

**CHARACTERISATION OF PHOSPHORYLATION-DEFICIENT
MUTANTS OF ENTEROPATHOGENIC *ESCHERICHIA COLI***

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Richard David Haigh BSc (Newcastle)

Department of Microbiology and Immunology
University of Leicester

January 1999

UMI Number: U532724

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U532724

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

Firstly I must extend my sincere thanks to my supervisor Peter Williams for his patience, encouragement and continued support over the years. I must also thank Terry Lymn and Barry Holland for their insight (?) in giving me a job in Leicester all those years ago, and for introducing me into the entirely unique institution that was the Genetics department.

I acknowledge both Tom Baldwin and Stuart Knutton for the work which formed the basis for this project and for much helpful advice. I also wish to thank the many people whose advice (both solicited and not!) has helped in the explanation of the odd bizarre result which may have cropped up over the last few years, in particular, Karl Wooldridge, Julian Ketley, Martin Goldberg, Primrose Freestone (special cake mention required), Arnoud van Vliet, Anne Wood, Stuart Clarke and Rob Kingsley. I would also like to thank Eli Orr for allowing me the invaluable use of his fluorescence microscope.

I thank all those who have made my time at Leicester more entertaining; this includes the last survivors of the old "Hard core" Julie Morrissey and Karl Wooldridge, those mere youngsters (metaphorically speaking) Jeni Lockett, Steve (MD) Saville, Paul Chilley (and of course Mrs Chilley), Dean Harris, Duncan Baird, and all those members of the Genetics Department, both past and present, who were always willing to go for just that one drink after work. I would also especially like to thank Ruth for support beyond the call of duty with respect to chocolate and biscuit supplies, proof reading (even though it's all prokaryotic stuff and makes no sense anyway) and generally putting up with me over the months while I've been writing - it is finished !

Finally I must thank the lovely people at the Wellcome Trust and all of my ex-lodgers at the (in)famous 12 Hamilton Street for their generous financial support.

In memory of my Mother

Abbreviations.

A ₆₀₀	absorbance at 600 nm
AE	attaching and effacing
AEEC	attaching and effacing <i>E. coli</i>
ATP	adenosine triphosphate
BFP	bundle-forming pilus
bp	basepair
BSA	bovine serum albumin
Ca ²⁺	calcium ions
CT	cholera toxin
Da	daltons
DA	diffuse adherence
DAEC	diffuse-adhering <i>E. coli</i>
dCTP	deoxyribo-cytosine triphosphate
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribo-nucleoside triphosphate
EAggEC	enteroaggregative <i>E. coli</i>
EDTA	diaminoethanetetra-acetic acid
EHEC	enterohaemorrhagic <i>E. coli</i>
EIEC	enteroinvasive <i>E. coli</i>
EPEC	enteropathogenic <i>E. coli</i>
ETEC	enterotoxigenic <i>E. coli</i>
FAS	fluorescent actin staining
kb	kilobase
kDa	kilodalton
LA	localised adherence
LEