfbA adds a single N-acetyl mannosamine sugar to the und-P-P-GlcNAc precursor. RfbB then continually adds glycosyl units to the growing polysaccharide chain. RfbB belongs to a family of enzymes which are believed to have dual function; possessing both transferase and export activity. Hence, it is possible that RfbB acts as a synthase, in polymerising the growing O chain and also may be involved in the export through the plasma membrane since homologues for either RfbX or the ABC transporter system have not been found (Whitfield, 1995).
1.7.4.2.1 Enterobacterial Common Antigen (ECA)

Enterobacterial common antigen is produced by all Enterobacteriaceae (Mäkelä and Mayer, 1976; Mayer and Schmidt, 1979; Kuhn et al., 1988). It is an amphipathic molecule consisting of a hydrophilic amino sugar chain linked to a hydrophobic -L-glycerophosphatidyl residue. The sugar chain of ECA has been found to bind to the core in rough LPS mutants (Mayer and Schmidt, 1979; Ramia et al., 1982). The amino sugar chain is composed of three sugars; N-acetyl glucosamine (GlcNAc), N-acetyl mannosaminuronic acid (ManNAcA) and 4-acetamido-4,6-dideoxy-D-galactose (4-FucNAc), and is responsible for the antigenicity of the molecule (Lugowski et al., 1983).

ECA is synthesised under the direction of the rfe and rff genes (Mäkelä et al., 1970; Mäkelä and Mayer, 1976; Rick and Young, 1982; Schmidt et al., 1976) and in some species rfb genes (Mäkelä and Mayer, 1974; Mäkelä et al., 1976; Lew et al., 1986). rfe has been shown to also be involved in the biosynthesis of S-LPS in some species (Mäkelä et al., 1974; Schmidt et al., 1976; Mäkelä and Mayer, 1976; Lew et al., 1986; Klena and Schnaitman, 1993; Stevenson et al., 1994) whereas the rff locus is dedicated to ECA biosynthesis (Lew et al., 1986).

1.7.4.3 Other genes required for O antigen synthesis

With the elucidation of the pathways for biosynthesis of lipid A, core and O antigen, interest has turned to the understanding of O antigen export and polymerisation, transfer to lipid A-core and the distribution of the chain length of O antigen. In this respect, four enzymes are of most interest, RfbX (O antigen assembly or transport), Rfc (the O antigen polymerase), RfaL (the O antigen ligase) and Rol (the regulator of O antigen chain length) (Fig. 1.11).
Figure 1.11 Model for mechanism of action of RfbX, Rol and Rfc proteins.

Schematic diagram representing the pathway of LPS biosynthesis and the location of proteins involved in O-unit synthesis (RfbBCAD), and assembly (RfbX, RfbF and RfbG), ligation of the first O-unit to the LipidA-core (RfaL), polymerisation of O units (Rfc) and control of O chain length (Rol).
Rfc/Rol/RfaL/LipidA-core

Lipid A-core

O-antigen repeat unit

Rfe/RfbF/RfbG

Bactoprenol-PP-sugar

Bactoprenol

UDP-N-acetyl glucosamine

TDP-rhamnose

UMP/TDP
1.7.4.4 RfbX

The rfbX gene is predicted to encode a membrane bound protein with twelve potential transmembrane domains. The most likely role for RfbX in O-antigen biosynthesis is the export of O units to the periplasmic face of the cytoplasmic membrane (Reeves, 1993; Schnaitman and Klena, 1993; Macpherson et al., 1995; Liu et al., 1996). rfbX mutants of S. dysenteriae 1 were isolated and were found to have an accumulation of undecaprenol pyrophosphate linked O units on the cytoplasmic side of the cytoplasmic membrane, suggesting a flippase activity for RfbX (Liu et al., 1996).

1.7.4.5 RfaL, O-antigen ligase

The O antigen ligase is responsible for the transfer of the growing O antigen chain to the lipid A -core structure (Mäkelä and Stocker, 1984). RfaL proteins from E. coli K-12 and S. enterica sv. Typhimurium are structurally similar but have little sequence identity (Schnaitman and Klena, 1993). Completion of the core requires the action of RfaK, this is then able to accept the newly synthesised O chains (Schnaitman and Klena, 1993). RfaL of different species can be used to ligate the O antigen of heterologous LPS in the presence of RfaK (Schnaitman and Klena, 1993). This may be due to a requirement for a species- specific completion of the core or it may reflect an interaction between RfaK and RfaL (Whitfield et al., 1997).

1.7.4.6 Rfc, O-antigen polymerase

The rfc gene of S. flexneri is located inside the rfb region (Morona et al., 1994) as is the case for E. coli O4 (Lukomski et al., 1996) and P. aeruginosa (de Kievit et al., 1995), this
contrasts the situation in *S. typhimurium* (Naide *et al*., 1965; Stocker and Mäkelä, 1971; Collins and Hackett, 1991). *rfc* lacks its own promoter and is likely to be transcriptionally coupled with the upstream genes (Morona *et al*., 1994; Lukomski *et al*., 1996). *Rfc* proteins of various species share many characteristics; lack of an obvious promoter, the possibility of transcriptional coupling and the predominance of modulating or rare codons in the first 23 amino acids. The molecular mass of the *Rfc* proteins does not vary greatly either with *S. flexneri* (43.7 kDa), *S. enterica* sv. *Typhimurium* (47.5 kDa) and *E. coli* O4 (45.5 kDa); all are hydrophobic proteins with twelve potential transmembrane regions (Morona *et al*., 1994). Despite the functional similarity, the overall sequences are different. In *E. coli* O4, complementation of *rfc* mutations by plasmid borne copies of *rfc* does not restore the phenotype to wildtype. It may be that large amounts of mRNA exhaust the limited pool of rare tRNAs from a high copy number complementing plasmid, in contrast, a poorly transcribed low copy number plasmid would allow translation to be completed (Lukomski *et al*., 1996).

### 1.7.4.7 Rol, regulator of O antigen chain length.

LPS of many species exhibit a preference for particular chain lengths, referred to as the modal distribution (Grossman *et al*., 1987; Mills and Timmis, 1988; Goldman and Hunt, 1990. A protein, Rol (Cld), has been found that is responsible for the regulation of O-antigen chain length (Macpherson *et al*., 1991; Batchelor *et al*., 1991; Bastin *et al*., 1993; Morona *et al*., 1995). The gene encoding this protein, *rol* (*cld*) is located between *his* and *gnd*, near the *rfb* operon (Batchelor *et al*., 1991; Bastin *et al*., 1993; Klena and Schnaitman, 1993; Morona *et al*., 1995) (Fig. 1.10), however, a plasmid-borne copy has also been identified (Stevenson *et al*., 1995).
The Rol protein is anchored into the cytoplasmic membrane via its amino and carboxy terminal ends, however, the majority of the protein is located in the periplasmic space. Mutations in this protein lead to random lengths of O chains with no modal distribution as seen in wildtype strains. The mechanism by which this process occurs is the subject of debate. It has been suggested that Rol acts as a molecular chaperone which interacts with the RfaL protein to assemble a complex. This complex would give a specific ratio of RfaL to Rfc, altering the overall kinetics of the ligation reaction to give the modal chain length. In the absence of Rol, the ratio of RfaL to Rfc are not fixed and hence random interaction results (Morona et al., 1995).

An alternative proposal for the action of Rol (or Cld) is that it acts in a time dependent manner with Rfc which can exist in two functional states (Bastin et al., 1993). One state favours polymerisation and the other favours transfer to RfaL which then ligates the O chain to the lipid A core. The timing of the polymerisation over transfer is a function of Rfc and the different modal lengths observed result from an alteration in timing of the switching of states. This process is analogous to the mechanism used by ribosomes and fatty acid synthesis (Bastin et al., 1993).

Although Rol proteins can be interchanged, each Rol will still be specific for its modal length and will impart this on any O antigen. This is currently being studied in the development of heterologous vaccines (Klee et al., 1997).

1.8 Lipopolysaccharide and the distribution of IcsA

The importance of smooth LPS (S-LPS) as a virulence factor has been increasingly studied, in particular its role in the correct localisation of IcsA (Gemski et al., 1972; Okamura et al., 1983; Sansonetti et al., 1983; Rajakumar et al., 1994; Sandlin et al., 1995, 1996).
Mutations in LPS biosynthetic genes were found not to affect the ability of \textit{S. flexneri} to invade or to multiply intracellularly but did affect their ability to spread from cell to cell (Okada \textit{et al.}, 1991 a and b). They identified three classes of mutants, those with i) altered core structures, ii) no O antigen (\textit{rfb}) and iii) decreased O antigen chain length (\textit{rfc}).

More recently, the relationship between the distribution of \textit{IcsA} on the bacterial surface and the requirement for long O chains has been examined specifically (Rajakumar \textit{et al.}, 1994; Sandlin \textit{et al.}, 1995).

To date, many of the LPS mutants studied were spontaneous mutants which had not been characterised at a genotypic level. To demonstrate the role of LPS in the unipolar localisation of \textit{IcsA}, two defined mutants (\textit{rfe} and \textit{galU}) were constructed. A mutation in \textit{galU}, which acts at the inner core, does not allow the completion of the core or the addition of O antigen and results in the circumferential distribution of \textit{IcsA}. Mutations in \textit{rfe} results in the inability of the organism to synthesis O antigen and produces an intermediate phenotype, with some polar distribution and some \textit{IcsA} on the sides of the bacterial body. The \textit{rfe} mutant unlike the \textit{galU} mutant produced small plaques on monolayers, whereas the \textit{galU} mutant produced no plaques at all (Sandlin \textit{et al.}, 1995). It was also demonstrated that the sugar composition (ie. glucosylation) of the O side chain did not affect the localisation of \textit{IcsA} on the surface, unlike the need for long O chains (Sandlin \textit{et al.}, 1996).

Mutations in \textit{rol} and \textit{rfbD} were constructed in \textit{S. flexneri} of serotype 2a (2457T) by insertional inactivation using a kanamycin cartridge (Van den Bosch \textit{et al.}, 1997). These mutants were examined with respect to distribution of \textit{IcsA} and virulence. It was observed that the \textit{rol} mutation affected the intra- and inter- cellular spreading ability of the organisms, and the expression of \textit{IcsA} on the cell surface. \textit{rol} mutants produced little or no \textit{IcsA} with approximately 10% resembling the parent strain. The properties of the \textit{rfbD} mutant were comparable to the \textit{galU} mutant (Sandlin \textit{et al.}, 1995) however, structurally it resembles the \textit{rfe} mutant. Both mutants were unable to form plaques in HeLa cell monolayers, however, some
Factin tail formation was observed. The rfbD mutant had IcsA located over its entire surface, although it was more concentrated at one end (Van den Bosch et al., 1997).

1.9 Chemical understanding of O-antigen of S. flexneri

Initially characterisation of the structure of S. flexneri O units was carried out using O polysaccharide purified by acid hydrolysis (Simmons, 1969). These polysaccharides were then subjected to chromatography and electrophoresis to determine the components comprising the repeat unit. A ratio of 1:1:2 of the hexose sugars, glucose, N-acetyl glucosamine and rhamnose was determined (Simmons, 1971). However, it was not until Kenne and colleagues carried out methylation studies coupled with techniques such as $^1$H-NMR and $^{14}$C-NMR that the actual composition of the O unit was elucidated (Lindberg et al., 1973; Kenne et al., 1977; 1978).

The molar ratio of L-rhamnose to N-acetylglucosamine was found to be consistently 3:1, and the structure of the basic tetrasaccharide repeat unit:

$$\rightarrow3)-\beta-D-GlcNAc-(1\rightarrow2)-\alpha-L-Rha-(1\rightarrow2)-\alpha-L-Rha-(1\rightarrow3)-\alpha-L-Rha-(1\rightarrow$$

I II III

Polysaccharides were subjected to a) methylation analysis which gives information regarding at which positions the neutral sugar residues are linked, b) N-deacetylation followed by methylation analysis reveals the glycosidic linkage of the modified amino sugar (ie positions through which the particular rhamnose sugar is substituted and c) N-deacetylation, deamination, which gives information on the positions to which the amino sugars are substituted (Kenne et al., 1978).

Substitutions of the basic repeat unit were also determined which result in the type and group antigens (Kenne et al., 1978). Confirmation of these structures was achieved using
monoclonal antibodies by Carlin and coworkers (Carlin et al., 1984; Carlin et al., 1986a; Carlin et al., 1986b; Carlin et al., 1987) (Table 1.3).

The structures that were recognised as the type antigens were determined: type I results from the addition of an α-linked glucose to the C4 of the GlcNAc (Carlin and Lindberg, 1986), type II antigen results from the addition of a glucose residue α-linked to the C4 of RhaIII (Carlin and Lindberg, 1983), type IV antigen is the α-linked glucose to C6 of GlcNAc and type V antigen results from an α-linked glucose to RhaII (Carlin and Lindberg, 1987). The group antigens have also been determined: group 3,4 (Carlin and Lindberg, 1987), group 7,8 resulting from the addition of an α-linked glucose to C3 of RhaI (Carlin and Lindberg, 1986) and the group 6 (Type III) antigen is the O-acetylation via C2 of RhaIII (Carlin and Lindberg, 1986).

1.10 Variation in O antigen

Variations in the O polysaccharide component of LPS of many bacterial species has been extensively studied. The O polysaccharide is the most antigenic component of the LPS and, as such, the most variable. Antigenic variation serves to enhance the virulence of the organism and to protect it from host response mechanisms. Many mechanisms exist by which alterations in O antigen can occur, including lysogenisation by temperate bacteriophages, spontaneous mutations in genes and horizontal gene transfer.

1.10.1 Vibrio cholerae and serotype conversion

V. cholerae can be divided into either the classical or the El Tor biotypes (Freeley, 1965; Sen et al., 1979) which in turn can be further subdivided into different serotypes, Inaba
### Table 1.3 Molecular structures of *S. flexneri* antigenic determinants

<table>
<thead>
<tr>
<th><em>S. flexneri</em> factor</th>
<th>Structure of determinant</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>α-glucosyl-(1→4)-β-N-acetylglucosamine</td>
</tr>
<tr>
<td>II</td>
<td>α-glucosyl-(1→4)-α-rhamnose (III)</td>
</tr>
<tr>
<td>III</td>
<td>2-O-acetyl-α-rhamnosyl-(1→3)-β-N-acetylglucosamine</td>
</tr>
<tr>
<td>IV</td>
<td>α-glucosyl-(1→6)-β-N-acetylglucosamine</td>
</tr>
<tr>
<td>V</td>
<td>α-glucosyl-(1→3)-α-rhamnose (II)</td>
</tr>
<tr>
<td>6</td>
<td>same as antigen III</td>
</tr>
<tr>
<td>7,8</td>
<td>α-glucosyl-(1→3)-α-rhamnose (I)</td>
</tr>
</tbody>
</table>

(reproduced from Simmons, 1986)
and Ogawa and the rarer Hikojima. The differentiation of these serotypes is determined by the expression of the antigens A, B and C. Inaba strains express A and C antigens, whereas Ogawa strains express A, B and some C antigens, Hikojima express high levels of all three antigens, however, these strains are unstable (Sakazaki and Tamura, 1971).

Interconversion between serotypes can occur, with Inaba strains gaining the ability to express B antigen and hence becoming Ogawa and Ogawa losing the B antigen and hence becoming Inaba. The latter being the most common conversion to occur (Shrivastava and White, 1947; Bhaskaran and Gorrill, 1957; Sheehy et al., 1966; Nobechi and Nakamo, 1967).

Recently, a single gene, rfbT, was identified that was found to be responsible for the serotype switching observed in V. cholerae (Stroehel et al., 1992). The rfbT gene is contained in both Inaba and Ogawa serotypes, however, differences between these genes results in the loss of expression of the B antigen in the Inaba serotype. Inaba strains arise from mutations in rfbT of Ogawa which results in a non-functional product. The reverse change requires a series of precise mutations and is less common (Stroehel et al., 1992).

1.10.1.1 V. cholerae O139

Traditionally, epidemics due to V. cholerae have been caused by strains of serotype O1, the recent cholera epidemic, however, was caused by a strain of a new serotype, O139 (Shimada et al., 1992). Genetic analyses of strains of this serotype led to the hypothesis that the O139 serotype arose from the transfer of DNA from a non-O1 to an O1 strain (Morris, 1994; Bik et al., 1995). The evidence for this was based upon the identification of a locus named otn possessing 7 open reading frames. Two genes otnA and otnB previously thought to be involved in O antigen biosynthesis were mutated and found to affect capsule synthesis. otnA and otnB were also identified in V. cholerae O69 and O141 serotypes (Bik et al., 1995; 1996)
The genes required for antigenic diversity were encoded by the rfb genes found to be linked to rfaD and located distal to the otn region (Stroher et al., 1997). The rfb region was sequenced and contains genes homologous to rfb genes of different serotypes. The otn region was found to be flanked by sequences homologous to O1 rfb DNA which may have been the sites involved in the recombination event which transferred the novel DNA into the O1 strain resulting in the O139 serotype (Stroher et al., 1997).

1.10.2 Salmonella enterica serovar Typhimurium and serotypic variation

Form variation of O antigens have been well characterised in Salmonella enterica sv. Typhimurium (Zinder, 1957). The structure of the repeat unit is:

\[
\begin{align*}
\text{Abe} \\
\rightarrow 2 \alpha - \text{Man} (1 \rightarrow 4 ) - \beta - \text{Rha} (1 \rightarrow 3) - \beta - \text{Gal} (1 \rightarrow)
\end{align*}
\]

O antigen factor 12₂ arises from the addition of an α-linked glucose residue onto the C4 of the galactose residue of the repeat unit (Tinelli and Staub, 1960; Hellerqvist et al., 1969). Glucosylation of the galactose residue is not an essential step in the biosynthetic pathway of O unit and in fact it is an unstable phenotype, alternating between a 12₂ positive (12₂ +) form to a 12₂ negative (12₂ -) form (Kauffmann, 1941). The gene oafR has been found to be associated with the 12₂ glucosylating genes (Mäkelä and Mäkelä, 1966; Mäkelä, 1973) however, the molecular mechanism of this modification is still unknown (Helander et al., 1992).

The abequose side group of the repeat unit is acetylated, conferring the O5 serotype (Slauch et al., 1995). Recently, a gene oafA was identified which is responsible for this O-acetylation. Mutation in oafA results in the loss of the O-acetyl group and hence a change in serotype, however, virulence is not affected. oafA exhibits homology to the O-acetylase gene, oac, of S. flexneri bacteriophage Sf6 and the nodX gene of Rhizobium legumonsarum.
Linkage analyses determined oac to map at 48.5 min near the rfb region at 45 min (Slauch et al., 1995).

\[
\begin{align*}
2OAcAbe \\
\downarrow \\
\rightarrow 2) -\alpha- Man (1\rightarrow 4) -\beta- Rha (1\rightarrow 3) - \beta- Gal (1 \rightarrow \\
\end{align*}
\]

factor O5

A second type of form variation exists in S. enterica sv. Typhimurium which also involves the \(\alpha\)-linked glucosylation of the galactose residue, however, this linkage is attached to C6 (Iwashita and Kanegasaki, 1973). This modification was found to be the result of lysogenisation by bacteriophage P22 which recognises O antigen (Zinder, 1958; Eriksson and Lindberg, 1977; Kita and Nikaido, 1973; Israel et al., 1972; Iwashita and Kanegasaki, 1976) (Fig. 1.12). Changes in O unit structure to express either O12₂ or O antigen factor 1 result in the inability of P22 to infect the strain (Iwashita and Kanegasaki, 1973). Neither of these antigens is normally fully expressed in the bacteria.

\[
\begin{align*}
\text{Abe} \\
\downarrow \\
\rightarrow 2) -\alpha- Man (1\rightarrow 4) -\beta- Rha (1\rightarrow 3) - \beta- Gal (1 \rightarrow \\
\end{align*}
\]

factor 12₂

\[
\begin{align*}
\text{Abe} \\
\downarrow \\
\rightarrow 2) -\alpha- Man (1\rightarrow 4) -\beta- Rha (1\rightarrow 3) - \beta- Gal (1 \rightarrow \\
\end{align*}
\]

factor 1

1.11 Bacteriophages which have their receptor in the O-antigen

To date, serotype converting bacteriophages which recognise the O antigen as a receptor all belong to the group C of the Bradley classification (Fig. 1.12) (Lindberg, 1977). Members of this group possess a hexagonal shaped head and a small tail with tail spikes attached (Bradley, 1967). The best characterised members of this group include bacteriophage P22 which lysogenises S. enterica sv. Typhimurium (Vieu et al., 1965; Israel et al., 1967;
Figure 1.12  Basic morphological types of bacteriophages.

Diagram is reproduced from Bradley (1967).
King et al., 1973) and the S. anatum bacteriophages ε^{15}, ε^{34} and g_{341} (Israel et al., 1972; Iwashita and Kanegasaki, 1973; 1975; 1976; Kanegasaki and Wright, 1973; Lindberg, 1977). The S. flexneri serotype converting bacteriophage Sf6 also belongs to this group (Lindberg et al., 1978).

1.11.1 Lysogenisation of group E Salmonellae

Bacteriophages which recognise the O-antigen as a receptor were originally identified in group E Salmonellae such as S. anatum (Bruner and Edwards, 1948). Bacteriophages ε^{15}, ε^{34} and g_{341} are able to mediate serotype conversion and all belong to group C of Bradley classification (Fig. 1.13) (Uetake et al., 1955, 1958; Vieu et al., 1965). These phages recognise their receptors in the O-antigen via their tail spikes (Lindberg, 1977).

The repeat unit of S. anatum is the tetrasaccharide:

\[
\begin{align*}
\text{OAc} \\
\rightarrow^6 \alpha- \text{Man} (1\rightarrow4) \beta- \text{Rha} (1\rightarrow3) \alpha- \text{Gal} (1\rightarrow)
\end{align*}
\]

composed of mannose, rhamnose and galactose and an O-acetyl group linked to the galactose residue which determines the O antigenic factor 3,10. Bacteriophage ε^{15} recognises S-LPS (Kanegasaki and Wright, 1973) and hydrolyses the linkage between the rhamnose and galactose sugars. Conversion of antigen 3,10 to antigen 3,15 occurs in a three step pathway which was determined by mutant analyses (Robbins and Uchida, 1962, 1965).

Bacteriophage ε^{15} encodes an enzyme which blocks the transacetylase synthesis so that O-acetyl groups are not transferred to the galactose residue of the O repeating unit (Robbins et al., 1965). A bacteriophage encoded polymerase is produced which alters the α linkage between the repeat units to β linkages. Concurrently, an inhibitor of the host polymerase is also produced (Losick and Robbins, 1967; Bray and Robbins, 1967). The
Figure 1.13 Structures of *Salmonella anatum* (subgroup E) O antigen and modifications mediated by serotype converting bacteriophages, $\epsilon^{15}$, $\epsilon^{34}$ and $g_{341}$.

Diagram is reproduced from Lindberg (1977)
<table>
<thead>
<tr>
<th>Bacterial strain</th>
<th>O-antigen</th>
<th>Structure of O-polysaccharide and site for phage enzyme activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. anatum</td>
<td>3, 10</td>
<td>d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\alpha\rightarrow6) d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\alpha\rightarrow6)</td>
</tr>
<tr>
<td>S. anatum ((g^{21}))</td>
<td>3, 10</td>
<td>d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\alpha\rightarrow6) d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp</td>
</tr>
<tr>
<td>S. anatum ((e^{19a}))</td>
<td>3, 15</td>
<td>d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\beta\rightarrow6) d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp</td>
</tr>
<tr>
<td>S. anatum ((e^{19}))</td>
<td>3, (10), 15</td>
<td>d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\beta\rightarrow6) d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp</td>
</tr>
<tr>
<td>S. anatum ((e^{15}, e^{34}))</td>
<td>3, (10), (15), 34</td>
<td>d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp (1^\beta\rightarrow6) d-Manp (1^\beta\rightarrow4) L-Rhap (1^\alpha\rightarrow3) d-Galp</td>
</tr>
</tbody>
</table>
S. anatum[ε\textsuperscript{15}] lysogens are no longer susceptible to infection by the same phage, due to the new O-antigen structure; however, this is now recognised by bacteriophage ε\textsuperscript{34} (Uetake et al., 1958).

Bacteriophage ε\textsuperscript{34} recognises and hydrolyses the galactose -mannose linkage in the repeat unit. Lysogenisation with ε\textsuperscript{34} results in the addition of a glucose group to the C-4 of the galactose residue which renders the strain resistant to superinfection due to a change in the receptor (Lindberg, 1977).

Lysogenisation of S. anatum, of serotype 3,10, by phage g\textsubscript{341} results in the lack or greatly reduced number of O-acetyl groups on the galactose residue in the repeating unit. It appears that infection by phage g\textsubscript{341} blocks transacetylase synthesis (Lindberg, 1977).

The three phages described have enzymatic activity, localised to their tail fibres, which enables them to cleave the O units at specific locations, thus providing access to the bacterial cell surface where they can then inject their genetic material. ε\textsuperscript{15} possesses endoglycosidase activity as does ε\textsuperscript{34}, however, g\textsubscript{341} appears to have an esterase which deacetylates the O polysaccharide chain (Takeda and Uetake, 1973) resulting in the greatly reduced acetylation of the LPS.

1.11.2 Serotype converting bacteriophages of S. flexneri

The discovery of serotype converting bacteriophages in Salmonella coupled with the observation of antigenic changes in S. flexneri, led to the search for a similar mode of conversion in the latter species.
1.11.2.1 Identification of *S. flexneri* bacteriophages

Attempts to isolate temperate bacteriophage began with cultures of each serotype being subjected to a freeze-thawing process (Matsui, 1958). The filtrate of serotype 4c was found to partially lyse cultures of serotypes 1a and 3b. The surviving organisms were plated onto solid medium and then assessed for changes to their antigenic status. Agglutination of lysogens of serotypes 1a and 3b with type I or type III and type IV antisera showed that the majority of the colonies reacted positively with anti-type IV sera. This indicated a change in the type antigen expressed by these strains. Similarly, absorption studies of antisera raised against the original parent strains (1a and 3b) using cultures of the lysogens showed that the original type antigen (type I or type III) remained (Matsui, 1958).

Confirmation of these results was obtained when the lysate of the strain of serotype 4c was used with strains of serotype 1a, 1b, 2a, 3b and Y. Each of the strains were found to be converted by the lysate to express they type IV antigen (Iseki and Hamano, 1959). The conclusion was that the serotype conversion observed was caused by a temperate bacteriophage as in the case in group E *Salmonellae*.

1.11.2.2 Identification of bacteriophages fII, f7,8 and fV

It was not until ten years later that two other bacteriophages were identified. These phages were isolated from a strain of *S. flexneri* of serotype 2b, NTCC4, which has two modifications of its basic repeat unit (Fig. 1.9). After 18 hours of growth, a culture of strain NTCC4 was heated for a period of 30 minutes and the lysate was found to produce two plaque morphologies, large and small, when plated with an indicator strain of serotype Y. Each plaque type was purified and lysogens isolated were assessed for reactivity to specific antisera.
Lysogens isolated from the larger plaque type were found to be converted from serotype Y to express the group 7,8 antigen indicative of strains of serotype X, whereas the lysogens from the smaller plaque type were found to convert strains to express the type II antigen (Giammanco, 1968). These phages were then named f7,8 and fIII respectively which indicates the group and type antigen they confer when they lysogenise strains of *Shigella flexneri* serotype X and Y.

Serotype conversion of a strain of serotype Y to 2b was achievable using phages f7,8 and fIII irrespective of the order in which the lysogenisation occurred. Bacteriophage f7,8 was able to infect with equal efficiency strains of serotype Y and 2a, and phage fIII was also able to infect strains of serotype Y as easily as serotype X.

Bacteriophage fIV was isolated in the same manner as f7,8 and fIII and was found to be capable of lysogenising strains of serotypes 1a, 2a, 2b, X and Y. Upon subsequent assessment of type antigens expressed by the lysogens, all were found to stably express type antigen 5. However, group antigens appeared to be expressed weaker than in wildtype strains of serotype 5 (Giammanco and Natoli, 1968). The reason for this was unknown.

### 1.11.2.3 Bacteriophage Sf6

To date, the best characterised *Shigella flexneri* bacteriophage is Sf6 (Gemski *et al*., 1975). Lysogenisation by Sf6 results in the expression of the group 6 antigen (Lindberg *et al*., 1972) due to the O-acetylation of the rhamnose III sugar of the repeat unit of strains of serotype X and Y (Clark *et al*., 1991; Verma *et al*., 1991). Sf6 also possesses endorhamnosidase activity and belongs to group C of the Bradley morphological classification (Lindberg *et al*., 1978). The genome of Sf6 has been mapped and has been found to be approximately 40 kb in size (Clark *et al*., 1991). The genome of Sf6 exhibits very good homology to that of *Salmonella* phage P22 with many genes found in identical locations (Chua, 1996).
1.11.3 Molecular basis for serotype conversion

Two genes have been identified which have been isolated from serotype converting bacteriophages of *S. flexneri*. The gene responsible for the O-acetylation, *oac*, has been cloned and sequenced and found upstream of the integrase gene. Transformation of a plasmid copy of the *oac* gene into strains of serotype X and Y is sufficient to convert these strains to express the group 6 antigen (Lindberg *et al.*, 1978; Clark *et al.*, 1991; Verma *et al.*, 1991), implying it does not require any accessory factors to mediate serotype conversion.

The second gene identified from a *S. flexneri* serotype converting phage was isolated from SfX, previously f7,8 (Giammanco, 1968; Verma *et al.*, 1993). Bacteriophage SfX mediates the glucosylation of the rhamnose I sugar of the repeat unit of strains of serotype 2a, 3b and Y, resulting in the expression of the group 7,8 antigen. The gene, *gtrX*, isolated from SfX, like *oac* has been identified as a single gene which acts alone (Verma *et al.*, 1993).

1.12 Aims of this study:

Although various serotype converting bacteriophages of *S. flexneri* have been isolated previously, few of these have been characterised to any extent. Primarily, the mechanism by which these phages mediate their serotype conversion is of utmost interest. In this study, the aim is to isolate the bacteriophage that encodes the type II antigen, SfII, and identify the gene(s) responsible. The mechanism of serotype conversion will be determined and characterisation of the proteins involved. Finally, attempts will be made to elucidate at which step in the complex *S. flexneri* biosynthetic pathway of LPS modification occurs.
Chapter Two

Materials and Methods

2.1 Bacterial strains, bacteriophages and plasmids

The *Shigella flexneri* strains used are listed in Table 2.1. Table 2.2 describes the *Escherichia coli* K-12 strains used in this study. The plasmids and phage cloning vectors which were used in this study are listed in Table 2.3. Bacteriophage Sf6 was provided by A. Lindberg, Stockholm, Sweden.

2.2 Maintenance of bacterial and bacteriophage strains

For long term storage, all strains were maintained as lyophilized cultures, stored in *vacuo* in sealed glass ampoules. When required, an ampoule was opened and its contents suspended in several drops of the appropriate sterile broth. Half the contents were then transferred to a 10 ml bottle of nutrient broth (NB) and incubated with aeration for 18 h at the appropriate temperature. The other half was streaked onto two nutrient agar plates and incubated for 18 h at the appropriate growth temperature. Antibiotics were added to the media when appropriate. If the colony form was uniform, single colonies were selected and picked off plates for subsequent storage or use.

Short-term storage of strains in routine use was as a suspension of freshly grown bacteria in glycerol (32% (v/v)) and peptone (0.6% (w/v)) at -70°C. Fresh cultures from glycerols were prepared by streaking a loopful of the glycerol suspension onto a nutrient agar plate (with or without antibiotic as appropriate) followed by incubation for 18 h just prior to use.
### Table 2.1 *Shigella flexneri* strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Serotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE569</td>
<td>1a</td>
<td>L. Beutin&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>PE568</td>
<td>1b</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE523</td>
<td>2a</td>
<td>C. Murray&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>PE655</td>
<td>2a</td>
<td>WRAIR&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>PE567</td>
<td>2a</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE574</td>
<td>2a</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE642</td>
<td>2a</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE790</td>
<td>2a</td>
<td>Alice Springs Hospital&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>PE824 (NCTC4)</td>
<td>2b</td>
<td>NTCC&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>PE571</td>
<td>3a</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE844</td>
<td>3b</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE566</td>
<td>4a</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE572</td>
<td>4b</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE780</td>
<td>5a</td>
<td>WRAIR</td>
</tr>
<tr>
<td>PE565</td>
<td>5b</td>
<td>L. Beutin</td>
</tr>
<tr>
<td>PE576</td>
<td>X</td>
<td>C. Murray</td>
</tr>
<tr>
<td>PE577</td>
<td>Y</td>
<td>C. Murray</td>
</tr>
<tr>
<td>PE870</td>
<td>Y</td>
<td>WRAIR</td>
</tr>
</tbody>
</table>

<sup>a</sup> L. Beutin, *Escherichia coli* Reference Laboratory, Robert Koch Institute, Berlin

<sup>b</sup> C. Murray, *Salmonella* Reference Laboratory, Institute of Medical and Veterinary Sciences, Adelaide, South Australia

<sup>c</sup> WRAIR, Walter Reed Army Institute for Research, Washington, DC., Virginia, USA

<sup>d</sup> Alice Springs Hospital Isolate, Alice Springs, Northern Territory, Australia

<sup>e</sup> NCTC, National collection of Type Cultures, Central Public Health Laboratory, London U.K.
Table 2.2 *Escherichia coli* strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source/ Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5</td>
<td>F· recA1 endA1 hsdR17 (rK, mK&lt;sup&gt;+&lt;/sup&gt;) supE44λ- thiI gyrA relA1</td>
<td>B.R.L. &lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>DH5α</td>
<td>F· supE44 ΔlacU169 (ϕ80lacZ ΔM15) hsdR17 recA1 endA1 deoR gyrA96 thi-1 relA1</td>
<td>B.R.L.</td>
</tr>
<tr>
<td>S17-1</td>
<td>pro hsdR RP4-2-Tc::Mu Km::Tn7</td>
<td>U. Priefer &lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>SM10λpir</td>
<td>A lysogen of SM10 (thi thr leu tonA lacY supE supE recA::RP4-2-Tc::Mu) used for conjugal transfer of plasmids with R6k replicon.</td>
<td>Kaniga et al., (1991)</td>
</tr>
<tr>
<td>E1196</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;, Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Taylor et al., (1989)</td>
</tr>
<tr>
<td>E2096</td>
<td>DH5α· pGP1-2</td>
<td>Laboratory Strain</td>
</tr>
<tr>
<td>C75A</td>
<td>tonA22 pho-64 ompF627 (T2&lt;sup&gt;R&lt;/sup&gt;) relA1</td>
<td>CGSC&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>CC118</td>
<td>Δ(ara, leu)7697 ΔlacX74 ΔphoA20 galE galK thi rpsE rpoB argE(am) recA1</td>
<td>Manoil, (1991)</td>
</tr>
<tr>
<td>S874</td>
<td>F· lacZ2286 trp49 upp12 relA1 D(sbcB-rfb)86 rpsL150 λ-</td>
<td>CGSC</td>
</tr>
</tbody>
</table>

<sup>a</sup> Bethesda Research Laboratories, Gaithersburg, Maryland USA
<sup>b</sup> Max Planck-Institut für Biologie, Tübingen, FRG
<sup>c</sup> Coli Genetic Stock Centre, Yale University, New Haven, Conn.
<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC18/19</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Vieira and Messing (1982)</td>
</tr>
<tr>
<td>pBluescript S/K</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pBC-KS</td>
<td>Cml&lt;sup&gt;R&lt;/sup&gt; derivative of pBluescript</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pK184/194</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Hohn and Collins, (1980)</td>
</tr>
<tr>
<td>pGP1-2</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Tabor and Richardson (1985)</td>
</tr>
<tr>
<td>pGEM-T</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, PCR product cloning vector 3kb</td>
<td>Promega</td>
</tr>
<tr>
<td>pJRD215</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;, 10.2 kb cosmid vector</td>
<td>Davison &lt;i&gt;et al.&lt;/i&gt;, (1987)</td>
</tr>
<tr>
<td>pGEM5zf+</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt;, 3 kb</td>
<td>Promega</td>
</tr>
<tr>
<td>pPM2101</td>
<td>Ap&lt;sup&gt;R&lt;/sup&gt; Tc&lt;sup&gt;R&lt;/sup&gt;</td>
<td>Sharma &lt;i&gt;et al.&lt;/i&gt;, (1989)</td>
</tr>
<tr>
<td>pRMCD28</td>
<td>&lt;i&gt;E. coli&lt;/i&gt; phoA&lt;sup&gt;*&lt;/sup&gt; in pWSK29</td>
<td>Daniels &lt;i&gt;et al.&lt;/i&gt;, (1998). (In press)</td>
</tr>
<tr>
<td>pRMCD70</td>
<td>&lt;i&gt;E. coli&lt;/i&gt; lac&lt;sup&gt;Z&lt;/sup&gt; in pWSK29</td>
<td>Daniels &lt;i&gt;et al.&lt;/i&gt;, (1998) (In press)</td>
</tr>
<tr>
<td>pUT/Km</td>
<td>mini-Tn5 Km&lt;sup&gt;R&lt;/sup&gt; cartridge in pUT</td>
<td>De Lorenzo &lt;i&gt;et al.&lt;/i&gt;, (1990)</td>
</tr>
</tbody>
</table>
Bacterial strains were prepared for long-term storage by suspension of several loopfuls in a small volume of sterile skimmed milk. Approximately 0.2 ml aliquots of this thick bacterial suspension were dispensed into sterile 0.25 x 4 inch freeze drying ampoules and the end of each ampoule was plugged with cotton wool. The samples were then lyophilized in a freeze drier. After the vacuum was released, the cotton wool plugs were pushed well down the ampoule and a constriction was made just above the level of the plug. The ampoules were evacuated to a partial pressure of 30 microns and then sealed at the constriction without releasing the vacuum. Finally the ampoules were labelled and stored at 4°C.

Bacteriophage stocks were passed through 0.45 µm Millipore membrane filters and stored at 4°C over 0.5% (v/v) chloroform.

2.3 Growth Media

The following nutrient media were used for bacterial cultivation. Nutrient broth (NB) (Oxoid), consisted of 3 g Bacto-beef extract, 5 g Bactopeptone with added sodium chloride (NaCl) (5 g) per litre distilled water. Nutrient agar (NA) was NB with the addition of 1.5% Oxoid Bacto-Agar. Luria broth (LB), composed of Bacto-tryptone (10 g/l) (Difco), Bacto-yeast (5 g/l) (Difco) and NaCl (5 g/l). Soft agar contained equal volumes of LB and LA. Minimal medium was also prepared as described by Miller (1972) and supplemented prior to use with MgSO₄ (0.2 mg/ml), glucose (2 mg/ml) and thiamine-HCl (5 mg/ml). Methionine assay medium (Difco) was reconstituted according to manufacturers instructions.

Antibiotics were added to broth and solid media at the following final concentrations: ampicillin (Ap), 50 µg/ml; chloramphenicol (Cm), 25 µg/ml; kanamycin (Km), 50 µg/ml; rifampicin (Rif), 200 µg/ml.

Incubations were at 37°C unless otherwise specified. Liquid cultures, were normally grown in 20 ml McCartney bottles. Optical densities (OD) were measured at 600nm.
2.4 Chemicals and reagents

Chemicals were analytical grade. Ethanol, methanol, propan-2-ol, iso-amyl-alcohol, hydrochloric acid, glycerol, phenol, polyethylene glycol-8000 (PEG), sodium dodecyl sulphate (SDS), sodium chloride, ammonium acetate and sucrose were from BDH Chemicals. Adenosine-5’-triphosphate sodium salt (ATP), dithiothreitol (DTT), herring sperm DNA, ethidium bromide, and Tris (Trizma base) were from Sigma (St. Louis, MO). Caesium chloride, ethylene-diamine-tetra-acetic-acid, disodium salt (EDTA), citric acid, calcium chloride, magnesium chloride and sodium hydroxide were obtained from Ajax Chemicals, NSW, Australia. Sarkosyl was obtained from Geigy and Tween 20 from Boehringer Mannheim.

Antibiotics were purchased from Sigma (ampicillin, kanamycin sulphate, rifampicin), and Calbiochem (tetracycline, chloramphenicol). All other anti-microbial agents (dyes, detergents and antibiotics) were purchased from Sigma Chemical Co., BDH Chemicals Ltd., Glaxo, or Calbiochem.

Electrophoresis grade reagents were obtained from Boehringer-Mannheim (acrylamide and ammonium persulphate) and BRL (ultra pure N,N'-methylen bis-acrylamide and urea).

The four deoxyribonucleotide triphosphates (dATP, dCTP, dGTP and dTTP) and their corresponding dideoxy-ribonucleotide triphosphate homologues (ddATP, ddCTP, ddGTP and ddTTP), X-gal (5-Bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside), IPTG (isopropyl-β-D-thiogalacto-pyranoside), p-nitrophenyl phosphate and 2-nitrophenyl-β-D-galactopyranoside (ONPG) were purchased from Boehringer-Mannheim.

Sequencing kits using either dye-labelled primer or dye-labelled terminators were purchased from Perkin Elmer Applied Biosystems (Foster City, California). [35S]-Methionine (1.270 Ci/mmole) was purchased from Amersham. Digoxigenin (DIG) DNA labelling and detection kits were purchased from Boehringer-Mannheim.
2.5 Enzymes and Immunoconjugates

Deoxyribonuclease I (DNase I) was obtained from Bresatec. Lysozyme from Sigma and pronase and proteinase K from Boehringer-Mannheim. All restriction endonucleases were purchased from either Boehringer-Mannheim, New England Biolabs or Progen and used according to the suppliers' instructions.

Other DNA modifying enzymes were purchased from the following suppliers: Progen (T4 DNA ligase, calf intestinal phosphatase) and Boehringer-Mannheim (DNA polymerase I, Klenow fragment of DNA polymerase I, and molecular biology grade alkaline phosphatase). Taq polymerase (Ampli Taq) was purchased from Perkin Elmer Cetus Corp.

Horseradish peroxidase-conjugated goat anti-rabbit IgG was obtained from Kirkegaard and Perry Laboratories Inc. Anti-digoxigenin-AP (Fab fragments), anti-PhoA and anti-LacZ were from Boehringer Mannheim.

2.6 Transformation procedure

Transformation was performed essentially according to the method described by Brown et al. (1979). E. coli K-12 strains were made competent for transformation with plasmid DNA as follows: an overnight culture (in NB) was diluted 1:20 into NB and incubated with aeration until the culture reached an A600 OD of 0.6 (4x 10^8 cells/ml). The cells were chilled on ice for 20 min, pelleted at 4°C in a bench centrifuge, resuspended in half volume of cold 100 mM MgCl₂, centrifuged again and resuspended in a tenth volume of cold 100 mM CaCl₂. The cells were allowed to stand for 60 min on ice before addition of DNA. Competent cells (0.2 ml) were then mixed with DNA (volume made to 100 μl with 1 x TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and left on ice for a further 30 min. The cell/DNA mixture was heated at 42°C for 2 min and then 3 ml NB was added followed by incubation with shaking at 37°C for 1-2 h. The culture was then plated onto selection plates.
directly or concentrated by centrifugation and plated. Cells with sterile buffer were included as a control.

2.6.1 Super competent cells

A 10 ml overnight culture in LB was diluted 1:20 into LB and incubated with aeration until an $A_{600}$ OD of $4 \times 10^8$ cells/ml was reached. The cells were chilled on ice for 5 min, pelleted at 4°C in a bench centrifuge and resuspended in 10 ml of solution a (30 mM KAc, 100 mM KCl, 10 mM CaCl$_2$, 50 mM MnCl$_2$, 15% glycerol). The solution was pelleted for 15 min at 4°C in a bench centrifuge and resuspended in 1 ml of solution b (10 mM MOPS, 75 mM CaCl$_2$, 10 mM KCl, 15% glycerol) and left on ice for 1-2 h. The final solution was aliquoted (100 µl) into 0.5 ml reaction tubes and stored at -70°C or used immediately.

2.7 Electroporation of S. flexneri and E. coli.

Electrocompetent DH5α and S. flexneri cells were freshly prepared according to the Bio-Rad protocol. An overnight broth of DH5α (500 µl) was used to inoculate 10 ml Luria broth and incubated at 37°C with aeration until the cells reached an $OD_{600}$ 0.5-0.8. The cells were centrifuged for 10 min at 5000 rpm and the supernatant discarded. The cells were resuspended in 5 ml of ice-cold sterile 10% (v/v) glycerol and then centrifuged again for 5 min at 5000 rpm. Finally, the cell pellet was resuspended in 1 ml of ice-cold 10% (v/v) glycerol and kept on ice. The cells were either stored at -70°C or used immediately.

In a sterile microfuge tube on ice, ca. 1-2 µl PCR product or plasmid DNA (in TE or sterile Milli-Q water) were mixed with 100 µl of electrocompetent E. coli or S. flexneri and then transferred to an ice-cold sterile E. coli Pulser TM cuvette (0.2 cm electrode gap, Bio-Rad). The gene pulser (Bio-Rad) was set at 25 µF and the pulse controller at 200 Ω. The E. coli and S. flexneri cells were pulsed at 2.5 kV with time constants of 4.6 - 4.8 msec.
Immediately after electroporation, 1 ml of LB was added to the cuvette, the contents mixed and transferred to a sterile microfuge tube, and then incubated at 37°C for 60-90 min. After centrifugation at 15,000 rpm for 1 min, the supernatant was discarded and the cells gently resuspended in 200 µl of sterile LB and plated onto NA containing appropriate antibiotic.

2.8 Bacterial conjugation

Overnight broth cultures grown in NB or LB were diluted 1:20 and grown to early exponential phase. Donor and recipient bacteria were mixed at a ratio of 1:10 and the cells pelleted by centrifugation (5000 rpm, 5 min, IEC bench centrifuge). The pellet was gently resuspended in 200 µl of broth and spread onto a cellulose acetate membrane filter (0.45 µm, type HA, Millipore Corp.) on a NA plate. This plate was incubated for 4 h at 37°C, when the bacteria on the filter were collected by vortexing the filter in 10 ml NB. After removing the filter, the cell suspension was centrifuged as described above, the cell pellet resuspended in 200 µl of sterile saline, and samples plated onto selective agar.

2.9 Plasmid DNA extraction procedures

Plasmid DNA was isolated by one of the three following procedures:

**Method 1:** Small scale plasmid purification was performed by the three step alkali lysis method using a modification of Garger et al., (1983). Overnight bacterial cultures (1.5 ml) were transferred to a microfuge tube, harvested by centrifugation (45 sec, Eppendorf), and resuspended in 0.1 ml of solution 1 (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). The addition of 0.2 ml of solution 2 (0.2 M NaOH, 1% (w/v) SDS), followed by a 5 min incubation on ice, resulted in cell lysis. After the addition of 0.15 ml of solution 3 (60 ml of 5 M potassium acetate, pH 4.8, 11.5 ml of glacial acetic acid and 28.5 ml of H2O) and a 5 min incubation on ice, protein, chromosomal
DNA and high molecular weight RNA were collected by centrifugation (90 sec, Eppendorf). The supernatant was transferred to a fresh tube and extracted once with TE-equilibrated phenol and once with diethyl ether. Plasmid DNA was precipitated by the addition of 2 volumes of 100% ethanol and a 2 min incubation at room temperature. The DNA was collected by centrifugation (15 min, Eppendorf), washed with 70% (v/v) ethanol and dried in vacuo. The pellet was resuspended in 50 μl of 1x TE.

**Method 2:** Small scale plasmid preparations were also prepared essentially as above however, following the removal of protein, chromosomal DNA and high molecular weight RNA, the supernatant was transferred to a fresh tube to which was added 0.25 ml 7.5 M Ammonium acetate pH 7.8. This was kept on ice for a further 15 min and following centrifugation (15 min, Heraeus Biofuge), the supernatant was again transferred to a fresh tube. Plasmid DNA was precipitated by the addition of 0.8 ml propan-2-ol with a 15 min incubation on ice. The DNA was collected by centrifugation (15 min, Heraeus Biofuge), washed with 70% (v/v) ethanol and dried in vacuo. The pellet was resuspended in 50 μl 1x TE or MQ.

**Method 3:** Large scale plasmid purification was performed by the three step alkali lysis method (Garger *et al*., 1983). Cells from a one litre culture were harvested (6,000 rpm, 15 min, 4°C, GS-3, Sorvall) and resuspended in 24 ml of solution 1 (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA). Freshly prepared lysozyme (4 ml of 20 mg/ml in solution 1) was mixed with the cell suspension and incubated at room temperature for 10 min. Addition of 55 ml of solution 2 (0.2 M NaOH, 1% (w/v) SDS), followed by a 5 min incubation on ice resulted in total lysis of the cells. After the addition of 28 ml solution 3 (60 ml 5 M potassium acetate, pH 4.8, to which was added 11.5 ml glacial acetic
acid and 28.5 ml of H₂O) and incubation on ice for 15 min, protein, chromosomal DNA and high molecular weight RNA were removed by centrifugation (8,000 rpm, 20 min, 4°C, GSA, Sorvall). The supernatant was then extracted with an equal volume of a TE saturated phenol, chloroform, isoamyl alcohol mixture (25:24:1). Plasmid DNA from the aqueous phase was precipitated with 0.6 weight of 100% (v/v) propan-2-ol at room temperature for 10 min and collected by centrifugation (10,000 rpm at 4°C, 35 min, GSA, Sorvall). After washing in 70% (v/v) ethanol, the pellet was dried in vacuo and resuspended in 4.8 ml 1 x TE. Plasmid DNA was purified from contaminating protein, chromosomal DNA and RNA by centrifugation on a two step CsCl ethidium bromide gradient according to Garger et al., (1983). The DNA band was removed by side puncture of the tube with a 19 gauge needle attached to a 1 ml syringe. The ethidium bromide was extracted using isoamyl alcohol. CsCl was then removed by dialysis overnight against three changes of 5 litres 1x TE at 4°C. DNA was stored at 4°C.

**Method 4:**

Triton X-100 cleared lysates were prepared from 10 ml overnight cultures by a modification of the procedure of Kahn et al. (1979). Cells were resuspended in 0.4 ml 25% (w/v) sucrose in 50 mM Tris-HCl, pH 8.0. Lysozyme (50 μl, 10 mg/ml freshly prepared in H₂O) and 50 μl of 0.25 M EDTA, pH 8.0 were added to cells in Eppendorf tubes and left to stand on ice for 15 min. 0.5 ml TET buffer (50 mM Tris-HCl, 66 mM EDTA, pH 8.0, 0.4% Triton X-100) was added followed by a brief mixing by inversion of the tubes. The chromosomal DNA was then pelleted by centrifugation (20 min, 4°C, Eppendorf). The supernatant was extracted twice with TE saturated phenol (pH 7.5) and twice with diethyl-ether. Plasmid DNA was precipitated by the addition of 0.6 volumes of propan-2-ol and allowed to stand at -70°C for 30 min. The
precipitate was collected (10 min, Eppendorf), washed once with 1 ml 70% (v/v) ethanol, dried in vacuo and resuspended in 50 μl 1 X TE buffer.

2.10 Preparation of *S. flexneri* or *E. coli* genomic DNA

Whole genomic DNA from either *S. flexneri* or *E. coli* was prepared according to Manning *et al.* (1986). Cells from a 20 ml overnight culture were pelleted in a bench centrifuge for 10 min and washed once with TES buffer (50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 50 mM NaCl). The pellet was then resuspended in 2 ml of 25% (w/v) sucrose, 50 mM Tris-HCl, pH 8.0 and 1 ml of lysozyme (10 mg/ml in 0.25 mM EDTA, pH 8.0) was added and the mixture incubated on ice for 20 min. TE buffer (0.75 ml) and 0.25 ml of lysis solution (5% (w/v) sarkosyl, 50 mM Tris-HCl, pH 8.0, 0.25 mM EDTA, pH 8.0) were added, together with 2 mg solid pronase. The mixture was gently mixed and incubated at 56°C for 60 min. This was followed by three extractions with TE-saturated phenol and two extractions with diethyl-ether. The genomic DNA was precipitated with four volumes of 100% ethanol and resuspended in 1 ml of 1 x TE.

2.10.1 Quick method for genomic DNA extraction

Cells from a 10 ml overnight culture were pelleted in a bench centrifuge for 10 min and resuspended in 3 ml saline. An equal volume of TE-saturated phenol was added and the mixture vortexed intermittently for 2 min. Following centrifugation at 5 K for 5 min, the aqueous phase was transferred to a clean 20 ml McCartney bottle and 3 ml cold 100% ethanol added. The genomic DNA was spooled using a pasteur pipette and washed in 500 μl of 70% ethanol and resuspended in 1 ml of TE or Milli Q water.
2.11 Analysis and manipulation of DNA

2.11.1 DNA quantitation

The concentration of DNA in solutions was determined by measurement of absorption at 260 nm and assuming an A_{260} of 1.0 is equal to 50 μg DNA/ml (Miller, 1972).

2.11.2 Restriction endonuclease digestion of DNA

Most cleavage reactions were done using the restriction enzyme buffer SPK (10x: 200 mM Tris-HCl pH 7.5, 50 mM MgCl₂, 5 mM dithiothreitol, 1 mM EDTA, 500 mM KCl and 50% glycerol). The remaining restriction digests were carried out using EB buffer (10 mM Tris-HCl, pH7.5, 6 mM MgCl₂, 1 mM DTT) as a basis, with either the addition of NaCl or KCl as described by the manufacturers. 0.1-0.5 μg of DNA or purified restriction fragments were incubated with 2 units of each restriction enzyme in a final volume of 20 μl, at 37°C, for 1-2 hr. The reactions were terminated by heating at 65°C for 10 min. Prior to loading onto a gel, a one tenth volume of tracking dye (15% (w/v) Ficoll, 0.1% (w/v) bromophenol blue, 0.1 μg/ml RNase A) was added.

2.11.3 Analytical and preparative separation of restriction fragments

Electrophoresis of digested DNA was carried out at room temperature on horizontal, 0.6%, 0.8% or 1% (w/v) agarose gels (Seakem HGT). Gels were run at 100 V for 4-5 hr in either 1x TBE buffer (67 mM Tris base, 22 mM boric acid and 2 mM EDTA, final pH 8.8), or 1 x TAE buffer (40 mM Tris acetate and 2 mM EDTA). After electrophoresis the gels were stained in distilled water containing 2 μg/ml ethidium bromide. DNA bands were visualised by trans-illumination with UV light and photographed using either Polaroid 667 positive film or Thermal paper (K65HM) for Mitsubishi Video Copy Processor.
For preparative gels Sea Plaque (Seakem) low-gelling-temperature agarose at a concentration of 0.6% (w/v) was used for separation of restriction fragments, which were recovered by the following methods:

**Method 1:** DNA bands were excised and the agarose melted at 65°C. Five volumes of 20 mM Tris-HCl, 1 mM EDTA, pH 8.0 buffer were added and the agarose extracted with phenol:water (1:1) and then phenol:chloroform (1:1). Residual phenol was removed with chloroform and the DNA precipitated with two volumes of ethanol and one tenth volume of 3 M sodium acetate, pH 5.0. DNA was collected by centrifugation (15 min, Eppendorf), washed once with 70% (v/v) ethanol and dried in vacuo before being resuspended in 1 x TE buffer.

**Method 2:** After electrophoresis the required DNA bands were excised and then placed inside dialysis tubing. This was then positioned in an electrophoretic tank filled with 0.5 x TAE buffer. A current was applied causing the DNA to move out of the gel and into the buffer contained in the dialysis tubing. The DNA was then extracted with an equal volume of TE saturated phenol and precipitated with two volumes of ethanol and one tenth volume of 3 M sodium acetate, pH 5.0.

**Method 3:** The Qiagen gel extraction kit was used according to the instructions provided by the manufacturer.

### 2.11.4 Calculation of restriction fragment size

The sizes of restriction enzyme fragments were calculated by comparing their relative mobility with that of EcoRI digested *Bacillus subtilis* bacteriophage SPP1 DNA. The calculated sizes of the SPP1 EcoRI standard fragments used differ from those published (Ratcliff *et al.*, 1979) using bacteriophage lambda and plasmid pBR322 as standards. The
sizes (kilobases, kb) used were: 8.5; 7.35; 6.1; 4.84; 3.59; 2.81; 1.95; 1.86; 1.51; 1.39; 1.16; 0.98; 0.72; 0.48; 0.36; 0.09.

2.12 DNA cloning procedures

2.12.1 Dephosphorylation of DNA using alkaline phosphatase

Restriction enzyme digested DNA was treated with alkaline phosphatase by the following method. 0.1-0.5 μg of digested plasmid DNA was incubated with 1 unit of alkaline phosphatase (Calf intestinal: CIP), for 30 min at 37°C. The reaction was terminated by the addition of EDTA, pH 8.0 to a final concentration of 3 mM followed by heating at 65°C for 10 min. The reaction mix was then extracted twice with TE saturated phenol and twice with diethyl ether. DNA was precipitated overnight at -20°C with two volumes of ethanol and 1/10 volume of 3 M sodium acetate pH 8.0. The precipitate was collected by centrifugation (15 min, Eppendorf), washed once with 1 ml 70% (v/v) ethanol, dried in vacuo and dissolved in 1x TE buffer.

2.12.2 Ligation of DNA

Ligation reactions with T4 DNA ligase were performed in 1x ligase buffer (20 mM Tris-HCl, 10 mM MgCl₂, 0.6 mM ATP, 10 mM DTT and BSA (100 mg/ml)) for cohesive ends, or 1 mM Tris, pH 7.5, 1 M MgCl₂, 50% PEG, 0.1 M ATP and 1 M DTT for blunt end ligations, and incubated at 10°C for 16 h. Restriction enzymes were heat inactivated at 65°C prior to ligation.

2.12.3 In vitro cloning

DNA to be subcloned (200 ng) was cleaved in either single or double restriction enzyme digests. This was combined with 20 ng of similarly cleaved vector DNA, then ligated with 2 units of T4 DNA ligase in a volume of 50 μl in a final buffer concentration of 20 mM.
Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM dithiothreitol (DTT), 0.6 mM ATP for 16 h at 4°C. The ligated DNA was then used directly for transformation of *E. coli* strains. Transformants were screened for insertional inactivation of the appropriate drug resistance gene (eg. Ap or Tc), wherever possible, prior to plasmid DNA isolation.

2.12.4 Cosmid Cloning

High molecular weight genomic DNA prepared from strain PE655 was partially digested with *Sau3AI*. Samples were taken at specific time intervals and aliquots were electrophoresed on a 0.8% (w/v) agarose gel in TAE buffer. DNA fragments of approximately 35-45 kb were excised from the gel, electro-eluted and extracted using TE-saturated phenol/chloroform. The purified DNA was ligated to pPM2101 which had previously been completely digested with *BamHI* and dephosphorylated with CIAP. Ligation was allowed to proceed initially at room temperature for 10 h, and then overnight at 4°C. The ligation mix was packaged into bacteriophage λ (Collins and Hohn, 1978) using the *in vitro* Packagene System (Promega) and then transduced into *E. coli* S17-1.

2.13 Sequencing methods

2.13.1 Sequencing using dye labelled primers

Sequencing reactions were carried out on 1 µg of double stranded plasmid DNA using the protocol provided Applied Biosystems. In dye-labelled primer sequencing the DNA was split into four tubes containing 160 ng (A and C) and 320 ng (G and T) of DNA respectively. To each tube Ready reaction mix and DNA template were added:
Reagent       | A | C | G | T
---|---|---|---|---
Ready Reaction Mix | 4μl | 4μl | 8μl | 8μl
DNA Template | 1μl | 1μl | 2μl | 2μl
Total Vol. | 5μl | 5μl | 10μl | 10μl

Each reaction was overlayed with 20 μl of light mineral oil and centrifuged briefly.

Samples underwent: 15 cycles (96°C 10 sec; 55°C 5 sec; 70°C for 60 sec), followed by 15 cycles (96°C 10 sec; 70°C 60 sec; 15 cycles total), and were then held at 4°C. Reactions were combined in 80 μl of 95% (v/v) ethanol with 3 μl of 3 M sodium acetate and precipitated on ice. DNA was pelleted at 13,000 rpm for 15 min (Hereaus bench microfuge). Samples were dried *in vacuo* and stored at -20°C.

2.13.2 Sequencing with dye-labelled terminators

Plasmid DNA was purified prior to dye terminator sequencing with kits supplied by Boehringer Mannheim. 0.5 ml thin walled tubes (Gene Amp, Perkin Elmer) containing 1-2 μg of template DNA and 3.2 pmol primer, made up to a final volume of 20 μl with 9.5 μl of "Go" pre-mix (Boehringer Mannheim) and sterile water, were overlaid with mineral oil (Nujol, Perkin Elmer) and subjected to 25 cycles (96°C 30 sec; 50°C 15 sec; 60°C 4 min) before adding 2 μl of 3 M sodium acetate and 50 μl of ice cold 100% ethanol and precipitating for 2 h at -20°C then washing with 70% ethanol and drying *in vacuo*.

2.13.3 Analysis of DNA sequences

The dried pellets for sequencing were stored at -20°C until required, when they were resuspended in 4.5 μl loading buffer (83% deionised formamide, 8.3 mM EDTA pH 8.0),
heated (95°C for 2 min), and electrophoresed on a 6% polyacrylamide-8M urea gel in an
Applied Biosystems 373A DNA sequencer. Raw sequencing data from the 373A automated
sequencer were analysed using the Applied Biosystems Seq Ed program version 6.0. Service
provided by Sequencing Laboratory, IMVS, Adelaide. Sequencing data were analysed using
the LKB DNA and protein analysis programs, DNASIS and PROSIS (Hitachi Software).

2.14 Polymerase Chain Reaction Protocol (PCR)

The protocol used for PCR is that described by Delidow (1993) for the generation of
PCR products with cohesive ends. The PCR reaction was performed in reaction tubes (0.5 ml,
Perkin Elmer) in a 50 µl volume containing Taq buffer (50 mM KCl; 10 mM Tris-HCl pH
8.3; 1.1 mM MgCl₂, 0.01% (w/v) gelatin), 2 µM each deoxynucleoside triphosphate (dNTP),
100 pmol each primer, 200 ng of plasmid template or genomic DNA and 2.5U Taq
polymerase (Perkin Elmer). The reaction was overlaid with a drop of light mineral oil (Nujol,
Perkin Elmer) and following an initial denaturation period of 5 min at 95°C, was subjected to
25 cycles of amplification (95°C for 30 sec, 55°C for 30 sec, 72°C for 1 min followed by a
final extension at 72°C for 5 min) using a DNA thermal cycler (Perkin Elmer). Following
PCR, the reaction was carefully removed from under the oil and the DNA precipitated by the
addition of 0.5 volumes of 3 M sodium acetate plus 2.5 volumes of 95% ethanol and
incubated at RT for 10 min. The DNA was collected by centrifugation at 15,000 rpm for 15
min (Heraeus Biofuge 15), washed with 70% ethanol, dried in vacuo and resuspended in 20 µl
of sterile water.

2.14.1 Synthesis of oligodeoxynucleotides

Oligodeoxynucleotides were synthesized on an Applied Biosystems 381A DNA
synthesizer in the trityl-off mode and butanol extracted prior to use. Reagents were purchased
from Applied Biosystems or Ajax Chemicals. The oligonucleotides used in this study are listed in Table 2.4. Also purchased from Sequencing Laboratory, IMVS, Adelaide.

2.15 Southern techniques

2.15.1 Preparation of DIG-labelled DNA probes

DNA fragment and plasmid probes were labelled with digoxigenin-11-dUTP (Boehringer Mannheim) according to the manufacturers protocol, using random-labelling or end-labelling (oligonucleotides and small fragments). Random-primed labelling was performed with heat denatured (95°C 10 min) DNA (10 ng - 3 μg) chilled on ice (3 min) prior to addition of 2 μl each of hexanucleotide mix and dNTP labelling mix (Boehringer Mannheim) and 2 U Klenow fragment DNA polymerase I. This was made up to 20 μl with sterile distilled water and held at 37°C for 60 minutes. The reaction was stopped with 2 μl of 0.2 M EDTA pH 8.0, and DNA precipitated with ethanol and LiCl (75 μl ethanol 100%; 2 μl LiCl 4.0 M) on ice before washing, vacuum drying and resuspension in 20 μl of TE (10 mM Tris-HCl, 1 mM EDTA pH 8.0).

2.15.1.1 DIG labelled PCR

The protocol used for DIG-labelled PCR is essentially that described in Section 2.14. The PCR reaction was performed in 0.5 ml reaction tubes in a 50 μl volume containing Taq buffer, 200μM each of deoxynucleoside triphosphate dATP, dCTP and dGTP,190 μM of dTTP and 10 μM of DIG-11-dUTP, 200 ng of plasmid template or genomic DNA and 2.5 U of Taq polymerase (Perkin Elmer). The reaction was overlaid with a drop of light mineral oil and subjected to the cycle detailed in Section 2.14.
### Table 2.4 Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>747</td>
<td>CCGGAATTCGGATCTTTATGAAGATATTC</td>
<td>gtrX primer with EcoRI site</td>
</tr>
<tr>
<td>748</td>
<td>CCGGATACCGTTTGGCTTTACATTAC</td>
<td>gtrX primer with BamHI site</td>
</tr>
<tr>
<td>860</td>
<td>TTCCCTATTTTGGATCTTTATTTTGGTCTC</td>
<td>rfc primer with BamHI site</td>
</tr>
<tr>
<td>2003</td>
<td>TAGGGATCCGGACGATATTGG</td>
<td>rfc primer with BamHI site</td>
</tr>
<tr>
<td>2059</td>
<td>TCCCTATTGTTGCTTTATTTTGGCTCCAG</td>
<td>rfc primer with SphI site</td>
</tr>
<tr>
<td>2060</td>
<td>TAGGCATCGCGACGATATTGG</td>
<td>rfc primer with SphI site</td>
</tr>
<tr>
<td>2210</td>
<td>AGATTTAGCGCCAGTCCCAGGCGC</td>
<td>primer in TnphoA</td>
</tr>
<tr>
<td>2254</td>
<td>TACGCGAGCTGCGAGAAG</td>
<td>primer in TnphoA</td>
</tr>
<tr>
<td>2211</td>
<td>CCAATCGCAATACGAGGGC</td>
<td>nt. 5051 → 5034&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>2325</td>
<td>CTTGAAAAGGAGGCGC</td>
<td>nt. 1834 → 1851</td>
</tr>
<tr>
<td>2326</td>
<td>TACAGGAAGTAGGAGGAGC</td>
<td>nt. 3457 → 3440</td>
</tr>
<tr>
<td>2327</td>
<td>CCAATCCCGATCAGTTCG</td>
<td>nt. 3711 → 3693</td>
</tr>
<tr>
<td>2328</td>
<td>GCCTGCGGCGGTTGTTGCTCA</td>
<td>nt. 3150 → 3167</td>
</tr>
<tr>
<td>2329</td>
<td>GAAGAGCATACTGAGCTGG</td>
<td>nt. 2796 → 2780</td>
</tr>
<tr>
<td>2330</td>
<td>GCAGGTCTCAGAACAGCC</td>
<td>nt. 2320 → 2300</td>
</tr>
<tr>
<td>2331</td>
<td>CCCAGTGTAAGTGTGTTTCAGC</td>
<td>nt. 2546 → 4527</td>
</tr>
<tr>
<td>2345</td>
<td>AGAGCGAAAGGATGCTTAAAGC</td>
<td>nt. 4932 → 4952</td>
</tr>
<tr>
<td>2346</td>
<td>CGATAAGCTGAAATAATGTTGG</td>
<td>nt. 4320 → 4339</td>
</tr>
<tr>
<td>2347</td>
<td>GTTAAACCTATACAGTGGG</td>
<td>nt. 1711 → 1694</td>
</tr>
<tr>
<td>2348</td>
<td>ACCGCATCTCTCGACC</td>
<td>nt. 2402 → 2420</td>
</tr>
<tr>
<td>2349</td>
<td>TAGCTCTCTGCGCAGATGC</td>
<td>nt. 3684 → 3702</td>
</tr>
<tr>
<td>2350</td>
<td>AGAGCGAATCCGCTGGGGCG</td>
<td>nt. 4073 → 4055</td>
</tr>
<tr>
<td>2369</td>
<td>GCAACCAGCAGCAGAATGC</td>
<td>nt. 1001 → 982</td>
</tr>
<tr>
<td>2615</td>
<td>TCAGGTATTCCGACAGTTT</td>
<td>nt. 867 → 884</td>
</tr>
<tr>
<td>2623</td>
<td>CGCGCCCTCGCAATTGCC</td>
<td>nt. 443 → 426</td>
</tr>
<tr>
<td>2624</td>
<td>AGCTTCCGCTTCATCTGG</td>
<td>nt. 390 → 408</td>
</tr>
<tr>
<td>2625</td>
<td>GCTTATGAAATCGACCTACC</td>
<td>binds to proA (ECAE000132)</td>
</tr>
<tr>
<td>2370</td>
<td>CATTACACCGGATGCAGCC</td>
<td>binds to yagB (ECAE000132)</td>
</tr>
<tr>
<td>2371</td>
<td>GCCTGACCCAGCTGGAAC</td>
<td>Primer for phoAllacZ fusions of Bgt</td>
</tr>
<tr>
<td>2719</td>
<td>GGTCGAGAATTGTGATACC</td>
<td>Primer for phoAllacZ fusions of Bgt</td>
</tr>
<tr>
<td>2720</td>
<td>GCCTGAGGTTTACCAAGAAC</td>
<td>Primer for phoAllacZ fusions of Bgt</td>
</tr>
<tr>
<td>2721</td>
<td>TCCTGCGAATACCGATGAGA</td>
<td>Primer for phoAllacZ fusions of Bgt</td>
</tr>
<tr>
<td>2722</td>
<td>ATGGTCTAGAAGGGATGCG</td>
<td>Primer for phoAllacZ fusions of Bgt</td>
</tr>
</tbody>
</table>

<sup>a</sup> Nucleotide ranges refer to Figure 4.3.
2.15.2 Southern transfer and hybridisation

Unidirectional transfer of DNA from agarose gels to Hybond-N+ nylon transfer membrane (Amersham) were performed as described by Southern (1975) and modified by Maniatis et al. (1982).

Prior to hybridization with DIG-labelled probe, filters were incubated for 4 hr at 42°C in a pre-hybridization solution containing 50% (v/v) formamide, 1% (w/v) skim milk, 7% (w/v) SDS, 250 µg/ml single stranded herring sperm DNA (Sigma), 5 x SSPE (Maniatis et al., 1982). Pre-hybridization fluid was discarded and replaced with fresh hybridization buffer. Denatured probe was added and hybridization allowed to occur for 16-24 h at 42°C.

Filters were washed twice with shaking at room temperature (RT) for 5 min in 2 x SSC, containing 0.1% (w/v) SDS. This was followed by two further 15 min washes in 0.1 x SSC plus 0.1% (w/v) SDS at 65°C. After drying in the air (15 min, room temperature), the filters were washed with buffer 1 solution (0.1 M Tris-HCl, 0.15 M NaCl pH 7.5) and then incubated with 5% skim milk in buffer 1 for 60 min at RT. Anti-DIG AP (alkaline phosphatase coupled to anti-DIG Fab fragments) (2 µl in 10 ml buffer 1) is then added for 30 min at RT. This is followed by 2 x 15 min washes with buffer 1 and 1 x 2 min wash with Buffer 3 (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM MgCl2). Alkaline phosphatase-labelled goat anti-rabbit antibody was detected by the addition of nitroblue-toluidine (NBT) and bromochloroindolyl phosphate (BCIP) as described.

2.16 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Western transfer

2.16.1 SDS-PAGE

For analysis by SDS-PAGE, bacterial suspensions (ca. 10^9/ml) were prepared by resuspending bacterial pellets in saline and then adding an equal volume of 2 x sample buffer (0.25 mM Tris-HCl pH 6.8, 2% (w/v) glycerol, 5% (v/v) β-mercaptoethanol, 15% (w/v)
bromophenol blue) prior to loading. Samples were heated at 100°C for 3 min before loading 10 μl/well onto a 15% polyacrylamide gel.

SDS-PAGE was performed on a 5% stacking and 15% separating polyacrylamide gels using a modification of the procedure described by Lugtenberg et al., (1975). Gels were 15 cm long, 11 cm wide and 1.5 mm thick. Samples heated at 100°C for 3 min in SDS sample buffer consisting of 25 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, and 15% (w/v) bromophenol blue prior to loading. Gels were electrophoresed through the gel at 150 V for 2-3 hr. Gels were stained overnight with Coomassie Brilliant Blue G250 in 50% (v/v) methanol and 10% (v/v) acetic acid, 10% (v/v) methanol and 10% (v/v) ethanol over 24 h.

Size markers (Pharmacia) were phosphorylase B (94 kD), bovine serum albumin (67 kD), ovalbumin (43 kD), carbonic anhydrase (30 kD), soybean trypsin inhibitor (20.1 kD) and alphalactalbumin (14.4 kD).

2.16.2 Western transfer and detection

Protein were transferred to nitrocellulose (Schleicher and Schuell) at 200 mA for 2 h in a Trans-blot cell (Bio-Rad). LPS is transferred onto nitrocellulose (Schleicher and Schuell) at 500 mA for 1 h at 4°C in a Trans-blot cell. The transfer buffer used consisted of 25 mM Tris-HCl, pH 8.3, 192 mM glycine and 20% (v/v) methanol (Towbin et al., 1979). The blot was incubated for 1 h in blotto (5% skim milk powder in TTBS (0.05% (v/v) Tween 20, 20 mM Tris-HCl, pH 7.4, 0.9% (w/v) NaCl)) to block non-specific binding sites before incubating in primary antiserum (diluted in blotto) for 2-16 h at RT. Unbound antibody was removed by washing the filter (3 x 10 min) in TTBS before incubation with HRP- conjugated goat anti-rabbit IgG (diluted 1:5000 in blotto) for 90 min at RT. Prior to detection, the filter was washed (4 x 5 min) with TTBS and then (2 x 5 min) with TBS (20 mM Tris-HCl, pH 7.4, 0.9% (w/v) NaCl).
The antigen-antibody complexes were then visualised using peroxidase substrate (9.9 mg 4-chloro-1-naphthol dissolved in 3.3 ml -20°C methanol added to 16.5 ml TBS containing 15 μl hydrogen peroxide) which was added and allowed to incubate for 10-15 min with shaking as described by Hawkes et al., (1982).

For ECL detection, filters were washed prior to detection in 20 ml PBS for 2 x 5 min, and then soaked in ECL detection reagent (Boehringer Mannheim) for 1 min in a transparent plastic bag. Finally, the filters were exposed to X-ray film at room temperature.

2.17 Lipopolysaccharide (LPS) preparation

LPS for analysis by SDS-PAGE followed by silver staining was prepared as follows. A volume of 1 ml of cells grown in liquid culture to the stationary phase (ca. 5 x 10⁹ cells/ml) was pelleted in an Eppendorf tube. The pellet was suspended in 50 μl of lysing buffer (2% (w/v) SDS, 4% (v/v) β-mercaptoethanol, 10% (v/v) glycerol, 1 M Tris-HCl pH 7.6, 0.1% (w/v) bromophenol blue), and boiled for 5 min. Proteinase K (10 μl of a 2.5 mg/ml solution in lysing buffer) (Hitchcock and Brown, (1983) was added and the mixture was incubated at 55°C for 2 h. Samples (10-20 μl) were heated to 100°C for 3 min prior to SDS-PAGE.

2.17.1 LPS-specific silver staining

Silver staining of LPS was performed using a modification of the method described by Tsai and Frasch (1982). Following electrophoresis the procedure used was to fix SDS-PAGE gels for 2-18 h in a solution of 40% (v/v) ethanol, 10% (v/v) acetic acid, then oxidised for 5 min with in a solution of 0.7% (v/v) periodic acid in 40% (v/v) ethanol, 10% (v/v) acetic acid. After oxidation the gels were washed 4 times for 30 min each with water and then stained for 10 min in a solution composed of 28 ml sodium hydroxide (0.1 M), 2 ml ammonium hydroxide (30% (w/v)) and 5 ml of silver nitrate (20% (w/v)). The stained gels were developed in a prewarmed (37°C) solution of 50 mg/ml citric acid and 0.05% (w/v)
formaldehyde. Staining was stopped using a solution of 20% (v/v) methanol and stored in Milli Q water.

2.17.2 Colony immunoblotting of LPS

The procedure used was a modification of that described by Towbin et al., (1979). 1 ml of overnight samples were centrifuged (5000 rpm, 10 min, bench centrifuge) and resuspended in 100 µl of saline. 5 µl of each sample was spotted onto nitrocellulose membrane (Scheicher and Schuell) and treated in the same manner as described for western immunoblotting, using 1: 1000 dilution of primary antiserum and 1: 5000 dilution of secondary antiserum goat anti-rabbit HRP. Detection was carried out using the hydrogen peroxide (as described in 2.16.2).

2.17.3 Bacterial agglutination assays

Fifty microlitres of appropriate antisera were diluted (serially) in 96 well round bottom microtitre trays. Overnight cultures were washed in 1 x PBS and resuspended to a final concentration of 5 x 10⁹ cells/ml, and 50 µl added to each well containing antiserum. The tray was incubated for 1-2 h at 37°C followed by incubation at 4°C for 2-16 h. Extent of agglutination was determined visually.

2.17.4 O antigen hydrolysis assay

2.17.4.1 Preparation of formaldehyde (formalin)-fixed cells

Overnight cultures of cells grown in 10 ml LB at 37°C were pelleted (5,000 rpm, 10 min, IEC bench centrifuge) and washed twice with saline. The pellet was then resuspended in a 1 ml of a 1% formalin solution in saline and place at 37°C for 1 h with aeration. Treated cells were pelleted (10¹⁰ cells), resuspended in 1 ml of saline and stored at 4°C.
2.17.4.2 O antigen hydrolysis assay

0.1 ml of formalin-fixed cells were mixed with equal volumes of Sf6c phage (10^{12} pfu) and TM buffer (10 mM Tris HCl pH 7.1, 10 mM MgSO_4). The suspension was then incubated at 37°C for 30 min. The cells were then centrifuged (15,000 rpm, 1 min, Heraeus microfuge) and washed twice in Milli-Q water. LPS samples were then prepared, subjected to SDS-20% PAGE and visualised by silver staining.

2.18 T7 polymerase / promoter expression system and L-[^35]S]-methionine labelling

The plasmid pGP1-2 carries the T7 RNA polymerase under the control of the lambda P_L promoter (Tabor and Richardson, 1985). This plasmid was transformed into E. coli strains containing a plasmid with the specific gene of interest under control of the T7 RNA polymerase promoter. A 10 ml LB broth with ampicillin and kanamycin was inoculated with a single colony and grown with aeration at 30°C overnight. The culture was subcultured 1:10 and incubated with constant aeration at 30°C. When an A_{600} OD of 0.6 was reached, the cells were pelleted and resuspended in 1.5 ml of M9 media supplemented with 0.01% of 18 amino acids (minus methionine) and grown with aeration at 30°C for a further 120 min. The cells are then incubated at 42°C for 20 min to induce the pGP1-2 P_L promoter by the inactivation of cI_{ts}, allowing expression from the λP_L promoter. Rifampicin was added to a final concentration of 200 mg/ml to inactivate the E. coli RNA polymerase and incubation was continued at 42°C for a further 20 min. The culture was then left for at least 2 h shaking at 37°C. Samples were then pulsed with 10 μCi of L-[^35]S]-methionine for 5 min at 30°C. 1 ml of culture was transferred to microfuge tubes, centrifuged (15 K, 2 min, Heraeus microfuge) to pellet the cells and resuspended in 100 μl of 1 x SDS sample buffer. 10 μl of sample was heated at 100°C for 2 min and loaded onto SDS-PAGE gels for analysis. Gels were subsequently stained with Coomassie G250.
2.18.1 Cell Fractionation

The cell fractionation procedure was a modification of that described by Osborn et al., (1972). Cells were grown in LB to mid-exponential phase at 37°C (50 ml, OD_{600} of 0.6). Cells were pelleted in a Beckman SS-34 rotor, (10,000 rpm, 10 min, 4°C) and resuspended in 1 ml of 20% (w/v) sucrose, 30 mM Tris-HCl pH 8.1, transferred to SM-24 tubes and chilled on ice. Cells were converted to sphaeroplasts with 0.1 ml of 1 mg/ml lysozyme in 0.1 M EDTA pH 7.3 for 30 min on ice. Cells were centrifuged as above and the supernatant collected (periplasmic fraction). The cell pellet was frozen in an ethanol dry ice bath for 30 min, thawed and dispersed vigorously in 3 ml 3 mM EDTA, pH 7.3. Cells were lysed with a Branson Ultrasonifier (45% cycle, intermittent), by successive freeze-thawing. Unlysed cells and large cell debris were removed by low speed centrifugation (7,000 rpm, 5 min, 4°C). The supernatant containing the membranes and the cytoplasm was centrifuged at 35,000 rpm in a 50 Ti rotor for 60 min at 4°C in a Beckman L8-80 ultracentrifuge. The supernatant (cytoplasmic fraction) was collected and the membrane pellet was resuspended in 1 ml 25 mM Tris-HCl pH 7.5. Five hundred microlitres of Triton solution (4% Triton X-100, 2 mM MgCl_{2}, 50 mM Tris pH 7.5) was added to an equal volume of the membrane sample which was vortexed for 30 min. The inner (soluble) membrane fraction was separated from the outer (insoluble) membrane fraction by centrifugation at 35,000 rpm for 90 min in a 50 Ti rotor (Beckman L8-80) at 4°C. The outer membrane fraction was dissolved in 0.25 M Tris pH 7.5.

2.18.2 Autoradiography

SDS-PAGE gels were dried on Whatman 3MM chromatography paper at 60°C for 2 h on a Bio-Rad gel drier. [^{35}S]-methionine autoradiography was performed at room temperature for 1 -7 days without intensifying screens using Kodak XR-100 film.
2.19 Alkaline phosphatase and β-galactoside assays

2.19.1 Alkaline phosphatase assays

Alkaline phosphatase assays were performed by a method modified from that previously described by Manoil (1991). Cultures were grown in LB with Ap at 37°C for 16 h then subcultured 1/20 into fresh LB containing Ap and incubated until the OD_{600} ca. 0.5. 1 ml of culture was centrifuged for 3 min at 14,000 rpm and washed in cold 10 mM Tris-HCl, pH 8.0, 10 mM MgSO₄ and the final pellet resuspended in 1 ml of cold 1 M Tris-HCl, pH 8.0, 1 mM iodoacetamide. The OD_{600} was measured by placing 300 μl in the well of a 96 well microtitre tray and reading on a DYNATECH MR 7000 microplate reader. Washed culture (200 μl) was added to 800 μl of 1 M Tris-HCl, pH 8.0, 0.1 mM ZnCl₂, 1 mM iodoacetamide and permeabilised by the addition of 50 μl 0.1% SDS and 50 μl chloroform. 275 μl of permeabilised cells were placed in duplicate wells of a microtitre tray and the reaction was started by the addition of 25 μl of 0.4% p-nitrophenyl phosphate (in 1 M Tris-HCl, pH 8.0). Optical densities were recorded at 410 nm (colour change) and 570 nm (cell debris). PhoA units were calculated using the standard equation (Manoil 1991), shown below.

2.19.2 β-galactosidase assays

β-galactosidase assays were performed as described in Baker et al., 1997. Cultures were grown in LB at 37°C with aeration for 16 h. They were cooled on ice for 20 min and their OD_{600} was measured (Pharmacia LDB-Ultraspec Plus Spectrophotometer). Dilutions were made using cold LB so as all cultures had an OD_{600} of 0.550 - 0.850. Diluted cultures were then aliquoted (300μl) into 96 well flat-bottomed microtitre trays and the OD_{600} read immediately. In a microfuge tube, diluted culture (500 μl) was added to 500 μl of Z buffer (60 mM Na₂HPO₄·7H₂O, 40 mM NaH₂PO₄·H₂O, 10 mM KCl, 1 mM MgSO₄·7H₂O, 50 mM β-mercaptoethanol adjusted to pH 7.0 and stored at 4°C). Cells were opened by addition of 2
drops of chloroform and 1 drop of SDS (0.1% w/v) and vortexing for 10 s. Aliquots (100 µl) of the bacterial lysates were placed in duplicate wells of a 96 well microtitre tray. The tray was incubated at 28°C for 15 min and the reaction started by addition of 2-nitrophenyl-\(\beta\)-D-galactopyranoside (ONPG) (200 µl) (0.7 mg/ml in Z buffer) to each well, with mixing. Optical densities were recorded at 410 nm (O-nitrophenol colour change) and 570 nm (cell debris). Enzyme units were calculated for each strain at the time point when the OD\(_{410}\) first exceeded 0.6 using the following equation.

\[
\beta\text{-galactosidase units} = \frac{1000 \times (\text{OD}_{410} - 1.75 \times \text{OD}_{570})}{\text{Time} \times \text{Volume} \times \text{OD}_{600}}
\]

2.20 Bacteriophage techniques

2.20.1 Construction of lysogens

The phage (10^9/ml) were spotted for single plaques onto a soft agar lawn of the bacterial strain to be lysogenized and the plates incubated at the appropriate temperature for 6 h. The centre of a turbid plaque was streaked for single colonies on the appropriate agar and the plate incubated at 37°C for 16 h. Single colonies were selected and purified three times by streaking for single colonies before assessing reactivity to type II antiserum by colony immunoblotting.

2.20.2 Phage assays

SfII phage lysates were assayed for phage titre by mixing 0.1 ml of phage diluted in LB, 0.1 ml of log phase indicator bacteria and incubating for 15 min in a 37°C water bath. 3 ml of melted (0.7% (w/v)) soft agar overlay was added and the mixture poured onto pre-warmed LA plates and incubated at 37°C for 16 h.
2.20.3 Preparation of phage stocks

2.20.3.1 Low titre phage stocks

Low titre stocks (10^9 -10^10 pfu/ml) of SfII were prepared by the Plate Lysate Method (Maniatis). Plate stocks were prepared from a single plaque. A single phage plaque was removed with a sterile pasteur pipette and resuspended in 1 ml of LB for 2 h at 4°C. The required bacterial cells (S. flexneri Y serotype PE577) were cultured in LB to late log phage (A_600 = 0.8). 0.1 ml of cells were mixed with 0.1 ml of eluted bacteriophage for 15 min at 37°C. The mixture was then plated with 3 ml of soft agar onto pre-warmed LA plates and incubated for 6 h at 37°C. The soft agar overlay was then scraped off and added to 10 ml LB and centrifuged (6,000 rpm, 10 min) to remove agar and bacterial debris. The supernatant was stored with 0.5% (v/v) chloroform at 4°C.

2.20.3.2 High titre phage stocks by liquid infection

Essentially as described by Maniatis et al., (1989), a fresh overnight culture of PE577 was diluted one hundred fold into two aliquots of 500 µl of LB and incubated at 37°C with aeration to A_600 = 0.5. The culture was infected at a multiplicity of addition of 0.1 with phage from a low titre SfII stock and incubation was continued at 37°C for 4 h or until lysis was evident. Chloroform (0.5% (v/v)) was then added, the culture was returned to 37°C for 15 min to lyse remaining cells. Bacterial debris was removed by centrifugation (10, 000 rpm, 4°C 20 min, JA-20 rotor) and the supernatant decanted. NaCl and PEG-8000 were added to a final concentration of 0.5 M and 10% (w/v) respectively, and precipitation was allowed to proceed overnight at 4°C. The precipitate was collected by centrifugation (9,000 rpm, 4°C, 20 min, JA-20 rotor), resuspended in 5 ml of TM buffer. If required the phage were purified by CsCl block gradient centrifugation.
2.20.3.3 CsCl block density gradient for preparation of high titre phage stocks

CsCl block density gradient centrifugation was used to prepare high titre SIII phage stocks for the preparation of DNA by phenol extraction.

Three CsCl solutions of density 1.45 g/ml, 1.5 g/ml and 1.70 g/ml were prepared in sterile TM and were used to form a block gradient by adding 4 ml of 1.45 g/ml solution and underlaying it with 1 ml of the 1.5 g/ml followed by the 1.7 g/ml solution in a 10 ml polycarbonate Oakridge tube. The high titre phage suspension in TM was carefully layered on top of the gradient and the tube centrifuged (45,000 rpm, 90 min, 4°C, Beckman Ti-50 rotor). The opaque phage band was collected by piercing the side of the tube with a syringe and dialysed three times against one litre of TM and stored at 4°C. A titre of $10^{12} - 10^{13}$ pfu/ml was usually obtained by this method.

2.20.3.4 Phenol extraction of bacteriophage DNA

A high titre phage stock ($10^{12} - 10^{13}$ pfu/ml) was diluted to 0.9 ml in TE and then 0.1 ml of 100 mM Tris-HCl pH 8.0, 100 mM NaCl, 100 mM EDTA, 5.0% SDS was added, followed by 5 mg of Proteinase K. After incubation at 37°C for 60 min, the solution was again diluted (to 5 ml) with TE, and an equal volume of TE-equilibrated phenol added. The mixture was gently shaken for 5 min and the phases were separated by centrifugation (7,000 rpm, 5 min, 20°C, JA-20 rotor). The aqueous phase was collected and re-extracted at least twice with an equal volume of TE-equilibrated phenol. The phenol phases were washed with an equal volume of TE, and sodium acetate (pH 4.6) was added to the pooled aqueous phases to a final concentration of 0.3 M, followed by 2.5 volumes of ethanol. DNA was precipitated at -20°C for 30 min and was collected by centrifugation (18,000 rpm, 20 min, 4°C, JA-20 rotor). The pellet was washed in 70% and 95% ethanol (v/v), dried in vacuo and dissolved in 1 ml TE and stored at 4°C.
**2.20.3.5 Bacteriophage sensitivity tests**

Bacteria were grown to stationary phase in either NB or LB and swabbed across a NA plate. Bacteriophages were spotted (ca. $10^6$ pfu in 5μl NB) onto the bacteria and the plates were incubated at 37°C for 16 h. Strains were scored as bacteriophage sensitive if a turbid plaque or lysis resulted.

**2.21 Electron microscropy (EM)**

For EM, bacteriophage were spotted onto a lawn of indicator bacteria and incubated for 6 h at 37°C or CsCl purified phage were spotted (5 μl) onto Whatman parafilm and Colloidin-coated copper grids (300 mesh, TAAB) were place coated side down for 1 min. The grid was then transferred to a 40 μl drop of 2% phosphotungstic acid (PTA) for 10 sec, blotted and stored in a dust free environment. Grids were examined using a Phillips TM-100 electron microscope.
Chapter Three

Isolation and characterisation of *S. flexneri* serotype converting bacteriophage SfII.

### 3.1 Introduction

Bacteriophage SfIII encodes the gene(s) required for the expression of the type II antigen. The type II antigen is characterised by the addition of a glucosyl residue to rhamnose III of the repeating units of the O polysaccharide of *S. flexneri*. In comparison with the other serotypes, strains of serotype 2a are of particular interest due to their greater prevalence, which may be an indication of their relative virulence (Noriega *et al.*, 1995). Many factors, including bacteriophage-encoded functions or an alteration in LPS, may be contributing to this feature.

The isolation of bacteriophage SfIII and its characterisation is required in order to understand the mechanism by which it infects and modifies the LPS. Determination of the gene(s) involved in serotype conversion and other functions encoded by this phage could provide information relating to its evolution and potential role in virulence of the organism.

This chapter describes the isolation of bacteriophage SfIII from a strain of serotype 2b, NCTC4 (Giammanco, 1968). The host range of the phage is assessed and the genome and region involved in serotype conversion are characterised. Bacteriophage SfIII lysogens have also been isolated and their LPS characterised by silver staining of PAGE gels and western immunoblotting.
3.2 Results

3.2.1 Attempts to isolate bacteriophage SfII.

Strains expressing the type II antigen contain the genome of the temperate phage SfII integrated into their chromosome and, therefore, the genes responsible for serotype conversion. Isolation of bacteriophage SfII was initially attempted by various techniques using PE655 of serotype 2a. Bacteriophage are often released spontaneously from lysogens (Lwoff and Gutmann, 1950); the supernatant of a culture of PE655 was harvested and used in a plaque assay with the indicator strain PE577 (serotype Y). As bacteriophage plaques were not obtained by this method, several other approaches were taken to induce phage from the lysogen. Physical and chemical stresses such as heat, chloroform, U.V. irradiation and the addition of mitomycin C all proved to be unsuccessful in the induction of bacteriophage. These methods did not result in release of bacteriophage from a variety of serotype 2a strains (PE523, PE567, PE574, PE642 and PE790) from the laboratory collection.

3.2.2 Cosmid cloning of serotype conversion genes from strain PE655

A cosmid bank was constructed using DNA from strain PE655, and the cosmid vector pPM2101 (Table 2.3) and mobilised into PE577. Colony immunoblotting with anti-type II serum was used to screen for serotype conversion from Y to 2a. Three cosmids, pRMM163, pRMM164 and pRMM165 which converted PE577 to serotype 2, were isolated. Subsequent attempts at subcloning these cosmids were unsuccessful as they were unstable, with spontaneous deletions occurring, resulting in loss of the serotype-converting ability.
3.2.3 Isolation of bacteriophage SfII from NCTC4

NCTC4, of serotype 2b, is lysogenised by phages SfX (f7,8) and SfII (fII) which are responsible for group 7,8 and type II antigens, respectively (Giammanco, 1968). A lysate from heat-induced NCTC4, when used in plaque assays, gave two plaque types; these were either large or small. Isolation and purification of either plaque type consistently resulted in the appearance of plaques of both types. It was previously reported that the two plaque morphologies could be separated, with large plaques being caused by SfX and small plaques by SfII (Giammanco, 1968), however, this could not be confirmed.

Lysogenisation of PE577 with the phage propagated from either large or small plaques resulted in the expression of the type II antigen. No conversion to serotype X was seen with PE577 (500 colonies screened, 492 type II positive, 8 remained group 3,4 positive only). Upon induction, the lysates from these lysogens were also able to mediate serotype conversion to serotype 2. The bacteriophage in the lysate from one lysogen RMM273 was therefore named SfII.

3.2.4 Purification and characterisation of bacteriophage SfII.

Bacteriophage SfII was purified by plating the heat-induced lysate from RMM273 with the indicator strain PE577 and isolating a single plaque. This procedure was repeated three times to ensure only a single plaque was isolated. The lysogens resulting from infection with plaque-purified SfII were assayed by colony immunoblotting with anti-type II serum and shown to express the type II antigen. Propagation of the phage was carried out on solid media followed by liquid media once a titre of $10^{10}$ pfu/ml had been reached. The purified phage were examined by electron microscopy (Fig. 3.1) and it was seen that SfII belonged to the
Figure 3.1  Electron micrograph of SfII.

Electron micrograph of bacteriophage SfII, stained with 3% phosphotungstic acid. Magnification 205 K. The head, neck and contractile tail are clearly visible. Two ‘curled’ tail fibres can also be seen. Morphologically, SfII belongs to the Bradley group A (1967).
group A of Bradley's classification possessing a hexagonal shaped head, a tail with a contractile sheath and tail fibres (Bradley, 1967).

3.2.4.1 Proteins of bacteriophage SfII

Purified SfII phage was solubilised in sample buffer, electrophoresed in an SDS-15% PAGE gel and stained with Coomassie Blue. Figure 3.2 shows three major protein bands at approximately 35, 36 and 40 kDa and minor bands at 15, 17, 19 and 67 kDa. The structures corresponding to these protein bands are unknown, however, in bacteriophage Mu, belonging to the same morphological group, the major head protein is 33 kDa, tail sheath protein 55 kDa and tail proteins range between 16 and 64 kDa (Howe, 1987). The bands shown in figure 3.2 may also be representative of these structures due to the morphological similarity between SfII and phage Mu.

3.2.4.2 O-antigen hydrolysis assays

O-antigen hydrolysis assays assess the ability of the bacteriophage to cleave at specific positions along tetrasaccharide repeat units. A shortened LPS results from the cleavage of the O-antigen which provides access to the surface of the bacterium. The phage then adsorbs to the surface and ejects its DNA. Bacteriophage Sf6 possesses the ability to cleave the O-antigen of strains of S. flexneri of serotypes Y and X (Lindberg et al., 1978, Van den Bosch et al., 1997). Assays using bacteriophage SfIII were carried out with strains PE577 of serotype Y and PE576 of serotype X (Fig 3.3). It can be seen that no differences could be detected in O-chain length between the control lanes (PE577 and PE576) without phage and the lanes (PE577 + SfII and PE576 + SfIII) with SfII added. Therefore, it can be concluded that during infection with phage SfII, cleavage of the O-antigen is unlikely to occur.
Figure 3.2  Coomassie Blue stained SDS-15% PAGE displaying the structural proteins of bacteriophage SfIII. Major proteins of sizes 35, 36 and 40 kDa and minor proteins of 15, 17, 19 and 67 kDa were detected and are arrowed.
**Figure 3.3** O-antigen hydrolysis assay. Silver stained SDS-15% PAGE showing LPS profiles of formalin-fixed *S. flexneri* PE577 and PE576 treated with whole SfII phage. Whole cell lysates of *S. flexneri*, treated with Proteinase K and subjected to electrophoresis.

Lanes contain:
- PE577 (serotype Y) untreated
- PE577 treated with SfII
- PE576 (serotype X) untreated
- PE576 treated with SfII

R-LPS - Lipid A-core alone

SR-LPS - Lipid A-core with a single repeat unit

S-LPS - Lipid A-core with long O-antigen chains
3.2.4.3 SfII lysogenisation of rfc mutant RMM109 of serotype Y

Bacteriophage SfII was used to infect rfc strain RMM109, turbid plaques were observed and single colonies isolated. Colony immunoblotting of these single colonies exhibited reactivity to type II antiserum, indicating that a single O-antigen repeat unit can be modified. It seems, therefore, that SfII phage does not require long O-antigen chains to recognise its receptor, and in fact a single O unit is sufficient. Conversely, SfII did not plaque on an rfbD strain (Van den Bosch et al., 1997) lacking O-antigen chains.

3.2.5 Characterisation of lysogens of bacteriophage SfII

Assessment of the host range of phage SfII revealed that strains of serotypes 1b, X and Y were susceptible to infection by phage SfII and could be serotype-converted (Table 3.1). Lysogenisation by SfII resulted in the expression of the type II antigen in serotypes 1b, X and Y as determined by colony immunoblotting. Further confirmation of the type II modification was achieved by bacterial agglutination using a specific type II antiserum (Table 3.2). The serotype converted lysogens were compared with a wildtype strain of serotype 2a, PE877 (2457T) which still agglutinates at a titre of 1:20 whereas the lysogens continue to agglutinate at a titre of 1:40, implying a significant and increased level of type II expression.

Strains of serotype Y are sensitive to lysogenisation by bacteriophage Sf6, which recognises receptors in the O-antigen and cleaves at particular positions along the polysaccharide chain. Strains of serotype 2a are not susceptible to infection by Sf6 phage and therefore lysogens of SfII would be expected to no longer be sensitive. However, when the SfII lysogens were assessed for their sensitivity to Sf6 phage by cross streaking, some residual sensitivity was observed.
**Table 3.1:** Host range of bacteriophage SfII. Bacteriophage SfII has the ability to convert strains of serotypes 1b, X and Y to express the type II antigen. The serotype (Roman numerals) and serogroup (Arabic numerals) of the strains are indicated in brackets for both the original and lysogenised organisms.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Type I</th>
<th>Type II</th>
<th>Type IV</th>
<th>Type V</th>
<th>group 6</th>
<th>group 3,4</th>
<th>group 7,8</th>
<th>Resultant Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE569 1a (I:4)</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE568 1b (I:6)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>2a (II: 3,4)</td>
</tr>
<tr>
<td>PE655 2a (II:3,4)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>unchanged</td>
</tr>
<tr>
<td>NCTC4 2b (II:7,8)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE571 3a (6, 7,8)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE645 3b (6, 3,4)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE566 4a (IV:3,4)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE572 4b (IV:6)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE780 5a (V:3,4)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE565 5b (V:7,8)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>unchanged</td>
</tr>
<tr>
<td>PE756 X (7,8)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>2b (II: 7,8)</td>
</tr>
<tr>
<td>PE577 Y (3,4)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>2a (II: 3,4)</td>
</tr>
</tbody>
</table>

*An unchanged resultant serotype implies that bacteriophage SfII is unable to recognise the O-antigen as a receptor in strains of that serotype.*
Table 3.2  Bacterial agglutination assay using anti-type II serum.

Serial dilutions of antiserum were incubated with $5 \times 10^9$ bacteria/ml of control strain PE877 (serotype 2a), parent strain PE577 (serotype Y) and independent SfII lysogens of PE577; RMM181, RMM182, RMM189, RMM268 and RMM273. The final titre of antiserum at which the bacteria still agglutinated is shown.

<table>
<thead>
<tr>
<th>Strains</th>
<th>Titre</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE877</td>
<td>1/20</td>
<td>2a</td>
</tr>
<tr>
<td>PE577</td>
<td>-</td>
<td>Y</td>
</tr>
<tr>
<td>RMM181</td>
<td>1/40</td>
<td>2a</td>
</tr>
<tr>
<td>RMM182</td>
<td>1/40</td>
<td>2a</td>
</tr>
<tr>
<td>RMM189</td>
<td>1/40</td>
<td>2a</td>
</tr>
<tr>
<td>RMM268</td>
<td>1/40</td>
<td>2a</td>
</tr>
<tr>
<td>RMM273</td>
<td>1/40</td>
<td>2a</td>
</tr>
</tbody>
</table>

- indicates negative agglutination
The LPS of strain PE577 (serotype Y) and SfII lysogens (RMM181, RMM182, RMM189, RMM268 and RMM273) were silver stained (Fig. 3.4, panel A). Compared with the PE577 LPS profile, the lysogens exhibit LPS bands with a slightly slower migration presumably due to the addition of the glucose residue which alters the LPS profile. The Western immunoblot (Fig 3.4, panel B) using anti-type II serum confirms that the lysogens have been converted to express the type II antigen when compared to the parent strain.

### 3.2.6 Isolation and cloning of PstI fragments of SfII genome

Bacteriophage SfII DNA was isolated and digested with the restriction enzymes BamHI, BglII, EcoRI and PstI (Fig. 3.5) to determine a restriction map. The genome of SfII was found to be 42.3 kb in size and the organisation of its PstI fragments is shown in figure 3.6. Each PstI fragment was isolated and cloned into pBluescript-SK or pBC-KS (Table 2.3) (Table 3.3), and individual restriction maps determined. Subsequently, each PstI clone was DIG-labelled and used in Southern hybridisations with BamHI, EcoRI and PstI digested SfII DNA to assign adjacent fragments. In this way, the genome was mapped and the order of fragments determined.

Difficulties were encountered in the cloning of PstI fragment 3, and several approaches were used. The 7.2 kb fragment was purified, followed by restriction with enzyme BamHI in an attempt to clone two separate PstI-BamHI fragments of sizes 2.4 and 4.8 kb, however, a 7.2 kb fragment flanked by BamHI ends was obtained (Fig. 3.7). Sequencing from both ends of this clone revealed homology to the integrase gene of bacteriophage P22, and subsequent analysis revealed a rearrangement had taken place during ligation.

Each PstI clone was transformed into PE577 and assessed for expression of type II antigen; it was found that none of these clones possessed the ability to mediate the serotype conversion from Y to 2a (Table 3.3).
Figure 3.4  Silver stained SDS-15%PAGE and Western immunoblot of lysogens of bacteriophage SfII.

Panel A: Silver stained gel of parent and SfII lysogens.
Lanes contain PE577, parent strain of serotype Y and SfII lysogens KMM181, RMM182, RMM189, RMM268 and RMM273. The addition of a glucose residue to the repeat unit alters the migration pattern of the individual LPS molecules. The different migration rates of the O-units of the Y serotype (closed arrow) and the 2a lysogens (open arrow) are particularly apparent at the level of one and two repeat units.

Panel B: Western immunoblot using anti-type II serum. The LPS of parent strain PE577 of serotype Y did not react with anti-type II serum, whereas the LPS of the SfII lysogens reacted with anti-type II serum.
Figure 3.5  Restriction digests of DNA from bacteriophage SfII. The genome of SfII was digested with the restriction enzymes BamHI, BglII, EcoRI and PstI. The marker track (SPP-1) has the sizes indicated in kilobase pairs (kb).
Figure 3.6  Bacteriophage SfII genome map

The SfII phage genome was restricted with various enzymes (Fig. 3.5) which were used to generate a restriction map. The organisation of the PstI fragments of the SfII genome is shown.
Figure 3.7 Isolation of fragment 3 of PstI digested SfII DNA.

A: Fragment 3 was initially obtained flanked by BamHI restriction sites at the ends. Dye primer sequencing using M13 forward and reverse primers revealed homology to the P22 integrase gene.

B: The correct orientation of fragment 3 and the direction of the integrase gene.
Re-arranged PstI-3 fragment

Correct arrangement of PstI-3
Table 3.3  *PstI* fragments of bacteriophage SfII

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Fragment</th>
<th>Size (kb)</th>
<th>Vector&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Serotype Converting&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRMM219</td>
<td>1</td>
<td>11</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM220</td>
<td>2</td>
<td>8.4</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM228</td>
<td>3</td>
<td>7.2</td>
<td>pBC-KS</td>
<td>-</td>
</tr>
<tr>
<td>pRMM221</td>
<td>4</td>
<td>5.5</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM222</td>
<td>5</td>
<td>4.6</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM224</td>
<td>6</td>
<td>3.6</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM225</td>
<td>7</td>
<td>1.55</td>
<td>pBS-SK</td>
<td>-</td>
</tr>
<tr>
<td>pRMM269</td>
<td>8</td>
<td>0.48</td>
<td>pBC-KS</td>
<td>-</td>
</tr>
</tbody>
</table>

<sup>a</sup> corresponds to the vector that the particular *PstI* fragment was cloned into
<sup>b</sup> corresponds to ability of clone to mediate serotype conversion of a Y serotype to 2 as determined by colony immunoblotting.
- indicates no serotype converting ability
3.2.7 Characterisation of PsfI fragments of SfII genome

DNA sequence from the ends of each of the 8 SfII PsfI fragments was obtained by dye primer sequencing of the subclones using M13 forward and reverse primers. Analysis of the sequence(s) obtained by comparison with BLASTN and BLASTX databases identified similarities with other genes and proteins (Table 3.4).

Significant homology to different sections of the *E. coli* K-12 chromosome was seen with fragments PsfI-3, PsfI-6 and PsfI-7, and some homology to known phage sequences was observed with PsfI fragments 4 and 8. Table 3.4 and Figure 3.8 detail the homology observed between the sequences obtained from the PsfI fragments (nt.) and the proteins identified by BLAST searches (aa).

*E. coli* K-12 ORF o189 (ECAE000214) was identified which spans fragments 3 and 7 ending in the latter which also encodes ORFs o112 (ECAE000214) and o445 (ECAE000214). Also encoded in fragment 7 is a homologue of the *E. coli* K-12 replication protein O (f122) (ECAE000324), and ORF f164 (ECAE000324) of unknown function which continues into fragment 6. A homologue of phage H-19B gene Q (AFO34975), involved in transcription of late genes via antiterminator activity (Mahdi *et al.*, 1996), is shared by fragments 4 and 6. PsfI fragment 8 encodes a holin homologue involved in lysis functions of bacteriophage phi-105 (L35561).

Figure 3.8 shows the location of sequence homology in relation to the PsfI fragments of the SfIII genome. The preliminary mapping of the replication protein O, gene Q and the holin lysis gene and their position in the genome is consistent with the organisation of phages such as λ, P22 and Mu.

Various subclones of the rearranged PsfI-3 pRMM228, pRMM253, pRMM254, pRMM255, pRMM256 and pRMM257, were isolated (Fig. 3.9) and the ends sequenced using dye primer sequencing. A P22 *int* homologue is encoded in the region flanking the *BamHI*
Table 3.4 Homologies exhibited by *Pst*I fragments of bacteriophage SfII

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Homology searcha</th>
<th>Nucleotide position in sequenceb</th>
<th>Amino acid position in search sequencec</th>
<th>% homologyd</th>
<th>Functione</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pst</em>I-6</td>
<td>F: f164 (ECAE000324)</td>
<td>nt. 1 - 264</td>
<td>f164 - aa 37-124</td>
<td>98% - 85/87</td>
<td>unknown</td>
</tr>
<tr>
<td>pRMM224</td>
<td>R: f362 (ECAE000253)</td>
<td>nt. 70-237</td>
<td>f362 - aa 307-362</td>
<td>91% - 50/55</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>H:19B (AF034975)</td>
<td>nt. 2- 49</td>
<td>H-19B - aa 1 - 16</td>
<td>75% - 12/16</td>
<td>Q protein - late gene synthesise</td>
</tr>
<tr>
<td><em>Pst</em>I-7</td>
<td>F: o112 (ECAE000214)</td>
<td>nt. 117-437</td>
<td>o112 - aa 1-106</td>
<td>70% - 74/106</td>
<td>unknown</td>
</tr>
<tr>
<td>pRMM225</td>
<td>o189 (ECAE000214)</td>
<td>nt. 1-117</td>
<td>o189 - aa 149-189</td>
<td>85% - 34/40</td>
<td>unknown</td>
</tr>
<tr>
<td></td>
<td>o455 (ECAE000214)</td>
<td>nt. 470-520</td>
<td>o455 - aa 1-17</td>
<td>76% - 13/17</td>
<td>unknown</td>
</tr>
<tr>
<td><em>Pst</em>I-3</td>
<td>R: f122 (ECAE000324)</td>
<td>nt. 114-398</td>
<td>f122 - aa 28-122</td>
<td>94% - 88/94</td>
<td>replication protein O</td>
</tr>
<tr>
<td></td>
<td>f164 (ECAE000324)</td>
<td>nt. 1-114</td>
<td>f164 - aa 1-38</td>
<td>97% - 37/38</td>
<td>unknown</td>
</tr>
<tr>
<td>*pRMM228</td>
<td>R: f362 (ECAE000323)</td>
<td>nt. 1-303</td>
<td>o306 - aa 1-303</td>
<td>86% - 261/303</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>dpm1 (D90915)</td>
<td>nt. 35-240</td>
<td>dpm1 - aa 54-259</td>
<td>44% - 91/206</td>
<td>dolichol-P mannosyl synthase</td>
</tr>
<tr>
<td>*pRMM228</td>
<td>R: o189 (ECAE000214)</td>
<td>nt. 1 - 431</td>
<td>o189 - 11-150</td>
<td>94% - 114/121</td>
<td>unknown</td>
</tr>
<tr>
<td><em>Pst</em>I-2</td>
<td>F: <em>S. cholerasuis</em></td>
<td>nt. 52- 83</td>
<td>transferase - aa 60-155</td>
<td>37% - 12/32</td>
<td>galactosyl transferase</td>
</tr>
<tr>
<td>pRMM220</td>
<td>galactosyl transferase</td>
<td>(S22623)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R: N.S.H.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3.4 cont.

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Homology search&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Nucleotide position in sequence&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Amino acid position in retrieved sequence&lt;sup&gt;c&lt;/sup&gt;</th>
<th>% homology&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Function&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>PstI -1</td>
<td>F:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRMM219</td>
<td>R:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PstI -5</td>
<td>F:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRMM222</td>
<td>R:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PstI -8</td>
<td>F: phi-105 holin (L35561)</td>
<td>nt. 40-279</td>
<td>holin - aa 6-82</td>
<td>42% - 32/76</td>
<td>lysis function</td>
</tr>
<tr>
<td>pRMM269</td>
<td>R:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pRMM221</td>
<td>R:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Homology searches were all performed in BLASTX
<sup>b</sup> nucleotides correspond to PstI fragment sequenced
<sup>c</sup> amino acids position correspond to amino acid number in retrieved sequence
<sup>d</sup> Percentage homology over the amino acids in c
<sup>e</sup> function of gene retrieved by BLASTX search
* Pst-3 fragment is flanked by BamHI ends
N.S.H. - no significant homology
R- Dye primer sequencing using M13 reverse primer
F- Dye primer sequencing using M13 forward primer
Figure 3.8 Map of fragments generated by PstI digestion of bacteriophage SfiI genome.

The alignment of the 8 PstI fragments was determined by cloning each of these fragments into vectors (pBLUESCRIPT or pBC-KS) in the orientation as indicated (f represents the M13 forward primer and r represents the M13 reverse primer). The fragments are aligned with respect to each other and homologies (DNASIS) exhibited by sequence obtained by partial dye primer sequencing to known sequences is indicated in the diagram.
Figure 3.9 Subclones of rearranged PstI fragment 3.

Subclones pRMM253, pRMM254, pRMM255, pRMM256 and pRMM257 of the rearranged fragment 3 pRMM228 were partially sequenced using dye primer sequencing. Region of homology with P22 integrase, orf10X1 of *S. flexneri* and Dpm1 of *S. cerevisiae* are indicated.
site. A homologue to the *S. flexneri* ORF10x1 (Morona *et al*., 1994) was found near the *EcoRI* site. Homology to the dolichol phosphate mannosyl synthase gene (*dpm1*) of *Saccharomyces cerevisiae* (Orlean *et al*., 1988) was also seen when sequencing was carried out from the 3' end of pRMM257 and weak homology to a *Salmonella* galactosyl transferase gene was identified in the *PstI-2* fragment. The region of DNA encompassing *PstI-2* and *PstI-3* was subsequently isolated.

### 3.2.8 Identification of a region of bacteriophage SfII with serotype-converting ability

A 4 kb *BamHI* fragment overlapping the *PstI-2* and *PstI-3* junction was isolated and cloned into pBC-KS resulting in pRMM264 (Fig. 3.10). Upon transformation of pRMM264 into PE577 conversion from Y to 2a was seen by colony immunoblotting using type II antiserum. Two derivatives of pRMM264, pRMM266 and pRMM268 (Fig. 3.10), which are *EcoRI* and *PstI* deletions respectively, were unable to mediate conversion to serotype 2 when introduced into PE577 as determined by slide agglutination and colony immunoblotting.

### 3.3 Summary

This chapter described the isolation of the serotype converting bacteriophage SfII which is responsible for the addition of a glucose residue to the RhaIII of the tetrasaccharide repeat unit of the LPS of *S. flexneri*, resulting in expression of the type II antigen. Morphologically bacteriophage SfII belongs to the group A of the Bradley classification and does not exhibit O-antigen hydrolytic activity. Characterisation of the genome of SfII phage revealed homology to regions of the *E. coli* K-12 chromosome and bacteriophage P22. A 4 kb region of DNA in pRMM264 which was able to mediate serotype conversion to type II.
Figure 3.10  Map of genome of SfII phage and serotype-converting plasmid pRMM264.

pRMM264 is a 4kb BamHI fragment encompassing the junction of PstI-2 and PstI-3, which contains the gene(s) required for conversion to type II. Plasmids pRMM266 and pRMM268 are subclones of pRMM264 and do not possess serotype converting ability.
Chapter Four

Sequence of pRMM264 and identification of the attachment region of bacteriophage SfII

4.1 Introduction

Temperate bacteriophages possess several features which allow them to integrate and excise from their hosts chromosome: an attachment site (attP), an integrase gene (int) and an excisionase gene (xis). Similarly, the bacterial host must possess the corresponding attachment site (attB) and must encode an integration host factor (IHF) (Campbell, 1992).

Integration of the phage genome into the bacterial chromosome occurs by site-specific recombination at the attachment sites and is mediated by the function of the integrase gene product. Excision of a temperate phage requires the action of both the integrase (int) and excisionase (xis) which are often located adjacent to the attP site.

This chapter describes the DNA sequencing of the serotype converting plasmid pRMM264. In addition, plasmid pRMM273 was isolated which encompasses the region upstream of the 5' BamHI site in pRMM264. Five open reading frames were identified: xis, int, ORF2, ORF3 and ORF4. The attachment site, attP, was also identified and localised to PstI fragment 3.
4.2 Results

4.2.1 DNA sequencing of the *BamH*I-*BamH*I insert in pRMM264

DNA sequencing of the insert in pRMM264 (Section 3.2.7) was carried out using three clones derived from *BamH*I, *PstI* and *EcoRI* subclones of the 4 kb *BamH*I fragment: pRMM267, pRMM262 and pRMM255 (Fig. 4.1). Dye primer sequencing was carried out on each of these plasmids using M13 forward and reverse primers. Oligonucleotide primers (Fig. 4.1, Table 2.4) were designed from the sequence obtained and used in Dye Terminator reactions to sequence the remainder of the fragment.

4.2.2 Open reading frames in 4 kb *BamH*I fragment of pRMM264

Three complete ORFs were found in the 4 kb *BamH*I fragment of pRMM264, with a further ORF located at the 5’ end which was found to be partial. By comparison with databases, this ORF was found to encode an integrase homologue (Section 4.2.3). In order to clone the complete *int* gene, plasmid pRMM273 (*BamH*I-*Bgl*II) was isolated. pRMM273 lies entirely within fragment *PstI*-3 and contains the fragment adjacent to pRMM264, sharing the *BamH*I site and containing the 5’ end of the partial ORF. Sequencing of this 1.4 kb fragment was carried out using oligonucleotide primers (Fig. 4.2, Table 2.4) and resulted in the identification of a fifth ORF (*xis*).

The combined 5.2 kb sequence of pRMM264 and pRMM273 is shown in Figure 4.3. The GenBank Accession number is AF021347.
Figure 4.1 Sequencing strategy for pRMM264.

Three subclones of pRMM264 were constructed:

pRMM255 (1.3 kb *BamHI*-*EcoRI*), pRMM262 (1.1 kb *EcoRI*-PstI) and pRMM267 (1.6 kb PstI-*BamHI*). M13 dye primer sequencing was used on pRMM264 and subclones and based on the sequence obtained, oligonucleotide primers (numbered) were designed to facilitate sequencing of both strands of the entire fragment.
Figure 4.2 Sequencing strategy of ORF1-integrase region

A 1.4 kb *BamHI-BgII* fragment encompassing the 5' end of the integrase gene (*int*) was isolated and cloned into pBC-KS. Oligonucleotide primers (numbered) were designed to sequence both strands.
Figure 4.3 Combined sequence of inserts in pRMM264 and pRMM273

The sequence data generated in the manner shown in Figures 4.1 and 4.2 were linked together via the BamHI site that is shared by the inserts in the two plasmids. The open reading frames and the translation in one letter amino acid code are shown. Start codons and restriction sites are indicated in bold and stop codons are underlined. The name of the ORF is shown before the start codon.
4.2.3 Open reading frames in 5.2 kb sequence

Five open reading frames were found in the 5.2 kb sequence which were all transcribed in the same direction (Fig. 4.4). Potential ribosome binding sites have been identified upstream of the start codons of the five ORFs (Table 4.1), however no sequences resembling E. coli consensus promoter sequences for the "Priibnow box" (TTGACA) and the "-35 region" (TATAAT) were observed. The start and stop codons of int, xis and the latter three ORFs; ORF2, ORF3 and ORF4 overlap, resulting in a shift in the reading frame for the next gene; this arrangement suggests that translation of these genes may be coupled.

4.2.4 Integrase (int) and Excisionase (xis)

Two open reading frames were identified between nt. 524-958 and nt. 835-1998 and were found to be transcribed in the same direction. The first of these exhibited a 87.8% (nucleotide level) and 87.9% (amino acid level) and homology to the excisionase gene (xis) of bacteriophage P22 and 68.8% homology to the region upstream of the integrase gene in the defective phage DLP12 which is located at 12 min on the E. coli K-12 chromosome. The second ORF also exhibited a high degree of homology to the integrase gene of phage P22 (X04052) and the defective phage DLP12 (M31074) (Table 4.2). The organisation of the two ORFs (xis and int) in bacteriophage SfII is identical to phage P22 (Poteete, 1988).

4.2.5 The attachment site, attP

Lysogens of SfII were further characterised by Southern hybridisation as shown in Figure 4.5. Chromosomal DNA of lysogens RMM181 and RMM273, the parent strain
Figure 4.4 Open reading frames of 5.4 kb sequence.

Plasmid pRMM264, spans PstI fragments 2 and 3 of the SfII genome. Plasmid pRMM273 is located upstream of pRMM264 and contained within PstI fragment 3. Sequencing of the total 5.4 kb of DNA identified five ORFs transcribed in the same direction. The first two ORFs show homology to excisionase and integrase genes of known phages and the latter three ORFs are discussed (Chapter 5).
**Table 4.1** Table of properties of ORFs encoded by 5.4 kb sequence

<table>
<thead>
<tr>
<th>Coding Region</th>
<th>M.W.</th>
<th>aa</th>
<th>pI</th>
<th>Hydropathy Index</th>
<th>G+C%</th>
<th>rbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>xis</td>
<td>nt 524-958</td>
<td>16,249</td>
<td>144</td>
<td>9.86</td>
<td>-0.64</td>
<td>47.6</td>
</tr>
<tr>
<td>int</td>
<td>nt 835-1998</td>
<td>44,942</td>
<td>387</td>
<td>10.23</td>
<td>-0.71</td>
<td>46.9</td>
</tr>
<tr>
<td>ORF2</td>
<td>nt 2260-2622</td>
<td>13,298</td>
<td>120</td>
<td>9.86</td>
<td>1.03</td>
<td>43.0</td>
</tr>
<tr>
<td>ORF3</td>
<td>nt 2619-3548</td>
<td>34,844</td>
<td>309</td>
<td>7.68</td>
<td>0.09</td>
<td>42.4</td>
</tr>
<tr>
<td>ORF4</td>
<td>nt 3545-5005</td>
<td>55,775</td>
<td>486</td>
<td>9.11</td>
<td>0.54</td>
<td>31.8</td>
</tr>
</tbody>
</table>

- M.W. = molecular weight
- aa = number of amino acids the ORF
- pI = determined using PROSIS
- Hydropathy Index = determined according to Kyte and Doolittle as implemented in PROSIS
- rbs = potential ribosome binding site identified at nucleotide position indicated
Table 4.2 Similarity of \textit{int} to other ORFs

<table>
<thead>
<tr>
<th></th>
<th>\textit{int}</th>
<th>P22\textit{int}</th>
<th>DLP12\textit{int}</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{int}</td>
<td>100</td>
<td>79.9 (N)</td>
<td>68.7 (N)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>90.2 (A)</td>
<td>72.7 (A)</td>
</tr>
<tr>
<td>P22\textit{int}</td>
<td>100</td>
<td></td>
<td>68.3 (N)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>71.1 (A)</td>
</tr>
<tr>
<td>DLP12\textit{int}</td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

The values shown represent both homology at the nucleotide level (N) and the amino acid level (A) over the entire length of the coding sequence. P22\textit{int} corresponds to the integrase gene of \textit{Salmonella} bacteriophage P22 (X04052), DLP12\textit{int}; corresponds to the integrase of a defective phage in the \textit{E. coli} K-12 chromosome at map position 12 min (M31074)
Figure 4.5  Southern hybridisation of lysogens of SfII.

Chromosomal DNA of two SfII S. flexneri lysogens (RMM181 and RMM273) were digested with PstI, transferred to nylon membrane and probed with DIG-labelled PstI digested SfII DNA. The arrows indicate the sizes of the PstI fragments, corresponding to PstI digested phage SfII DNA.

The marker track EcoRI digested SPP-1 phage DNA is detected using DIG-labelled SPP-1 (sizes are 8.5, 7.3, 6.1, 4.8, 3.6, 2.8, 1.95, 1.86, 1.51, 1.3, 1.16, 0.98 and 0.72 kb) and is also shown in this figure.
PE577, and phage SfII were digested with \textit{Pst}I and then probed with DIG-labelled SfII DNA. The genome of phage SfII is restricted into 8 fragments by the enzyme \textit{Pst}I, all of which can be clearly seen in the lane labelled SfII. The parent strain, PE577 shows weak homology to SfII, whereas, RMM181 and RMM273, two SfII lysogens, show seven of the eight \textit{Pst}I fragments, with \textit{Pst}I-3 missing in both. \textit{Pst}I-3 may contain the attachment region (\textit{attP}) of phage SfII, which is required for the integration of the phage genome into the host chromosome and is subsequently split upon integration (see Fig. 4.6). Inspection of the sequence (Fig. 4.3) identified the \textit{attP} site which is located on \textit{Pst}I fragment 3 between \textit{int} and ORF2 at nt 2043-2064 (5'CTTCGCATTACGAATTATAAG3') (Fig. 4.3) identical to that of phage P22 (Leong \textit{et al.}, 1985).

A region of DNA encompassing the 3' end of \textit{int}, \textit{attP} and the 5' end of ORF2 of pRMM264 was DIG-labelled using PCR with primers \#2326 and \#2330 (Table 2.4) and used to probe \textit{Pst}I digested chromosomal DNA of several independent lysogens (Section 3.2.5). A single fragment of 8.5 kb in size was seen in the parent track, PE577, (Fig. 4.7A) whereas in the tracks containing the lysogens (RMM181, RMM182, RMM189 and RMM273), two fragments of 9.2 and 6.5 kb were detected which correspond to the hybrid \textit{attPlattB} sites formed as a result of integration (Fig. 4.6).

Bacteriophage P22 integrates into the \textit{Salmonella typhimurium} chromosome adjacent to \textit{proAB} at 7 min. The same sequence is found in the \textit{E. coli} K-12 (ECAEO00132) genome where the tRNA gene \textit{thrW} (nt. 8811-8886) is encoded adjacent to \textit{proA,B} and contains the \textit{attB} site (nt. 8841-8855, 5'ATTCGTAATGCGAAG3'). This region was amplified by PCR (primers \#2370, 2371) and DIG-labelled and used to probe \textit{Pst}I digested chromosomal DNA of the lysogens. Figure 4.7B shows that \textit{thrW} hybridises to a fragment of approximately 9.2 kb in the parent strain PE577, whereas in the SfII lysogenic strains the fragment containing \textit{thrW} (or \textit{attB}) is located on a larger fragment of 9.2 kb in size. A much fainter band is detected at approximately 6.5 kb; the reason this fragment stains weakly is not known,
Figure 4.6  Integration of phage genome into host chromosome via attachment sites; $attP$ and $attB$.

The phage attachment site ($attP$) is contained within $PstI$ fragment 3 at the location indicated in Figure 4.2 and the distances from the $PstI$ sites are shown. The host attachment site ($attB$) lies within the tRNA gene $thrW$ where recombination between the two sequences occurs resulting in a hybrid site.
Bacterial chromosome

Lysogen

SfII

PstI

5436 bp

attP

1764 bp

PstI

9.2 kb

6.5 kb

PstI

PstI

B

P'

attB

PstI

SfII

PstI
Figure 4.7 Southern hybridisation to identify attachment site, *attP*.

Southern hybridisations using chromosomal DNA from:

lane 1: PE577, parent strain

lane 2: RMM181, SfII lysogen

lane 3: RMM182, SfII lysogen

lane 4: RMM189, SfII lysogen

lane 5: RMM273, SfII-lysogen

using **A**: DIG- labelled *attP* and

**B**: DIG- labelled *thrW* as probes.

SPP-1 markers are shown and the sizes of the fragments are labelled.
however, the two fragments correspond to the sizes seen when probing the same lysogens with the \textit{attP} fragment, as would be expected (Fig. 4.7A). These data are consistent with the original mapping of the attachment site (\textit{attB}) in the \textit{Shigella} and \textit{Salmonella} chromosomes (Petrovskaya and Licheva, 1982; Campbell, 1992).

### 4.3 Summary

The sequencing of the insert in the serotype converting plasmid pRMM264 identified 3 complete ORFs and one partial ORF corresponding to the integrase gene. Isolation and sequencing of a fragment (in pRMM273) adjacent to pRMM264 resulted in the complete sequencing of the integrase gene and in addition the excisionase gene of bacteriophage SfIII. This chapter also demonstrates the splitting of attachment sites \textit{attP} and \textit{attB} upon integration of the phage genome. Bacteriophages SfIII and P22 are morphologically different, however, they have identical integration and attachment systems suggesting that they were derived from a common ancestor.
Chapter Five

Localisation of the serotype converting genes of bacteriophage SfII

5.1 Introduction

The isolation and characterisation of bacteriophage SfII and lysogenic strains has been previously described. Lysogenisation of *S. flexneri* strains of serotype X, Y and 1b by phage SfII resulted in the expression of type II antigen and Southern hybridisation revealed the presence of the integrated phage genome in the host chromosome. The genome of phage SfII was isolated to identify the gene(s) responsible for serotype conversion.

The 42.3 kb phage chromosome was restricted with *PstI* into 8 fragments, which were cloned and introduced into a *S. flexneri* strain of serotype Y; none of which were able to mediate serotype conversion. Preliminary sequencing of these clones led to interest in the region spanning *PstI* fragments 2 and 3. A 4 kb *BamHI* fragment was isolated and cloned (pRMM264) which spanned the shared *PstI* site of these two fragments. When pRMM264 was introduced into *S. flexneri* Y strains, expression of type II antigen was detected. Sequence analysis showed that pRMM264 encoded 3 complete open reading frames and the 3' end of the integrase gene.

This chapter describes the three complete ORFs; ORF2, ORF3 and ORF4 and the homology exhibited by these ORFs to other genes and proteins. A requirement for two
proteins encoded by ORF3 and ORF4, to mediate efficient serotype conversion is also demonstrated.

5.2 Results

5.2.1 ORF2

The second ORF located at nt 2260-2622 (Fig. 4.3) encodes a highly hydrophobic protein with a predicted molecular weight of 13,298 Da (Table 4.1). BLASTP searches revealed a high degree of homology to ORFs in *E. coli*, *S. flexneri* and bacteriophage P22 (Table 5.1, Fig. 5.1). All four predicted proteins have four potential transmembrane domains (Fig. 5.2). The function of these proteins is as yet unknown, however, ORF2 does not appear to play a role in the serotype converting ability of SfII phage.

5.2.2 ORF3 bactoprenol glucosyl transferase (*bgt*)

ORF3 lies between nt. 2619-3548 (Fig. 4.3) and encodes a protein of a predicted weight of 34,844 daltons. The amino terminal two thirds of ORF3 were found to be hydrophilic and two hydrophobic regions were detected near its carboxy terminal end; the latter are potential transmembrane domains (Fig. 5.3). ORF3 was found to be homologous to o306 and o322 of *E. coli* K-12, o318 and o331 of *Synechocystis*, csbB of *Bacillus subtilis* and o342 of *Mycobacterium tuberculosis* (Table 5.2). The function of these proteins is unknown, but alignment of these ORFs (Fig 5.4) allowed identification of the amino terminal motif sequences, DXSXD and DXD, which are present in Dpm1 of *Saccharomyces cerevisiae*. These motifs may represent the active sites of these proteins. Dpm1 is a well characterised
Table 5.1  Similarity of ORF2 to other ORFs

<table>
<thead>
<tr>
<th></th>
<th>ORF2</th>
<th>AE323o120</th>
<th>P22o120</th>
<th>ORF10x1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF2</td>
<td>100</td>
<td>88.4 (N)</td>
<td>71.3 (N)</td>
<td>52.6 (N)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>87.5 (A)</td>
<td>77.5 (A)</td>
<td>46.0 (A)</td>
</tr>
<tr>
<td>AE323o120</td>
<td>100</td>
<td>71.9 (N)</td>
<td>79.2 (A)</td>
<td>51.5 (N)</td>
</tr>
<tr>
<td>P22o120</td>
<td></td>
<td></td>
<td>100</td>
<td>52.4 (N)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42.5 (A)</td>
</tr>
<tr>
<td>ORF10x1</td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

The values shown represent both homology at the nucleotide level (N) and amino acid level (A) over the entire length of the coding sequence. AE323o120 corresponds a region of *E. coli* K-12 (ECAE000323); P22o120 corresponds to an ORF of 120 aa in the P22 *attP* region (M10893) and ORF10x1 is from *S. flexneri* (Morona et al., 1994).
Figure 5.1 Alignment of ORF2 with homologous proteins.

The amino acid sequences of the following proteins were aligned using the program CLUSTAL using the default settings. The proteins used are shown as: SFIIorf2, SfII phage ORF2; AE3230120 (ECAE000323), E. coli K-12; P220120 (M10893), P22 phage and ORF10X1 (X71970), S. flexneri. Amino acid positions identical to those in SFIIorf2 are shaded; *, identical amino acids at this position in all the ORFs; †, similar amino acids at this position in all the ORFs.
<table>
<thead>
<tr>
<th></th>
<th>SFIorf2</th>
<th>AE323o120</th>
<th>P22o120</th>
<th>SFRFBI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLKL--------</td>
<td>FKYTSIGVLNTLHWWVFVCICYAHTSQALANFTGFVVAVSFSPAN 53</td>
<td>FKYTSIGVLNTLHWWVFVCICYVAHTNQALANFAGFVVAVSFSPAN 53</td>
<td>FAKYTSIGVLNTLHWWVFVCICYVAHTNQALANFSGFVIAVSFSPYAN 53</td>
</tr>
<tr>
<td></td>
<td>MLKL--------</td>
<td>FKYTSIGVLNTLHWWVFVCICYVAHTNQALANFAGFVVAVSFSPAN 53</td>
<td>FKYTSIGVLNTLHWWGFACCVYGMHHTQALANFSGFVIAVSFSPYAN 53</td>
<td>FAKYTSIGVLNTLHWWVFVCICYVAHTNQALANFSGFVIAVSFSPYAN 53</td>
</tr>
<tr>
<td></td>
<td>MLKGKLLTSSFSYFLIGIVNTALHWGVFYACYNNLAFGQGRSNIVGFIACAATFSSPAN 60</td>
<td>MLKGKLLTSSFSYFLIGIVNTALHWGVFYACYNNLAFGQGRSNIVGFIACAATFSSPAN 60</td>
<td>MLKGKLLTSSFSYFLIGIVNTALHWGVFYACYNNLAFGQGRSNIVGFIACAATFSSPAN 60</td>
<td>MLKGKLLTSSFSYFLIGIVNTALHWGVFYACYNNLAFGQGRSNIVGFIACAATFSSPAN 60</td>
</tr>
<tr>
<td></td>
<td>SFRFBI</td>
<td>SFRFBI</td>
<td>SFRFBI</td>
<td>SFRFBI</td>
</tr>
<tr>
<td></td>
<td>ARFTFKASTTAMRYMYVGFMGILSVIVGWAADKCSLPPIVT六个月FSISLVCFGVYSKF 113</td>
<td>ARFTFKASTTAMRYMYVGFMGILSVIVGWAADKCSLPPIVT六个月FSISLVCFGVYSKF 113</td>
<td>ARFTFNATTTLRYMYVGFMGTLASAVVGMADQCSLPPLIT六个月FSISLVCFGIFYSRP 120</td>
<td>ARCSFKVSATKARYFIFFGAMSYLFGVLFDLALSPIFTLFTSFSLPSLVGLYCASKY 120</td>
</tr>
<tr>
<td></td>
<td>SFIorf2</td>
<td>AE323o120</td>
<td>P22o120</td>
<td>SFRFBI</td>
</tr>
<tr>
<td></td>
<td>I V F R D A K</td>
<td>I V F R D A K</td>
<td>I V F R D I R</td>
<td>F I F R --</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.2  Comparison of hydrophobicity plots of ORF2 and homologues.

The hydrophobicity profiles of a number of predicted proteins were generated in PROSIS and compared by alignment using AUTOCAD. The Kyte and Doolittle values were used. The proteins used are shown as SFIOrf2, SfII phage ORF2; AE3230120 (ECAE000323), E. coli K-12; P220120 (M10893), P22 phage and ORF10X1 (X71970), S. flexneri. The x axis in each profile is marked at the position of every 30 amino acids. The y axis indicates the relative hydrophobicity index.
Figure 5.3  Comparison of hydrophobicity plots of ORF3 and homologues.

The hydrophobicity profiles of a number of predicted proteins are compared by the use of the program PROFILEGRAPH (Hofmann and Stoffel, 1992). The Gunther von Heijne (von Heijne, 1992) values were used. The proteins used are shown as SFIIorf3, SfII phage ORF3; 323o306 (ECAE000323), *E. coli* K-12; SYN318 and SYN0331 (D64006 and D90915), *Synechocystis* ORFs of 318 and 331 amino acids respectively; BSCSBB (Z79701), *Bacillus subtilis* csbB gene (D64006); MYTo342, *Mycobacterium tuberculosis* ORF of 342 amino acids; 315o322 (ECAE000315), *E. coli* K-12, DPMI, (J04184), *Saccharomyces cerevisiae* dpm1. The x axis in each profile is marked at the position of every 10 amino acids. The y axis indicates the relative hydrophobicity index.
Table 5.2. Similarity of ORF3 to other ORFs

<table>
<thead>
<tr>
<th></th>
<th>ORF3</th>
<th>323o306</th>
<th>SYNo318</th>
<th>SYNo331</th>
<th>BSCSBB</th>
<th>MYTo342</th>
<th>315o322</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF3</td>
<td>100</td>
<td>89.9 (306)</td>
<td>47.5 (301)</td>
<td>43.7 (302)</td>
<td>39.8 (304)</td>
<td>31.9 (304)</td>
<td>30.9 (311)</td>
</tr>
<tr>
<td>323o306</td>
<td>100</td>
<td>47.5 (304)</td>
<td>41.1 (299)</td>
<td>32.5 (301)</td>
<td>31.8 (305)</td>
<td>31.1 (309)</td>
<td></td>
</tr>
<tr>
<td>SYNo318</td>
<td>100</td>
<td>43.5 (308)</td>
<td>43.9 (301)</td>
<td>30.9 (307)</td>
<td>33.8 (305)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYNo331</td>
<td>100</td>
<td>38.0 (303)</td>
<td>29.0 (303)</td>
<td>31.8 (305)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSCSBB</td>
<td></td>
<td>100</td>
<td>28.2 (319)</td>
<td>29.9 (308)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYTo342</td>
<td></td>
<td>100</td>
<td></td>
<td>25.7 (307)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>315o322</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values shown in the table represent homology at amino acid level over the number of amino acids presented in brackets. ORF3; SfII phage ORF3; 323o306 (ECAE000323), E. coli K-12; SYNo318 (D64006) and SYNo331 (D90915), Synechocystis ORFs of 318 and 331 amino acids respectively; BSCSBB (D64006), Bacillus subtilis csbB gene; MYTo342 (Z79701), Mycobacterium tuberculosis ORF of 342 amino acids; 315o322 (ECAE000315), E. coli K-12.
Figure 5.4  Alignment of proteins homologous to ORF3.

The amino acid sequences of the following proteins were aligned using the program CLUSTAL (Higgins and Sharp, 1988) using the default settings. Some manual alignment was also used in the regions of gaps. The proteins used are shown as: SFIOrf3, SfII phage ORF3; 323o306 (ECAE000323), E. coli K-12; SYN318 and SYNo331 (D64006 and D90915), Synechocystis ORFs of 318 and 331 amino acids respectively; BSCSBB (Z79701), Bacillus subtilis csbB gene (D64006); MYTo342, Mycobacterium tuberculosis ORF of 342 amino acids; 315o322 (ECAE000315), E. coli K-12, DPMI, (J04184), Saccharomyces cerevisiae dpm1. Amino acids positions identical to those in SFIOrf3 are shaded; *, identical amino acids at this position in all the ORFs; ., similar amino acid at this position in all the ORFs. Motif sequences are indicated by bold type, and the two hydrophobic domains identified in SFIOrf3 are indicated by a line above the corresponding amino acids.
dolichol phosphate mannosyl synthase which catalyses the formation of dolichol-phosphate mannose from dolichol-phosphate and GDP-mannose (Orlean et al., 1988). The motifs identified are similar to those present in glycosyl transferases (Saxena et al., 1995, Keenleyside and Whitfield, 1995). ORF3 was named bactoprenol glucosyl transferase (bgt) based on the above similarities and on complementation data (Section 5.3).

Interestingly, a partially sequenced ORF (105bp) at the 5' end of the attP (X04052) region of P22 phage displays very good homology to the 3' end of ORF3 (79.8% DNA and 82.4% protein) over the entire length (Fig. 5.5). The organisation of the region surrounding the partial ORF suggests that it may have a similar function to the region identified in SfII. Sequencing of this region is not complete but it is highly likely to encode the genes required for expression of the Salmonella O-antigen factor 1 (Iwashita and Kanegasaki, 1973).

5.2.3 ORF4 - glucosyl transferase gtrII

ORF4 encodes a highly hydrophobic protein (Table 4.1) which shows no significant homology to protein sequences by BLASTX searches, however, weak similarity to RfbP, a galactosyl transferase of S. enterica, was detected over a small region of the protein. Alignment of ORF4 with gtrX and o443 of E. coli K-12 identified conserved regions between these proteins (Fig. 5.6), although overall homology between these proteins is only about 20% using FASTA as implemented in PROSIS (Table 5.3). Hydropathy plots suggest 10 to 11 potential transmembrane domains implying that ORF4 may be an integral membrane protein (Fig. 5.7). This ORF has been designated gtrII based on complementation tests and the phenotype conferred by this gene (Section 5.3).
Figure 5.5 Homology of ORF01 of P22 to ORF3 of phage SFII

A schematic representation of the attP region of phage P22 is shown. A partially sequenced ORF, ORF01 exhibits homology to ORF3 of bacteriophage SFII at the nucleotide and amino acid level. An alignment of the sequences is shown, and the region of the P22 genome (X04052) is indicated.
Figure 5.6  Alignment of ORF4 with functionally related proteins.

The amino acid sequences of the following proteins were aligned using the program CLUSTAL using the default settings. The proteins used are shown as: SFIIorf4, SfII phage ORF4; GTRX (L05001), SfX phage glucosyl transferase gene; AE3230443 (ECAE000323), E. coli K-12. Amino acid positions identical to those in SFIIorf4 are shaded; *, identical amino acids at this position in all the ORFs; ., similar amino acids at this position in all the ORFs.
### Table 5.3 Similarity of ORF4 to other ORFs

<table>
<thead>
<tr>
<th></th>
<th>ORF4</th>
<th>gtrx</th>
<th>o443</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF4</td>
<td>100</td>
<td>20.2</td>
<td>18.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(356)</td>
<td>(356)</td>
</tr>
<tr>
<td>gtrx</td>
<td></td>
<td>100</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(214)</td>
</tr>
<tr>
<td>o443</td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

Values shown in the table represent homology at amino acid level over the number of amino acids presented in brackets. ORF4; SfII phage ORF4; gtrx; SfX phage gtrx (L05001) and o443; *E. coli* K-12 (ECAE000323).

Values determined by FASTA as implemented in PROSIS.
Figure 5.7  Comparison of hydrophobicity plots of ORF4 and functional homologues.

The hydrophobicity profiles of a number of predicted proteins are compared by the use of the program PROFILEGRAPH (Hofmann and Stoffel, 1992). The Gunther von Heijne (von Heijne, 1992) values were used. The proteins used are shown as: SFIIorf4, SfiI phage ORF4; GTRX (L05001), SfX phage glucosyl transferase gene; AE3230443 (ECAE000323), *E. coli* K-12. The x axis in each profile is marked at the position of every 10 amino acids. The y axis indicates the relative hydrophobicity index.
5.3 Complementation analyses

5.3.1 Serotype conversion in *S. flexneri* with plasmid constructs

Table 5.4 details the various constructs used in *S. flexneri* strain PE577 of serotype Y and *E. coli* K-12 strain DH5α to observe their serotype converting ability. Reactivity to various antisera (group 3,4; type II; group 7,8) was assessed by colony immunoblotting.

Plasmids pRMM270, pRMM271 and pRMM272 (Fig. 5.8) were derived by amplification of *bgt* and *gtrII* by PCR using primers (#2349, #2376 (*bgt*), #2325, #2407 (*gtrII*) and #2325, #2349 (*bgt/gtrII*) respectively) (Table 2.4) and ligated with pGEM-T. Upon transformation into PE577, no serotype conversion was seen with either of these two genes alone. However, when both genes were on a single fragment, conversion to serotype 2a was observed (Table 5.4). ORF2 was not included in any of these constructs.

5.3.2 Serotype conversion with *gtrX* and *bgt*

The *gtrX* gene of bacteriophage SfX has been previously shown to mediate partial serotype conversion in the absence of other genes. For this reason *gtrX* was PCR amplified using primers (#747, #748) (Table 2.4) based on the published sequence, cloned into pUC18 and pK184 resulting in plasmids pRMM174 and pRMM180, respectively. When introduced into PE577, neither plasmid resulted in serotype conversion to serotype X. However, when pRMM180 was introduced into PE577 already containing *bgt* on plasmid pRMM270, conversion to serotype X occurred (Table 5.4). This is consistent with the data observed for *gtrII*. These data imply that ORF2 does not appear to play a role in serotype conversion.
Table 5.4 Complementation in *S. flexneri* and *E. coli* strains using various plasmids.

<table>
<thead>
<tr>
<th>Strain/ plasmid</th>
<th>Reactivity to antisera in colony immunoblot</th>
<th>Serotype</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>anti-3,4</td>
<td>anti-II</td>
<td>anti-7,8</td>
</tr>
<tr>
<td>PE577</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM264</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM266</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM268</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM270</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM271</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM180</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PE577 + pRMM268 + pRMM180</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>PE577 + pRMM270 + pRMM180</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DH5α + pRMA154</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DH5α + pRMA154 + pRMM174</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>DH5α + pRMA154 + pRMM271</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

+ indicates positive reaction to antiserum in colony immunoblot
- indicates no reactivity to antiserum by colony immunoblot
**Figure 5.8** Plasmids pRMM270, pRMM271 and pRMM272

Plasmid pRMM264, and the subclones pRMM266 and pRMM268, is shown in this diagram. In addition, plasmids pRMM270, pRMM271 and pRMM272 which contain PCR amplified copies of \( bgt \), \( gtrII \) and both \( bgt \) and \( gtrII \), respectively are shown.
The function of \textit{bgt} was therefore necessary in the modification of the LPS, whereas the serotype specificity was determined by the particular glucosyl transferase gene (\textit{gtrII} or \textit{gtrx}).

5.3.3 Serotype conversion using \textit{E. coli} K-12 DH5\(\alpha\) ORF3 homologue, o306

Homology searches identified a \textit{bgt} homologue, called o306, located at 54 minutes (ECAE000323) in \textit{E. coli} K-12. To investigate if this homologue was functional, the \textit{S. flexneri} \textit{rfb} genes on pRMA154 (Macpherson \textit{et al.}, 1991) were introduced into an \textit{E. coli} K-12 strain DH5\(\alpha\). It was found that this strain reacted in a colony immunoblot to type IV antigen, as previously described (Morona \textit{et al.}, 1995). This implies that o306 of \textit{E. coli} K-12 is capable of adding a glucosyl residue to the GlcNAc sugar of \textit{S. flexneri} O-antigen. Subsequent addition of \textit{gtrx} (pRMM174) resulted in the expression of the group 7,8 antigen, suggesting that o306 was able to perform the same function as \textit{bgt}. Type II antigen, however, was not detected when \textit{gtrII} was introduced into the same strain. Inability of \textit{gtrII} to glucosylate the LPS was possibly due to the position of the existing modification, as glucosylation of both the GlcNAc and adjacent RhaIII sugar is not found in any of the naturally occurring \textit{S. flexneri} serotypes (Fig. 1.9).

5.4. Serotype conversion in \textit{S. flexneri} \textit{rfc} mutants

Little is known about the particular stage in the biosynthetic pathway of LPS, at which the bacteriophage-encoded modifications (i.e. glucosylation and O-acetylation) are added to the O-antigen repeat units of \textit{S. flexneri}. The function of the Rfc protein is to polymerise the O antigen repeat units into long O antigen chains. A mutation in \textit{rfc} results in the addition of a single O unit, mediated by RfaL, to the lipid A core structure, resulting in a semi-rough (SR)
LPS phenotype. The ability of plasmid constructs to mediate serotype conversion was observed, using strains with mutations in the O antigen polymerase gene.

5.4.1 rfc mutants of *S. flexneri* strains of serotype Y and X.

O-antigen polymerase mutants (*rfc*) of *S. flexneri* strain PE577 of serotype Y have been previously isolated (Morona *et al.*, 1994). *rfc* mutant, RMM109, was isolated which possessed a spontaneous point mutation which was not defined genetically. Strains of serotype X are also sensitive to infection by Sf6 phage, therefore, spontaneous Sf6R mutants of serotype X were isolated in the same manner as described for Y strains (Morona *et al.*, 1994). Spontaneous Sf6R serotype X strains were obtained and assessed for their LPS phenotype. Mutants displaying various phenotypes (ie *rfa*, *rfb* and *rfc*) were obtained as shown in Figure 5.9.

5.4.2 Serotype conversion of spontaneous *rfc* mutants

Strain RMM168 was chosen as a representative *rfc* mutant of a serotype X strain and used with RMM109, an *rfc* mutant from a serotype Y strain. Plasmids harbouring the various serotype conversion genes were introduced into both strains and assessed for their ability to convert the single O-unit present in these strains by slide agglutination and colony immunoblotting. Table 5.5 shows that the introduction of plasmid pRMM264 results in the expression of the type II antigen in both X and Y *rfc* mutants. The expression of group 7,8 by the introduction of a plasmid copy of *gtrX*, additionally requires the presence of a copy of *bgt* in RMM109. O-acetylation in both strains is mediated by the *oac* gene product, resulting in the appearance of the group 6 antigen.
Proteinase K digested whole cell lysates of Sf6 resistant isolates of PE576 of serotype X were electrophoresed in an SDS-15% PAGE and silver stained. The parent strain PE576 is shown, exhibiting wildtype LPS, and the five Sf6R strains, whose inferred genotypes are: RMM166 (rfa), RMM168 (rfc), RMM165 (rfb), RMM167 (rfb) and RMM169 (rfc).
### Table 5.5  Serotype conversion in *S. flexneri* rfc mutants of serotypes Y and X

<table>
<thead>
<tr>
<th>Strain</th>
<th>type II</th>
<th>group 6</th>
<th>group 7,8</th>
<th>Genes on plasmid</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMM109: rfc mutant of serotype Y</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>RMM454: RMM109 [pRMM264]</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>bgt/grtII</td>
<td>2a</td>
</tr>
<tr>
<td>RMM455: RMM109 [pCC142]</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>oac</td>
<td>3b</td>
</tr>
<tr>
<td>RMM168: rfc mutant of serotype X</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>RMM473: RMM168 [pRMM264]</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>bgt/grtII</td>
<td>2b</td>
</tr>
<tr>
<td>RMM474: RMM168 [pCC142]</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>oac</td>
<td>3a</td>
</tr>
</tbody>
</table>

+ indicates a positive reaction in a colony immunoblot
- indicates a negative reaction in a colony immunoblot
Confirmation of these results was achieved by assessing the expression of the type and group antigens in the serotype converted strains by bacterial agglutination assays (Table 5.6). Each strain was prepared as described in section 2.18.3, and assessed for agglutination with various dilutions of antisera. The data obtained via this method correlated well with the colony immunoblotting results which showed that a single O-repeat unit could be modified to give serotype conversion either by whole phage (Section 3.2.4.3), or by the addition of plasmid copies of serotype conversion genes (Table 5.5).

5.5 Distribution of \textit{bgt} and \textit{gtrII} in \textit{S. flexneri}

Analysis of the distribution of \textit{bgt} and \textit{gtrII} in \textit{S. flexneri} was carried out by using PCR-DIG-labelled \textit{bgt} (primers \#2349, \#2376) and \textit{gtrII} (primers \#2325, \#2407), and probing chromosomal DNA of all \textit{S. flexneri} serotypes and \textit{E. coli} K-12 in Southern hybridisations (Figs. 5.10, 5.11). It was found that \textit{bgt} homologues existed in all serotypes of \textit{S. flexneri} and \textit{E. coli} K-12. More than one copy of \textit{bgt} exists within any strain tested of \textit{S. flexneri} and \textit{E. coli} K-12. \textit{gtrII} hybridised strongly only to DNA from strains expressing the type II antigen (i.e. serotypes 2a and 2b) and weakly to strains of serotype 4a, 4b, 5a and 5b. When all \textit{S. flexneri} serotypes were analysed by PCR using \textit{bgt} or \textit{gtrII} specific primers, \textit{bgt} was amplified in serotypes 2a, 2b and X, and \textit{gtrII} was amplified only in serotypes 2a and 2b. This may be due to the specificity of the primers which were designed based on the sequence from bacteriophage SfII, lysogenic to strains of serotype 2. Additionally, the complementation data suggests that the \textit{bgt} homologues detected in PE577 are not functional.
Table 5.6  Bacterial agglutination assay using *S. flexneri* typing serum

Serial two-fold dilutions of antiserum were incubated with $5 \times 10^8$ bacterial. Control strains used were PE877 (serotype 2a), PE576 (serotype X), PE844 (serotype 3b) and *rfc* mutants RMM109 and RMM168 containing serotype converting plasmids.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Titre</th>
<th>Antisera used</th>
<th>Genes on plasmid</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMM109 <em>rfc</em> mutant of serotype Y</td>
<td>-</td>
<td>Type II</td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>RMM454</td>
<td>$&gt;1/10$</td>
<td>Type II ' bgt/gtrII</td>
<td></td>
<td>2a</td>
</tr>
<tr>
<td>RMM109[pRMM264]</td>
<td>1/40</td>
<td>Group 7,8</td>
<td>bgt/gtrX</td>
<td>X</td>
</tr>
<tr>
<td>RMM455</td>
<td>1/40</td>
<td>Group 6</td>
<td>oac</td>
<td>3b</td>
</tr>
<tr>
<td>RMM168 <em>rfc</em> mutant of serotype X</td>
<td>&lt;1/10</td>
<td>Group 6</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>RMM473</td>
<td>$&gt;1/80$</td>
<td>Type II</td>
<td>bgt/gtrII</td>
<td>2b</td>
</tr>
<tr>
<td>RMM168[pRMM264]</td>
<td>1/40</td>
<td>Group 6</td>
<td>oac</td>
<td>3a</td>
</tr>
<tr>
<td>PE877 wildtype serotype 2a</td>
<td>1/20</td>
<td>Type II</td>
<td></td>
<td>2a</td>
</tr>
<tr>
<td>PE844 wildtype serotype 3b</td>
<td>$&gt;1/80$</td>
<td>Group 6</td>
<td></td>
<td>3a</td>
</tr>
<tr>
<td>PE576 wildtype serotype X</td>
<td>$&gt;1/80$</td>
<td>Group 7,8</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Figure 5.10  Southern hybridisation using DIG- labelled ORF3 (bgt)

Chromosomal DNA of representatives of all *S. flexneri* serotypes and two *E. coli* K-12 strains were digested with the restriction enzyme *Pst*I and probed with DIG-labelled ORF3 (bgt). The hybridisation pattern shows that *bgt* is present in multiple copies in the genome of many serotypes.
Figure 5.11  Southern hybridisation using DIG- labelled ORF4 ($gtrII$)

Chromosomal DNA of representatives of all *S. flexneri* serotypes were digested with the restriction enzyme *NsiI* and probed with DIG-labelled ORF4 ($gtrII$). Strong staining is seen with serotypes 2a and 2b and weaker staining is seen in serotypes 4a, 4b, 5a and 5b.
5.6 Comparison of genome organisations in phage P22 and E. coli K-12

The five ORF’s identified in phage SfII have extensive homology to E. coli K-12 and bacteriophage P22 sequences. The organisation of these genes is shown in Figure 5.12. The first ORF in all three genomes corresponds to an excisionase, followed by an integrase gene, adjacent to either an attachment site in the phages or tRNA gene in E. coli K-12. ORF2, whose function is unknown, follows immediately, and exhibits a high degree of similarity to o120 of E. coli K-12, an ORF of 120aa of P22 and ORF10x1 of S. flexneri. The next ORF named bgt in SfII has a homologue in E. coli K-12, o306, which is likely to mediate the same function as bgt. Immediately adjacent to o306 lies an ORF of 443 aa, o443, which has been shown to have a similar hydropathy profile to GtrII and GtrX (Fig. 5.7). This may be the protein which is responsible for the addition of the glucose residue onto GlcNAc (section 5.3.3) when the S. flexneri rfb genes are introduced into E. coli K-12.

A partial ORF of 36 aa at the 5’ end of the attP region of P22 phage has also been identified. This genetic organisation may be common to all serotype converting phages (except Sf6) of S. flexneri, and the genes seen in this region of the E. coli K-12 chromosome near 51 minutes may be remnants of a similar defective phage.

5.7 Summary

In this chapter two genes have been identified that are involved in serotype conversion. The genes, bgt and gtrII, act together to mediate the addition of a glucosyl residue onto the rhamnose III of the repeat unit, resulting in the appearance of the type II antigen. Sequencing of these genes revealed homology to known genes in the Salmonella phage P22 and the E. coli K-12 chromosome. Interestingly, the organisation of these genes appears to be conserved.
Figure 5.12 Comparison of genome organisation of phage SfII, *E. coli* K-12 and phage P22.

The genomes of bacteriophages SfII and P22 and the *E. coli* K-12 chromosome are compared. The arrows indicate the direction of transcription. The integrase genes are all of the same size and immediately adjacent is either the attachment site *attP* or tRNA. ORF2 and homologues AE323o120 and P22o120, all of 120 aa in size, are indicated followed by ORF2 (bactoprenol glucosyl transferase) in SfII, o306 in *E. coli* K-12 and the partial ORF of 36 aa in P22. There is an ORF in *E. coli* K-12 of 443 aa whose function is yet to be determined and is downstream of o306 and in the same position as *gtrII*. 
**SfII**

\[ \text{xis} \quad \text{int} \quad \text{attP} \quad \text{ORF2} \quad (\text{bgt}) \quad o309 \quad \text{glucosyl transferase (gtrII)} \]

**E. coli**

\[ \text{xis} \quad \text{int} \quad \text{thrW} \quad o120 \quad o306 \quad o443 \]

**P22**

\[ \text{xis} \quad \text{int} \quad \text{attP} \quad o120 \quad o1 \]
across these species, implying that possibly a common evolutionary origin exists for all three loci.

Semi-rough mutants of *S. flexneri* have been used to determine the timing of the modifications by assessing the ability of plasmids encoding serotype converting genes to modify a single repeat unit. *S. flexneri rfc* mutants may also provide a useful tool, to understand the stage at which modifications are added. The cellular location of the proteins involved in the addition of the modifications to the O-antigen is of interest and is addressed in the next chapter.
Chapter Six

Characterisation of S. flexneri Type II negative mutants and Bgt protein analyses

6.1 Introduction

Two genes carried by bacteriophage SflI, bgt encoding bactoprenol glucosyl transferase, and gtrII encoding glucosyl transferase, are required for expression of the type II antigen in S. flexneri. To gain an understanding of the requirement of these genes for serotype conversion, chromosomal bgt or gtrII mutants were isolated in a S. flexneri strain of serotype 2a. The result of the mutations was a loss of expression of the type II antigen.

To date, the cellular location of proteins involved in O-antigen modification has not been determined. The previous chapter describes the similarity that exists between the SflI phage Bgt and the S. cerevisiae Dpm1 proteins. Dpm1 is a dolichol phosphate mannosyl synthase and is the best characterised of the family of glycosyl transferase proteins. Based on analysis of protein sequence, it has been predicted that Dpm1 is associated with the membrane of the endoplasmic reticulum (Preuss et al., 1991).

This chapter describes the characterisation of chromosomal bgt and gtrII mutants and the restoration of type II antigen expression by the introduction of plasmids containing the genes bgt and/or gtrII into the mutant strains. Protein analysis using the T7 polymerase/promoter system and cell fractionation was carried out on the Bgt protein, and topology studies conducted using fusions to PhoA and LacZ.
6.2 Results

6.2.1 Isolation of type II antigen negative mutants of *S. flexneri*

Mutants of a *S. flexneri* strain of serotype 2a (PE655) were isolated that no longer expressed the type II antigen. These mutants were isolated in the laboratory of Dr. R. Morona as follows: RMM254 (Rif<sup>R</sup>, PE655) was conjugated with E1196 (TnphoA donor) (Table 2.2), and Rif, Km resistant exconjugants were selected. The exconjugants were harvested, cultured and then agglutinated with anti-type II serum. Bacteria that did not agglutinate with the type II serum were then re-grown and the agglutination process repeated. Single colonies were isolated and then assessed for their reactivity to anti-type II and group 3,4 sera by whole cell immunoblotting. Several type II antigen negative, group 3,4 positive mutants (RMA901, RMA902, RMA903, RMA905, RMA906, RMA907, RMA910 and RMA911) of RMM254 were obtained.

6.2.2 Characterisation of type II antigen negative derivatives of RMM254

Strains RMA901, RMA902, RMA903, RMA905, RMA906, RMA907, RMA910 and RMA911 lacking the type II antigen were characterised by several methods.

Southern hybridisations using DIG-labelled *bgt* and *gtrII* probes were used to assess the nature of the TnphoA insertions in the eight mutants. Initially, chromosomal DNA from the type II negative mutants was isolated and digested with the restriction enzyme *EcoRV*. *EcoRV* does not cleave within *bgt* and was therefore suitable for screening for mutants in the *bgt* gene. Fig 6.1 shows hybridisation of DIG-labelled *bgt* to *EcoRV* digested chromosomal DNA from parent RMM254, and mutants RMA903, RMA906, RMA907, RMA910 and
Figure 6.1 Southern hybridisation using DIG-labelled bgt

Chromosomal DNA of parent strain RMM254 and Type II negative mutants RMA903, RMM906, RMA907, RMA910 and RMA911 were digested with the restriction enzyme EcoRV. The probe used was PCR amplified and DIG-labelled bgt using primers #2349 and #2376, which hybridised to a fragment of ca. 6.2 kb in the parent strain and ca. 14 kb in the mutant strains indicating the insertion of the 7.7 kb TnphoA. bgt homologues are detected in multiple copies in the genome.
RMA911. A fragment of ca. 6.2 kb in size corresponding to wildtype bgt, was detected in RMM254, whereas the mutants (RMA903, RMA906, RMA907, RMA910 and RMA911) uniformly display a fragment of ca. 14 kb in size suggesting a 7.7 kb TnphoA transposon (Manoil and Beckwith, 1985) insertion in bgt. Interestingly, multiple copies of bgt can be seen in the genome of strain RMM254. These multiple copies may correspond either to more than one copy of bgt or homologues, however, they are non-functional, demonstrated by the lack of expression of the type II antigen.

Similarly, the type II negative mutants (RMA901, RMA902 and RMA905) which did not display a size increase when probed with bgt were then hybridised with a DIG-labelled gtrII. Chromosomal DNA of RMA901, RMA902 and RMA905 was digested with enzyme NsiI, which does not restrict within gtrII, and probed with DIG-labelled gtrII. Insertions within this gene were identified when the 10.5 kb fragment containing gtrII increased in size to ca. 18.2 kb as a result of incorporating the 7.7 kb TnphoA transposon (Fig. 6.2).

RMA903 (bgt::TnphoA) (Fig. 6.1) and RMA901 (gtrII::TnphoA) (Fig. 6.2) were chosen as the representative strains to be used for further characterisation and complementation analyses.

Silver stained SDS-15% PAGE gels were used to analyse the LPS profiles of mutants RMA901 and RMA903. These mutants do not express the type II antigen which implies that they lack the glucose sugar linked to the third rhamnose of the repeat unit (Table 1.3, Fig. 1.9). This would result in their LPS displaying a slightly decreased molecular mass when compared to the parent 2a strain and, therefore, an altered migration. Figure 6.3 (A) is a silver stained SDS-PAGE gel which shows the parent strain, RMM254, RMA901 and RMA903. The individual LPS bands of the mutants migrate slightly faster and to a lower position in the gel than those of the parent strain.
Figure 6.2 Southern hybridisation using DIG-labelled gtrII

Chromosomal DNA of parent strain RMM254 and Type II negative mutants RMA901, RMA903 and RMA905 were digested with the restriction enzyme NsI. The probe used was PCR amplified and DIG-labelled gtrII using primers #2325 and #2407, which hybridised to a fragment of ca. 10.5 kb in the parent strain and ca. 18.2 kb in the mutant strains.
Figure 6.3 Silver stained SDS-15% PAGE and Western immunoblot of Type II negative mutants and complemented strains.

A: Silver stained SDS-15% PAGE gel displaying LPS patterns of:
   - RMM254, parent strain of serotype 2a
   - RMA901, type II negative mutant
   - RMM334, SfII lysogen of RMA901
   - RMM400, RMA901 + pRMM264
   - RMM463, RMA901 + pRMM271
   - RMA903, type II negative mutant
   - RMM336, SfII lysogen of RMA903
   - RMM468, RMA903 + pRMM264
   - RMM465, RMA903 + pRMM270

B: Western immunoblot of the same strains in the same order using anti-type II serum.
To confirm the loss of the type II antigen, western immunoblots were also carried out (Fig. 6.3B). The parent strain, PE254 reacts to the polyclonal anti-type II serum, whereas the two mutant strains, RMA901 and RMA903 did not react with this antiserum.

To confirm the Southern hybridisation data, primers designed to both *gtrII* (Table 2.4, #2325, #2407) and *bgt* (Table 2.4, #2349, #2376) were used in separate PCR reactions with chromosomal DNA from strains RMA903 and RMA901. Product sizes of 930 bp for *bgt* and 1.4 kb for *gtrII* were expected; PCR with RMA903 resulted in a 1.4 kb fragment when *gtrII*-specific primers were used, however, no product observed when *bgt*-specific primers were used. A 930 bp product was observed when RMA901 chromosomal DNA was used in a PCR amplification reaction with *bgt*-specific primers, however, with *gtrII*-specific primers no product was seen. The absence of a PCR product in strains RMA903 and RMA901 when amplified with *bgt* and *gtrII* primers, respectively, is due to the size of the region of DNA between the two primers, resulting from the insertion of the 7.7 kb TnphoA. Table 6.1 shows that the PCR data correlates with the Southern hybridisation data, and confirms the location of the TnphoA insertions in RMA901 and RMA903.

### 6.2.2.1 Inverse PCR of TnphoA insertion in RMA901

Further characterisation of mutant RMA901 was carried out using the inverse PCR (IPCR) technique, in order to identify the insertion point of the transposon TnphoA. Southern hybridisations were carried out on EcoRV digested chromosomal DNA from strains RMM254 and RMA901, using DIG- labelled Tn5 as a probe (Fig. 6.4). The enzyme EcoRV does not restrict within the *bgt* gene or within the TnphoA. The transposon Tn5 carries the same kanamycin resistance cartridge as that in TnphoA and can be used to hybridise to the latter. A low level of hybridisation to this probe is seen in the parent strain and the same hybridisation pattern is observed in RMA901. However, in RMA901, the probe hybridised to a new
Table 6.1 Characterisation of TnphoA insertion mutants

<table>
<thead>
<tr>
<th>Strain</th>
<th>LPS</th>
<th>Western immunoblot</th>
<th>PCR</th>
<th>Southern hybridisation</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMA901</td>
<td>altered migration</td>
<td>negative for type II antiserum</td>
<td>negative for ( bgt )</td>
<td>increased size when hybridised with ( bgt )</td>
<td>Y</td>
</tr>
<tr>
<td>RMA903</td>
<td>altered migration</td>
<td>negative for type II antiserum</td>
<td>negative for ( gtrII )</td>
<td>increased size when hybridised with ( gtrII )</td>
<td>Y</td>
</tr>
</tbody>
</table>
Figure 6.4 Southern hybridisation using DIG-labelled Tn5

Chromosomal DNA of parent RMM254 and type II negative mutant RMA901 were digested with the restriction enzyme EcoRV and then hybridised with DIG-labelled Tn5. The hybridisation patterns show that digestion of RMA901 chromosomal DNA with EcoRV results in the localisation of Tn5 on a fragment of 8.5 kb in size.
fragment of ca. 8.5 kb in size, which corresponds to the 7.7 kb TnphoA and 0.8 kb of adjacent DNA. Chromosomal DNA from RMA901 was digested with EcoRV and ligated with T4 DNA ligase. The IPCR fragment expected to be amplified using primers #2210 and #2211 (Table 1.4) which read out of the end of TnphoA, is 0.8 kb in size. The resultant 0.8 kb product was cloned into the vector pGEM-T and subjected to DNA sequencing using M13 primers. This revealed that the insertion point is after nt. 4639 (Fig. 4.3) in gtrII.

6.2.3 Complementation studies

Various constructs were used for complementation analysis with strains RMA901 and RMA903. Plasmid pRMM264 (bgt/gtrII) and subclones pRMM266 (gtrII) and pRMM268 (bgt) described in Section 3.2.7 (Fig. 3.10), and PCR derived constructs pRMM270 (bgt), pRMM271 (gtrII) and pRMM272 (bgt/gtrII) (Section 5.3, Fig. 5.8), were used as described in Table 6.2. The complementation was initially performed by bacterial agglutination and colony immunoblotting using anti-type II and anti-group 3,4 sera.

Lysogenisation of RMA901 (gtrII::TnphoA) and RMA903 (bgt::TnphoA) by SfiI phage restored expression of the type II antigen. Additionally, introduction of pRMM264, into RMA901 and RMA903 also restored expression of type II antigen. However, plasmid pRMM266 only restored expression of type II antigen in RMA901 and similarly plasmid pRMM268 complemented only in RMA903.

When plasmids pRMM270, pRMM271 and pRMM272 were introduced into the mutant strains, complementation was achieved with pRMM272, which contains both bgt and gtrII, in both RMA901 and RMA903. However, when pRMM270 and pRMM271 were introduced, only partial complementation was achieved. In the case of strain RMM463 [RMA901 + pRMM271], containing a plasmid copy of gtrII, staining in the colony immunoblot was weak when compared to the same strain containing the other plasmid
Table 6.2 Complementation of Type II negative strains RMA901 and RMA903

<table>
<thead>
<tr>
<th>Strain/ plasmid</th>
<th>anti-3,4</th>
<th>anti-II</th>
<th>Serotype</th>
<th>Gene</th>
<th>Sf6 sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMA901 (gtrII::TnphoA)</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>bgt</td>
<td>S</td>
</tr>
<tr>
<td>RMM334: RMA901 + SfII (lysogen)</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMM400: RMA901+ pRMM264</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMM424: RMA901+ pRMM266</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMM426: RMA901+ pRMM268</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>bgt</td>
<td>S</td>
</tr>
<tr>
<td>RMM462: RMA901 + pRMM270</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>bgt</td>
<td>S</td>
</tr>
<tr>
<td>RMM463: RMA901 + pRMM271</td>
<td>++</td>
<td>+/-</td>
<td>2a</td>
<td>gtrII</td>
<td>P</td>
</tr>
<tr>
<td>RMM464: RMA901 + pRMM272</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMA903 (bgt::TnphoA)</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>gtrII</td>
<td>S</td>
</tr>
<tr>
<td>RMM336: RMA903 + SfII (lysogen)</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMM468: RMA903+ pRMM264</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
<tr>
<td>RMM469: RMA903+ pRMM266</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>gtrII</td>
<td>S</td>
</tr>
<tr>
<td>RMM470: RMA903+ pRMM268</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt</td>
<td>R</td>
</tr>
<tr>
<td>RMM465: RMA903 + pRMM270</td>
<td>++</td>
<td>+</td>
<td>2a</td>
<td>bgt</td>
<td>R</td>
</tr>
<tr>
<td>RMM466: RMA903 + pRMM271</td>
<td>++</td>
<td>-</td>
<td>Y</td>
<td>gtrII</td>
<td>S</td>
</tr>
<tr>
<td>RMM467: RMA903 + pRMM272</td>
<td>++</td>
<td>++</td>
<td>2a</td>
<td>bgt/gtrII</td>
<td>R</td>
</tr>
</tbody>
</table>

++ indicates positive reaction to antiserum in colony immunoblot comparable to wildtype
+ indicates positive reaction to antiserum in colony immunoblot weaker than wildtype
+/- indicates positive reaction to antiserum in colony immunoblot weaker than wildtype but stronger than negative
- indicates no reactivity to antiserum by colony immunoblot
P partial sensitivity to Sf6
R resistance to Sf6
S resistance to Sf6
constructs (pRMM264, pRMM266 and pRMM272). This was also true for strain RMM465, (RMA903 containing the complementing plasmid pRMM270), where only weak staining was observed, compared with complementing plasmids pRMM264, pRMM268 and pRMM272. It may be possible that an insertion in either bgt or gtrII exerts a polar effect downstream or upstream, respectively, and, therefore, both bgt and gtrII must be supplied for successful complementation.

Sensitivity to Sf6 phage of the complemented mutants was also determined (Table 6.2). It was observed that complementation of either mutation by SfII lysogenisation or complementation with plasmids pRMM264 or pRMM272 restored resistance to Sf6, however, partial sensitivity was observed in strains RMM426, RMM462, RMM469 and RMM466 corresponding with their reduced reactivity to anti-type II serum.

6.2.3.1 LPS and western immunoblot analyses of complemented mutants RMA901 and RMA903

Comparisons of migration patterns of LPS on SDS-PAGE were made. In addition, western immunoblotting was carried out with the complemented strains RMM334, RMM400, RMM463, RMM336, RMM468 and RMM465 (Fig 6.3B) using anti-type II serum. The type II negative mutants RMA901 and RMA903 did not react with anti-II serum. Lysogenisation with bacteriophage SfII restored expression of type II antigen in both strains, as did plasmid pRMM264. Introduction of bgt and gtrII into the strains produced differing results. pRMM271 (gtrII) did not restore detectable type II antigen expression in RMA901, whereas pRMM270 (bgt) could to some extent with RMA903. This correlates well with the data in Table 6.2. An explanation for the observation made (Table 6.2, Figure 6.3B) is that the transposon insertion in these strains may be exerting a polar effect on bgt upstream of gtrII.
and not allowing efficient expression of bgt, thus complementation of this defect is only achievable by addition of both genes (bgt and gtrII).

### 6.3 Detection of bgt and gtrII gene products

To detect proteins encoded by genes bgt and gtrII, plasmids pRMM270, pRMM271 and pRMM272, (pGEM-T clones of bgt, gtrII and bgt|gtrII respectively (Section 4.3)), were used. The plasmids were introduced into E. coli K-12 strain E2096 harbouring pGP1-2, generating strains RMM446, RMM447 and RMM448. pGEM5Zf+ is the vector from which pGEM-T is derived and was used as the control plasmid in RMM449.

The in vivo T7 polymerase-driven expression system of Tabor and Richardson (1985) was used for expression of Bgt and GtrII (Section 2.19). The autoradiograph (Fig. 6.5) shows one unique protein encoded by RMM446 containing bgt on pRMM270 compared to the control strain RMM449 containing plasmid pGEM5Zf+. A strongly labelled protein migrating at ca. 34 kDa, consistent with the predicted size of Bgt was detected (Fig. 6.5) (Table 4.1). A 34 kDa protein was also observed in the lane containing strain RMM448 (pRMM272) which contained both bgt and gtrII. However, it was weakly labelled in this strain and was therefore not shown.

A labelled protein band at approximately 55 kDa was expected for GtrII with strains RMM447 and RMM448, however, this was not detected. One possible explanation is that the low G+C% content of gtrII of 31.8% (Table 4.1) coupled with the number of rare codons in the first 25 amino acids regulates translation of gtrII in a manner similar to that found for the O-antigen polymerase gene, rfc, which is also not visualised in Coomassie Blue stained gels or by autoradiography after L-[^35S]-methionine labelling (Morona et al., 1994; Daniels et al., 1998).
Figure 6.5 Expression of Bgt using the T7 polymerase/promoter system.

The T7 polymerase/promoter system (Tabor and Richardson, 1985) was used to overexpress proteins encoded on genes contained on plasmid pRMM270 (bgt), which was then labelled with $[^{35}S]$-methionine. Autoradiography of strains RMM449 [E2096 + pGEM5Zf+] and RMM446 [E2096 + pRMM270] electrophoresed on an SDS-15% PAGE exhibited bands of the predicted size of Bgt of 34 kDa in the track containing RMM446.
6.3.1 Cellular location of Bgt in *E. coli*

The predicted role of Bgt is to facilitate the transfer of glucose residues from the charged nucleotides to the lipid carriers prior to their transfer via the glucosyl transferase to the O antigen acceptor, suggesting it may be active in the cytoplasmic space. Analysis of hydropathy plots (Fig. 5.3) suggested the Bgt protein has two membrane spanning domains near the carboxy terminal end and a large hydrophilic region at the amino terminus; these characteristics imply the protein may be membrane associated.

To determine the subcellular location of Bgt, the overexpressed protein was L-[³⁵S]-methionine labelled and fractionated. Figure 6.6 shows the whole cell and cell fractions of strain RMM446 [E2096 + pRMM271] and the control strain RMM449 [E2096 + pGEM5Zf+] into the components periplasmic fraction (PF), cytoplasmic fraction (CF), whole membrane (WM) and inner (IM) and outer (OM) membrane fractions (for RMM446). The 34 kDa Bgt protein appears to be localised primarily in the cytoplasmic and whole membrane fractions and in the inner membrane. The presence of Bgt in the cytoplasm and in the inner membrane implies that it may exist in both a soluble form as well as a membrane associated form and may be a peripheral membrane protein.

6.3.2 Studies on the topology of Bgt using *phoA* and *lacZ* fusions

Analysis of the amino acid sequence of Bgt using the Kyte and Doolittle program (PROSIS) (Kyte and Doolittle, 1982) shows two clear hydrophobic regions located at aa 231-255 and aa 263-285. To determine the orientation of the protein with respect to the membrane, four oligonucleotide primers (#2719, #2720, #2721 and #2722 Table 2.4) were designed to construct in-frame fusions with vectors pRMCD28 (‘*phoA*) and pRMCD70 (‘*lacZ*) to different positions in Bgt. The fusions were located at the carboxy terminal end
Figure 6.6 Fractionation of cells expressing the Bgt protein

Strain RMM446 [E2096 + pRMM270] was labelled with $[^{35}S]$-methionine and fractionated into cellular components; whole cells (WC), periplasmic fraction (PF), cytoplasmic fraction (CF), whole membrane (WM), outer membrane (OM) and inner membrane (IM). Samples representing $10^8$ bacteria were electrophoresed on an SDS-15% PAGE gel. The autoradiograph shows a distinct product at ca. 34 kDa corresponding to the presence of plasmid encoded $bg$ in the WC, CF, WM and IM fractions.
near the predicted hydrophobic domains (fusion 1 at nt. 669, aa 223; fusion 2 at nt. 783, aa 261; and fusion 3 at nt. 927, aa 309) (Fig. 6.7). Six plasmids resulted (pRMM274, pRMM275, pRMM276, pRMM277, pRMM278 and pRMM279), which correspond to the three amplified products in vectors pRMCD28 and pRMCD70, respectively. Sequencing using primers (#2211 and #2254) which read outwards from the phoA and lacZ genes, respectively, confirmed the fusion points of the constructs.

Plasmids pRMM274, pRMM275, pRMM276, pRMM277, pRMM278 and pRMM279 were introduced into E. coli strain CC118 to give RMM475, RMM476, RMM477, RMM478, RMM479 and RMM480. The strains were assessed for either alkaline phosphatase or β-galactosidase activity on NA containing ampicillin and BCIP and BCIG, respectively (Table 6.3). Alkaline phosphatase is inactive in the cytoplasm and β-galactosidase is inactive in the periplasm, implying that fusions 1 and 3 are cytoplasmically located and fusion 2 is located in the periplasm. The location of the fusion junctions are shown on the topological model of Bgt (Fig. 6.8).

6.3.2.1 Tm-predict model of Bgt topology

The Tm-predict package (Hoffman and Stoffel, 1993) was also used. It predicted three inside to outside transmembrane helices, involving aa 50-68, 228-249 and 264-283, and three outside to inside helices, involving aa 52-68, 228-246 and 264-280, with the first in each case considered insignificant. Two models were suggested for the transmembrane topology by the program when the transmembrane helix length was set between 17 and 23 amino acids. Both models had two transmembrane helices with the “strongly preferred” model having the amino and carboxy terminal in the periplasm and the “alternative model” with both termini in the cytoplasm. Overall the activities correlate well with the “alternative model” predicted by the Tm-predict program (Fig. 6.8).
Figure 6.7 Hydropathy plot of ORF3

Primers #2719, #2720, #2721 and #2722 were designed incorporating PstI and XbaI restriction sites to enable the construction of fusions of Bgt of various lengths to PhoA and LacZ by cloning into vectors pRMCD28 and pRMCD70, respectively. Resultant plasmids derived from cloning of the different products into the vectors are listed, and the ability of these constructs to restore expression of type II antigen in RMA901(bgt::TnphoA) by colony immunoblotting was also assessed.
ORF3

Fusion 1
#2719

Fusion 2
#2720

Fusion 3
#2721

pRMCD28/70

pRMM274/277

pRMM275/278

pRMM276/279

Serotype Conversion in RMA901

- - +
Table 6.3 Characterisation of fusions using colour indicator plates, western immunoblotting of cell fractions and enzyme assays.

<table>
<thead>
<tr>
<th>Fusion</th>
<th>nt. position</th>
<th>aa</th>
<th>BCIP</th>
<th>BCIG</th>
<th>localisation</th>
<th>( \beta)-galactosidase</th>
<th>alkaline phosphatase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>669</td>
<td>223</td>
<td>-</td>
<td>+</td>
<td>cytoplasmic</td>
<td>4615 ± 168.65</td>
<td>56 ± 11</td>
</tr>
<tr>
<td>2</td>
<td>783</td>
<td>261</td>
<td>+</td>
<td>-</td>
<td>periplasmic</td>
<td>125 ± 23</td>
<td>280 ± 31.09</td>
</tr>
<tr>
<td>3</td>
<td>927</td>
<td>309</td>
<td>-</td>
<td>+</td>
<td>cytoplasmic</td>
<td>2535 ± 176</td>
<td>6 ± 1.73</td>
</tr>
<tr>
<td>Control*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5393 ± 168.65</td>
<td>423 ± 30.98</td>
</tr>
</tbody>
</table>

* Controls used were DH5 for \( \beta\)-galactosidase activity and E2308 for alkaline phosphatase activity
Figure 6.8 Topological model of *S. flexneri* Bgt protein.

This model is based on *bgt::phoA* and *bgt::lacZ* defined fusions and computer prediction data. The amino acids in the transmembrane and periplasmic domains are represented by their single letter code whereas the cytoplasmic domain is represented by the predicted structures as determined in PROSIS (Garnier *et al.*, 1978). The location of the fusions are indicated and the amino acid positions numbered.
6.3.2.2 Western immunoblotting of Bgt::PhoA and Bgt::LacZ fusions

To assess the stability and subcellular location of the Bgt::PhoA and Bgt::LacZ fusion proteins, western immunoblotting was performed using: A) anti-alkaline phosphatase and (B) anti-β-galactosidase sera (Fig. 6.9). Whole cell and cell fractions (periplasm, cytoplasm and whole membrane) of the three Bgt::PhoA fusions and three Bgt::LacZ fusions were tested. Only Bgt::PhoA fusion 2 (RMM476) (aa 261) panel (A), and Bgt::LacZ fusions 1 (RMM478) (aa 223) and 3 (RMM480) (aa 309) were readily detected. Fractionation of the cells showed that the fusion proteins were located in the cytoplasmic and membrane fractions. The truncated proteins of 24 kDa (fusion #1) and 29 kDa (fusion #2) in size and the full length 34 kDa (fusion #3) were fused to the 46.5 kDa PhoA and 94 kDa LacZ proteins, respectively. Predicted fusion products sizes were 71 kDa (#1), 75 kDa (#2) and 80 kDa (#3) for PhoA and 118 kDa (#1) and 122 kDa (#2) and 128kDa (#3) for LacZ. Some breakdown products reactive with anti-PhoA and anti-LacZ were detected. In addition, a non-specific band at ca. 58 kDa cross-reacting with the antiserum is seen in Fig. 6.9B that is common to all lanes, including strain DH5 harbouring the control plasmid pRMCD70. Comparisons of CF and WM fractions of fusions 1 (RMM478) and 3 (RMM480) with LacZ show more association of the former with the cytoplasmic fraction, and whole membrane fraction for the latter.

Enzyme activities for alkaline phosphatase and β-galactosidase were measured for each of the fusion proteins (Table 6.3). Generally, the periplasmically located alkaline phosphatase fusions have 5-10 fold greater activity than those located in the cytoplasm, and the cytoplasmically located β-galactosidase fusions exhibited greater activity than those located in the periplasm. Control strain DH5 was induced by the addition of 2 mM of IPTG.

Plasmids pRMM274, pRMM275 and pRMM276 were introduced into RMA901 (bgt::TnphoA) and assessed for restoration of expression of type II antigen. The resultant strains were colony immunoblotted using anti-type II serum and it was determined that only pRMM276 was able to complement the defect in RMA901. The two truncated Bgt fusions
Figure 6.9  Western immunoblots of *E. coli* CC118 strains producing Bgt::PhoA (A) and Bgt::LacZ (B) fusion proteins. After electrophoresis on an SDS-10% polyacrylamide gel and transfer to nitrocellulose, proteins were detected using either A) anti-PhoA or B) anti-LacZ serum.

Whole cell (WC) and cell fractions (periplasmic, PF; cytoplasmic, CF; whole membrane, WM) of strains

(A)  RMM475 - CC118 [pRMM274] (fusion 1 in pRMCD28)
     RMM476 - CC118 [pRMM275] (fusion 2 in pRMCD28)
     RMM477 - CC118 [pRMM276] (fusion 3 in pRMCD28)
     RMM481 - CC118 [pRMCD28]

(B)  RMM478 - CC118 [pRMM277] (fusion 1 in pRMCD70)
     RMM479 - CC118 [pRMM278] (fusion 2 in pRMCD70)
     RMM480 - CC118 [pRMM279] (fusion 3 in pRMCD70)
     RMM482 - CC118 [pRMCD70]
### A

<table>
<thead>
<tr>
<th>C75a</th>
<th>RMM481</th>
<th>RMM475</th>
<th>RMM476</th>
<th>RMM477</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>PF</td>
<td>CF</td>
<td>WM</td>
<td>WC</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>DH5</th>
<th>RMM482</th>
<th>RMM478</th>
<th>RMM479</th>
<th>RMM480</th>
</tr>
</thead>
<tbody>
<tr>
<td>WC</td>
<td>PF</td>
<td>CF</td>
<td>WM</td>
<td>WC</td>
</tr>
</tbody>
</table>

### Notes
- The images show Western blot analyses comparing various strains (C75a, RMM481, RMM475, RMM476, RMM477) and mutations (DH5, RMM482, RMM478, RMM479, RMM480).
- Each lane represents a different condition or mutation, with molecular weight markers indicated at 175, 83, 62, 47.5, 32.5, and 25.
were unable to complement, possibly due to their inability to associate with the cytoplasmic membrane. This suggests that insertion into the cytoplasmic membrane is necessary for function.

6.4 Summary and conclusions

Two genes have been identified, bactoprenol glucosyl transferase (bgt) and glucosyl transferase (gtrII), which are essential for successful modification of S. flexneri O-antigen. Mutations in either of these genes were shown to abolish expression of type II antigen. Complementation could be achieved by the lysogenisation by using either SfIII phage, or via the addition of the wildtype copy of either bgt or gtrII (in most cases).

Fractionation of cells with L-[35S]-methionine-labelled Bgt localised this protein to the cytoplasmic and membrane fractions, whereas topological analysis using fusions to phoA and lacZ reporter genes have identified that Bgt is highly likely to be located in the cytoplasmic membrane, and is anchored by the two transmembrane segments located near the carboxy terminus.
CHAPTER SEVEN

DISCUSSION

7.1. Introduction

*Shigella flexneri* is the causative agent of bacillary dysentery in humans and poses a major health problem in many developing countries. Children under the age of 5 years are particularly susceptible and comprise the majority of those killed by this pathogen each year. Adults who have previously suffered from bacillary dysentery are also susceptible to subsequent infection by other serotypes of the same species, but appear to be protected against infection from the same serotype. Therefore, an understanding of the molecular basis for serotype diversity would assist greatly in the development of a vaccine.

The different serotypes of *S. flexneri* (Fig. 1.9) arise from lysogenisation of strains by serotype converting bacteriophages such as Sf6 (Clark *et al.*, 1991; Verma *et al.*, 1991), SfX (Verma *et al.*, 1993), SfII (this study), SfV (Huan *et al.*, 1997a; 1997b) and SfI (Bastin *et al.*, 1997). These bacteriophages have genes which encode enzymes for either glucosylation or O-acetylation via defined linkages at specific positions on the repeat units of the O-antigen of the lipopolysaccharide (LPS). The O-antigen of LPS acts as the receptor for these serotype converting bacteriophages and is modified during lysogenisation, resulting in an alteration of the receptor and hence resistance to superinfection by the same bacteriophage.

The LPS of *S. flexneri* has been determined to be of significance in the virulence of the organism (Formal *et al.*, 1970; Sansonetti *et al.*, 1983; Okada *et al.*, 1991a; Sandlin *et al.*, 1995, 1996; Van den Bosch *et al.*, 1997). As stated, infection with any serotype of *S. flexneri*
results in immunity to that serotype. However, the individual is still susceptible to infection by other serotypes. A study conducted by Formal and coworkers using monkeys immunised with *S. flexneri* of serotype 2a showed resistance to subsequent challenge with a strain of serotype 2a but not to *S. sonnei* (Formal et al., 1991).

In this thesis, bacteriophage SfII has been isolated and characterised with respect to its host range, serotype converting ability and its genome has been physically mapped. The genes responsible for the serotype conversion mediated by this phage were identified and sequenced. Analysis of one of the gene products (Bgt) were carried out and a cellular location for the modification reaction has been proposed. Finally, the stage at which the modifications are added in *S. flexneri* LPS biosynthesis has been investigated.

### 7.2 Bacteriophage SfIII and other serotype determining bacteriophages

Isolation of bacteriophage from various strains of different serotypes of *S. flexneri* proved to be difficult and several approaches were taken to induce (the excision of) these phages. It appears that some bacteriophages, once integrated, may undergo a deletion resulting in their subsequent inability to excise despite various methods of induction. Mutants of bacteriophage λ (*λ*<sub>xii</sub>) have been isolated which are able to form plaques and lysogens but are defective in excision (Guarneros and Echols, 1970). Other mutations which may result in defective excision include *int* mutations or an alteration of the prophage attachment site.

The serotype converting phage SfIII belongs to group A of the Bradley morphological classification (Bradley, 1967) and is unable to hydrolyse the O polysaccharide of the strains it infects. Other serotype converting bacteriophages of *S. flexneri* have also been classified according to this scheme: Sf6 and SfX belong to group C and SfV belongs to group B. O-antigen hydrolysis assays using Sf6 phage have shown that long O chains of *S. flexneri* strains of serotype X and Y are cleaved during Sf6 infection (Gemski et al., 1975) presumably due to
its endorhamnosidase activity (Chua, 1996). Phage SfII does not possess O-antigen hydrolytic activity and it is unknown whether SfX and SfV perform the same function. It would be of interest to observe whether endoglycosidic activity could be correlated to a morphological group.

Bacteriophages SfX and SfV, like SfII are able to mediate the addition of a glucose residue to the O-repeating units of S. flexneri LPS. Lysogenisation by phage SfX results in the addition of a glucose residue to rhamnose I of the repeat unit and hence production of the group 7,8 antigen, while SfV mediates the glucosylation of rhamnose II and the production of the type V antigen. Although these phages belong to different morphological groups, SfII and SfV share remarkable sequence homology in the region encoding serotype conversion (discussed later).

Bacteriophage Sf6 appears to share more in common with the serotype converting bacteriophages identified in Salmonella anatum, e$^{15}$, e$^{34}$ and g$^{34}$, and P22 of Salmonella typhimurium (Lindberg, 1977). These phages belong to Bradley group C and are able to hydrolyse the LPS of their host strain by specific enzymes located in their tail spike proteins (Lindberg, 1977). These phages mediate inhibition of acetylation of the galactose residue of the repeat unit by e$^{15}$ and g$^{34}$, and alteration of α linkages by the former, while e$^{34}$ is able to add a glucosyl residue to the galactose sugar in the repeating unit. Bacteriophages, e$^{15}$, e$^{34}$, g$^{34}$ and P22 do not recognise semi-rough isolates of their host cells, implying one O-unit is insufficient as a receptor (Lindberg, 1977). This contrasts the situation seen with phage SfII which recognises the single repeat unit of rfc mutants, lysogenises the strain and modifies the single repeat unit of the SR-LPS.

Bacteriophage SfII morphologically resembles bacteriophage Mu; both phages possess an icosahedral head, neck structure, contractile tail and tail fibres. The latter is not capable of mediating serotype conversion. However, does possess the ability to alter its host range based
on the orientation of the invertible G DNA region (Kamp et al., 1978). The morphological similarity between the two bacteriophages is of interest when comparing the sizes of proteins of purified SfII phage with those of phage Mu. Major proteins of sizes 35, 35 and 40 kDa and minor proteins (15-67 kDa) were identified in SfII (Fig. 3.2) which may correspond to specific structures such as the head (33 kDa), tail sheath (55 kDa) and tail spike proteins (16-64 kDa) of Mu (Howe, 1987). Complete sequencing of the SfII genome and comparisons with databases will aid prediction of the sizes of the structural proteins and determination of their location within the chromosome of the phage.

### 7.3 Lysogens of bacteriophage SfII

Bacteriophage SfII is able to infect and form lysogens on strains of serotypes 1b, X and Y. It recognises an as yet uncharacterised receptor in the O-antigen of these serotypes, which is then blocked by the type II modification. SfII lysogens all express the type II antigen as demonstrated by immunoblotting and bacterial agglutination assays (Table 3.1, Fig. 3.4).

It has been reported that strains of serotype Y carrying plasmid copies of glucosyl transferase gene of bacteriophage SfX (gtrX) resulted in conversion to serotype X (Verma et al., 1993). However, when assessed by agglutination using anti-group 7,8 and anti-group 3,4, reactivity to the original group 3,4 antigen remained (Verma et al., 1993). When plasmid copies of gtrY, of bacteriophage SfV, were introduced into strains of serotype Y, residual Sf6 sensitivity was observed, indicating that possibly not all of the repeat units had been modified by the addition of a glucosyl residue (Huan et al., 1997a). It was concluded that the serotype conversion mediated by these plasmids containing the serotype converting genes was partial.

Bacteriophage SfII lysogenisation of a strain of serotype Y (I;3,4) results in a 2a (II;3,4) serotype and retains the original group 3,4 antigen. Therefore, the 'completeness' of the conversion mediated cannot be assessed by continued reactivity to group 3,4 antiserum but
can be determined using bacterial agglutination with type II antiserum. In addition, sensitivity to bacteriophage Sf6 was used to determine the extent of serotype conversion. *S. flexneri* lysogens of SfII and strains carrying serotype converting plasmids showed continued sensitivity to Sf6 phage, however, they continued to agglutinate at a greater dilution of type II antiserum than a natural isolate. In comparison, a clinical isolate of serotype 2a showed complete resistance to Sf6 phage implying that a difference exists between natural isolates of serotype 2a and Y strains that have been converted *in vitro* to express the type II antigen. The partial sensitivity to Sf6 phage may be due to mixed populations of lysogens in which the phage has excised from a proportion of cells. It appears that phage sensitivity assays using Sf6 phage are not an accurate means of determining serotype conversion.

### 7.4 Genome mapping of bacteriophage SfII

Comparison of phage SfII with well-characterised phages such as P22, λ and Mu, revealed a great deal of similarity in the organisation of their genomes (Fig 7.1). Structurally the genomes of bacteriophages λ and P22 are almost perfectly conserved (Poteete, 1988).

The functions associated with DNA replication (*PstI*-6 and 7) are encoded adjacent to the integration region (*intlxis*) and the bacteriophage attachment site, *attP* (*PstI*-3), which is followed by the serotype converting genes in phage SfII (Fig. 3.8) and SfV (Huan *et al.*, 1997b). In phage P22, the *al* gene which has been implicated in antigenic conversion mediated by this phage, is found in the same location (Poteete, 1988), whereas in phage Mu, the corresponding region encodes the invertible G segment (Howe, 1987).

If the organisation of phage genomes is similar then, it is proposed that the genes encoding biosynthesis of tail proteins would be located on *PstI* fragments 2 and 1 and the genes for head biosynthetics would be located in *PstI* fragments 1 and 5. Extensive sequencing needs to be completed to confirm the precise location of these genes. Sequence analysis of
Figure 7.1  Comparison of organisation of bacteriophages SfII, P22, λ and Mu.

The organisation of the genomes of phages P22, λ and Mu are shown. The regions of the chromosome that determine DNA replication, integration, head and tail structures and late control genes have been indicated. Preliminary sequencing of bacteriophage SfII PstI fragments show homology to genes with functions indicated. The location of the SfII head and tail structural genes has also been postulated.
<table>
<thead>
<tr>
<th>SfII</th>
<th>DNA replication</th>
<th>Integration</th>
<th>Serotype Conversion</th>
<th>? Tail</th>
<th>? Head</th>
<th>Lysis</th>
<th>Late Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0</td>
<td>xis int attPbgt gtrII</td>
<td>holin</td>
<td>Q</td>
<td>42.3 kb</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P22</th>
<th>DNA replication</th>
<th>Integration</th>
<th>Antigenic Conversion</th>
<th>Tail</th>
<th>Head</th>
<th>Lysis</th>
<th>Late Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>18,12</td>
<td>int xis</td>
<td>al</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13,19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43.6 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>λ</th>
<th>DNA replication</th>
<th>Recombination</th>
<th>Non essential</th>
<th>Tail</th>
<th>Head</th>
<th>cos</th>
<th>Lysis</th>
<th>Late control</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>xis int attP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.5 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mu</th>
<th>Integration</th>
<th>Replication</th>
<th>atp</th>
<th>Tail</th>
<th>Non essential</th>
<th>Head</th>
<th>Lysis</th>
<th>Late control</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ORF3 in bacteriophage SfV, revealed ca. 30% homology to tail fibre proteins of 3 different species, Bacteriophage P2 gene G, Shigella boydii orfB175 and phage \( \lambda \) tfa gene (Huan et al., 1997a). ORF3 is located downstream of the serotype converting genes which is consistent with the predicted organisation of the SfII genome.

Psrl fragment 8 was found to encode a homologue to the holin (lysis) gene of phage phi-105. Further downstream, a homologue of gene Q, spanning fragments 4 and 6 was identified. Q regulates late gene transcription, consistent with the position of homologous genes in P22, \( \lambda \) and Mu. In addition, bacteriophage SfII has sequences with a great deal of sequence similarity to the E. coli K-12 genome indicating a common evolutionary origin with the defective phage DLP12 present in the E. coli K-12 genome. SfII possesses P22 xis, int, attP and bgt homologues and may, therefore, be a close derivative of this phage.

7.5 Bacteriophage SfII attachment site

The attachment site of SfII was located in Psrl fragment 3 as predicted by Southern hybridisations of chromosomal DNA of lysogens (Fig. 4.5). Sequence analysis revealed the attP site of phage SfII to be identical to that of P22 (Section 4.2.4), with high similarity between the integrase (int) and excisionase (xis) genes, as also seen with the attP, int and xis genes of SfV.

The integration site of phage SfII had been determined to map near the lac-pro loci by conjugation experiments with S. flexneri and E. coli K-12 (Petrovskaya and Nevskaya, 1975; Petrovskaya and Licheva, 1982). It has now been determined that this location is due to the use of identical attP and attB sites. These data were confirmed by Southern hybridisations using chromosomal DNA from the SfII lysogens and probing with a DIG-labelled fragment amplified from this region of the chromosome. Two fragments were observed to hybridise to this probe which is consistent with splitting of the attachment sites upon integration of SfII
phage (Fig. 4.6). Therefore the attP and attB sites of bacteriophage SfII and *S. flexneri*, respectively, have been localised and are the same as P22.

### 7.6 A two step pathway for serotype conversion

Two genes have been identified, *bgt* and *gtrII*, which are required for the successful glucosylation of *S. flexneri* O-antigen. The predicted functions of the gene products were bactoprenol glucosyl transferase (*Bgt*) which transfers the glucose residue from the activated nucleotide to bactoprenol. The glucose is then transferred by the glucosyl transferase (*GtrII*) onto the O-unit acceptor. This would therefore involve a two-step pathway.

A two step pathway for the glucosylation of O-antigen had previously been independently proposed by two groups.

\[
\text{UDPG} + \text{lipid-P} \leftrightarrow \text{lipid-P-G} + \text{UDP} \quad (1)
\]

\[
\text{Lipid-P-G} + \text{antigen} \rightarrow \text{antigen-G} \quad (2)
\]

One group studying bacteriophage ε³⁴, which adds a glucose residue to the galactose sugar of the repeating unit (factor 34) of *Salmonella anatum* (Wright, 1971) and the other in the study of the *Salmonella* O-antigen factor 12₂ (Nikaido *et al.*, 1971). A lipid linked intermediate was detected in both instances and involved transfer of the glucose residue from the activated nucleotide sugar UDP-glucose to bactoprenol before addition to the LPS acceptor. The lipid intermediate is the C₅₅ polyisoprenol alcohol (bactoprenol) which is also used in O-antigen and peptidoglycan synthesis. However, the linkages that attach the O-unit sugars are pyrophosphate, whereas the linkage attaching the glucose residues are phospho-diester bonds (Wright, 1971).

*Salmonella enterica* sv. Typhimurium can be glucosylated on the galactose sugar via an α-4 linkage resulting in 12₂ specificity, or via an α-6 linkage resulting in factor 1, mediated
by bacteriophage P22 (Wright, 1971). As a result of the studies in this thesis, it is proposed that a lipid intermediate and two step pathway exists for the glucosylation event, resulting in factor 1.

Mutants of bacteriophage ε34 were isolated that fall into two categories: a) those defective in their ability to transfer glucose from UDP to the lipid acceptor and b) those defective in the transfer of glucose from the glucosyl lipid to the O-antigen (Wright and Barzilai, 1970). These can be equated to mutations in bgt and gtr genes respectively. Strains lysogenised by bacteriophage ε34 carrying either of these mutations do not express factor 34, similarly, S. flexneri mutants carrying chromosomal mutations in bgt (RMA903) and gtrII (RMA901) did not express the type II antigen. It is proposed that the first step in the pathway in mediated by the product of the bgt gene, which transfers the glucose sugar from the activated nucleotide onto the bactoprenol, while the gtrII product transfers glucose from the bactoprenol onto the ‘acceptor’ O-antigen-bactoprenol (Fig. 7.2)

7.7 Dpm1- dolichol phosphate mannosyl synthase

Dolichol phosphate mannose synthase Dpm1 of S. cerevisiae is a well characterised member of a group of glycosyl transferases. Dpm1 catalyses the formation of dolichol-phosphate-mannose from dolichol-phosphate and GDP-mannose (Orlean et al., 1988). This protein is encoded in 801 nt, is 30 kDa in size and is predicted to be anchored in the membrane of the endoplasmic reticulum. Membrane spanning regions are located either near the amino and carboxy terminal ends, or if the amino terminal end functions as a signal sequence, then the protein is anchored only by the carboxy terminus and faces the cytoplasm and is oriented toward the lumen (Orlean et al., 1988).

Dpm1 and Bgt both exhibit potential membrane spanning domains at the carboxy terminal end and amino terminally located active site motifs (DXSXD and DXD) (Fig. 5.4),
Figure 7.2 Model for modification events in LPS biosynthetic pathway of *S. flexneri*

Schematic diagram representing the pathway of LPS biosynthesis and the location of proteins involved in O-unit synthesis (RfbBCAD), and assembly (RfbX, RfbF and RfbG), ligation of the first O-unit to the LipidA-core (RfaL), polymerisation of O-units (Rfc), control of O chain length (Rol) and modification (Bgt and GtrII).
Rfc/Rol/RfaL/LipidA-core

Lipid A-core

[\ldots]_n

O-antigen repeat unit

GtrII

Bactoprenol-P-glucose

Bgt

Bactoprenol-PP-sugar

Bactoprenol

Gall

Rfe/RfbF/RfbG

Bactoprenol

UMP

UTP

UDP-glucose + Pi

GalU

glucose-1-PO_4

UDP-N-acetyl glucosamine

TDP-rhamnose

UMP/TDP
which are predicted to be located in the cytoplasmic space where they can then interact with their substrate. In addition, Southern hybridisation data suggests that in both cases more than one copy of the gene or homologues performing similar function exist in the genome of the strain.

Enzymes belonging to this group of transferases are involved in the formation of β-linkages with the sugar they are transferring (Saxena et al., 1995). This has been examined in *Salmonellae* with factor 34 where the glucose is linked to UDPG via an α-linkage and linked to the lipid intermediate via a β-linkage and finally linked to the LPS via an α-linkage. An inversion event between the two steps of the pathway occurs which results in the correct configuration of the branch group (Wright, 1971).

### 7.8 Complementation analyses

Previous studies have identified the *gtrX* gene of bacteriophage SfX, which when introduced into strains of serotype Y, was able to mediate their partial conversion to serotype X. When this gene was introduced into PE577 of serotype Y, no serotype conversion was detected. However, when *bgt* was also provided to the strain using a different plasmid, conversion to X was observed. Similarly, the action of *gtrII* and *gtrV* require the presence of *bgt* (or *orf5*) for the expression of type II and type V antigen, respectively (Huan et al., 1997b). This is consistent with the two-step pathway.

Type II negative mutants RMA901 and RMA903 possessing TnphoA insertions in the *gtrII* and *bgt* genes, respectively, were complemented by the addition of various plasmid constructs (Table 5.4) and lysogenisation with SfII (Table 6.1). Plasmid pRMM272 containing both *bgt* and *gtrII* restored production of the type II antigen in both mutants. RMA901 (*gtrII::TnphoA*) was not complemented by *gtrII* alone but type II antigen was seen
with the introduction of both genes. This implied the possibility of a polarity effect being exerted by TnphoA on the upstream bgt gene.

ORF2 is located upstream of bgt and gtrII and does not appear to be required to mediate expression of type 2 antigen. The lack of requirement of ORF2 in the serotype conversion experiments was evident by the production of type II antigen in the absence of ORF2 (Section 5.3.1). Homologues of ORF2 exist in E. coli K-12 (ECAE000324), S. flexneri (orf10X1) (Morona et al., 1994) and Salmonella bacteriophage P22 (M10893), however, a function has not yet been defined for this open reading frame. This contrasts the observed requirement for this ORF (orf6) by phage SfV and various serotype converting plasmid constructs. A plasmid containing orf4, orf5 but not orf6 of phage SfV which equate to bgt, gtr and orf2 of SfII, when introduced into a S. flexneri Y strain, resulted in the expression of type V and group 3,4 antigens. The lysogens when assessed by Sf6 phage continued to show sensitivity; the authors concluded that the conversion was partial due to the absence of orf6. Our data show that ORF2 is not required for the stable expression of the type 2 antigen and the continued sensitivity to Sf6 phage by SfII lysogens may be due to prophage excision or plasmid instability.

A functional homologues of bgt exists in the E. coli K-12 chromosome, indicated by the ability of gtrX to add the group 7,8 modification in the absence of a plasmid copy of bgt (Table 5.5). The chromosomal gene o306 was sufficient to substitute for bgt in the two step pathway. E. coli K-12 also contains a copy of a glucosyl transferase, encoded by o443, which may be responsible for the addition of a glucose residue to the GlcNAc of the repeat units, resulting in the expression of the type IV antigen (Table 5.5) (Morona et al., 1995).
7.9 Comparison of organisation of genomes

The organisation of genes int, o120, o306 and o443 on the chromosome of E. coli K-12 is similar to that observed for their homologues in both phages SfII and P22. In the case of phage P22 which mediates the addition of a glucose residue to the galactose sugar of the S. enterica sv. Typhimurium repeat units (Lindberg, 1977), the serotype converting region (aI) has not been completely sequenced. The aI gene may be a functional homologue of the glucosyl transferase genes (gtr) of bacteriophages Sfl, Sftr, SfV and SfX. A partially sequenced ORF has been identified and named ORFo1 (Section 5.2.2) which exhibits significant homology with both bgt and o306 implying that further sequencing of this region may identify a homologue to bgt. The organisation of this region of the genome is conserved across the 3 species (Fig. 5.12).

7.10 Bgt and GtrII Proteins

GtrII could not be detected using the L-[35S]-methionine labelled and T7 polymerase/promoter system; this may be due to the number of rare codons used extensively throughout the protein. This compares with the situation observed for the Rfc protein, which despite numerous attempts, has failed to be over-expressed (Daniels et al., 1998). The possibility exists that GtrII is not required in great amounts, due to the specific function it carries out in transferring a glucose residue to RhaIII of the repeat unit. In contrast, Bgt function in pathways other than in LPS biosynthesis and, therefore, would be necessary in greater quantities.

Bgt was predicted to be a membrane bound protein facing the cytoplasm, with two potential membrane spanning domains located at the carboxy terminus. Cellular fractionation was carried out using an E. coli strain in which Bgt had been labelled with
L-[35S]-methionine. Autoradiography revealed Bgt to be located in both the whole membrane and cytoplasmic fractions.

In-frame fusions of truncated and complete forms of Bgt to the reporter genes *phoA* and *lacZ* were constructed. Western immunoblotting of fractionated cells confirmed the localisation of Bgt to the cytoplasmic membrane. Of the 3 fusion proteins only the full length fusion-protein was able to complement *bgt* mutant RMA903, implying that the membrane association plays an important role in the function of the protein, i.e., both membrane spanning domains are essential.

### 7.11 Addition of modification

The localisation of Bgt to the cytoplasmic membrane raises the issue concerning which stage of the LPS biosynthetic pathway the O-antigen becomes modified. In *S. enterica* sv Typhimurium, O-hapten (i.e., long chains of O-antigen not associated with lipid A-core) was found to be glucosylated to the 12₂ form, however semi-rough (*rfc*) mutants of the same strain were found not to contain glucose residues (Nikaido *et al.*, 1971). Further to this, the first 6 repeat units of the O-antigen of strains expressing this O factor do not appear to possess any modification (Helander *et al.*, 1992).

It may be possible that the addition of a glucosyl residue to the repeating units may alter the configuration sufficiently such that they are no longer recognised by the O-antigen polymerase and/or O-antigen ligase. It was postulated that these enzymes function by recognising a specific conformation of the O-unit and any alteration to that conformation, such as the addition of a glucosyl residue, may result in the O-unit no longer being recognised by these enzymes. It was therefore proposed, though not determined, by Wright (1971) and Nikaido *et al.*, (1971) that the terminal repeating unit of the polysaccharide chains in strains expressing factor 34 and 12₂, respectively are not modified.
In contrast to this, rfc mutants of *S. enterica* sv. Typhimurium were able to express factor O5 (Slauch *et al.*, 1996). This also appears to be the case in *S. flexneri*, where rfc mutants still express O-acetylation of the abequose sugar and are also able to accept further modification when added on plasmid copies. Modification of O-units in *S. flexneri* appears to occur at the level of a single repeat unit.

### 7.12 Conclusions

The isolation of bacteriophage SfII has provided information regarding the molecular mechanism by which modifications are carried out by the serotype converting bacteriophages of *S. flexneri*. Some indication to the origin of this bacteriophage has been gained by the preliminary sequencing that has been carried out, however, more information would be obtained by complete sequencing of the genome. It may be that the observed virulence of strains of serotype 2a may be due to genes carried on the serotype converting phage, in a manner similar to the recently isolated CTX phage of *V. cholerae* responsible for the cholera toxin (Waldor and Mekalanos, 1996). The similarities observed between bacteriophages SfII and SfV are of interest also in the elucidation of the evolutionary origins of these bacteriophages.

A two step pathway for the addition of the glucosyl residue to the repeating units of *S. flexneri* has been determined. In addition, the site of action of one of the proteins, Bgt, involved in the modification event provided more information regarding the steps involved in LPS biosynthesis in *S. flexneri*.

Finally, the characterisation of the genes responsible for the different serotypes in *S. flexneri* will provide a tool from which to understand and tackle the issue of developing a vaccine to combat this enteric disease. The identification of an additional gene, bgt, required for the expression of the glucosylations can be incorporated in the construction of serotype
conversion cassettes (SCCs). SCCs will contain the wildtype \textit{attP} site and integrase genes of the \textit{S. flexneri} bacteriophage SfII and introducing \textit{bgt} along with each of the \textit{gtr} genes. In this way, an isogenic set of \textit{S. flexneri} serotypes will be constructed with the serotype conversion genes integrated into the correct site in the host chromosome where they will be stably expressed. This system eliminates the majority of the phage genome, which to date, remain uncharacterised. This isogenic set may then be used as the basis of vaccine strains which could immunise against all serotypes and thus provide complete immunity to \textit{S. flexneri}. 
BIBLIOGRAPHY


Giammanco, G. (1963) Conversion lysogénique de caractères antigéniques de Shigella flexneri (antigène II de type et complexe 7,8 de groupe) Ann. de l’Institut Pasteur 114: 63-76.


p5  Section 1.4 Change to “After passage through the M cell, the organism is engulfed by a macrophage and encapsulated within a phagolysosome, from which the organism escapes shortly after”.

p86. The mutants were isolated in the laboratory of Dr. R. Morona by technical assistant Elizabeth Parker.

Table 2.2. Strain E1196 contains plasmid pRT733.

Figure 3.4. The label S-LPS in this figure should read SR-LPS. The difference observed in migration rates of the first repeat unit of RMM181 compared with the other mutants is due to differences in the amount of sample loaded. More sample has been loaded for RMM181 which has created the appearance of a slower migrating band, however, in the other LPS molecules the migration rates appear identical.

Figure 4.5. The 7.2 kb PstI-3 band is not present in RMM273, this strain behaves in every way as a lysogen. Hybridisation seen with a fragment at approximately the same position as the PstI-3 fragment may be due to excision of bacteriophage from the lysogens during culture, which would result in the regeneration of the PstI-3 fragment.

Table 3.3 details the PstI fragments and the corresponding sizes in kb pairs.

p 83. Summary Section 5.7. The gene products of bgt and gtrII appear to act together to mediate the addition of a glucosyl residue onto the rhamnose III of the repeat unit, resulting in the appearance of the type II antigen.

p 92. One possible reason for there being less Bgt in the inner membrane than in the whole membrane preparation is that protein was lost during the successive steps leading to the isolation of the inner membrane.

Figure 6.3. The LPS of strains RMM334 and RMM468 appear to migrate more slowly due to higher population of LPS molecules present. More sample was loaded in these tracks. It is known in this system that overloading results in upward smearing of bands.

Table 2.1 Virginia omitted from footnote c

p57, line 9 Cells were lysed by a Branson Ultrasonifier (45% cycle, intermittent) and by successive freeze-thawing.

2.20.2 No corrections necessary

p37 Section 2.3 paragraph 3. Liquid cultures were normally grown in 20 ml McCartney bottles.

p48 Section 2.13.1 Sequencing reactions were carried out on 1 µg of double stranded plasmid DNA using the protocol provided by Applied Biosystems.
Examiner #1

p5  Section 1.4  Change to “After passage through the M cell, the organism is engulfed by a macrophage and encapsulated within a phagolysosome, from which the organism escapes shortly after”.

p86.  The mutants were isolated in the laboratory of Dr. R. Morona by technical assistant Elizabeth Parker.

Table 2.2. Strain E1196 contains plasmid pRT733.

Figure 3.4. The label S-LPS in this figure should read SR-LPS. The difference observed in migration rates of the first repeat unit of RMM181 compared with the other mutants is due to differences in the amount of sample loaded. More sample has been loaded for RMM181 which has created the appearance of a slower migrating band, however, in the other LPS molecules the migration rates appear identical.

Figure 4.5. The 7.2 kb PstI-3 band is not present in RMM273, this strain behaves in every way as a lysogen. Hybridisation seen with a fragment at approximately the same position as the PstI-3 fragment may be due to excision of bacteriophage from the lysogens during culture, which would result in the regeneration of the PstI-3 fragment.
Table 3.3 details the PstI fragments and the corresponding sizes in kb pairs.

p 83. Summary Section 5.7. The gene products of bgt and gtrII appear to act together to mediate the addition of a glucosyl residue onto the rhamnose III of the repeat unit, resulting in the appearance of the type II antigen.

p 92. One possible reason for there being less Bgt in the inner membrane than in the whole membrane preparation is that protein was lost during the successive steps leading to the isolation of the inner membrane.

Figure 6.3. The LPS of strains RMM334 and RMM468 appear to migrate more slowly due to higher population of LPS molecules present. More sample was loaded in these tracks. It is known in this system that overloading results in upward smearing of bands.

Table 2.1 Virginia omitted from footnote c

p57, line 9  Cells were lysed by a Branson Ultrasonifier (45% cycle, intermittent) and by successive freeze-thawing.

2.20.2 No corrections necessary

p37 Section 2.3 paragraph 3. Liquid cultures were normally grown in 20 ml McCartney bottles.

p48 Section 2.13.1 Sequencing reactions were carried out on 1 µg of double stranded plasmid DNA using the protocol provided by Applied Biosystems.
p73 remove line break from line 4 after 'resembling'

p76 Section 5.1 paragraph 2; none of which was able to mediate serotype conversion.

p80 Section 5.4 line 1. remove comma after LPS.

p84 line 3 The timing of the modifications is being assessed. The semi-rough mutants were used to observe whether long O chains were required before modifications could be added, this was not found to be the case.

Examiner #2

Figure 3.4 addressed in Examiner #1 comments

Figure 3.6 The genome of SfII is presumably packaged by the headful mechanism due to its homology to P22 phage.

p74 (line 4 from bottom) should be 8.5 kb not 9.2 kb.

Figure 4.7B Weak hybridisation to the 6.5 kb band is seen because the attB site which was detected is located approximately 100 bp from the end of the 3 kb PstI fragment probed. Hence, a low signal is detected using this attB probe.

Chapters 5, 6 and 7- no changes required.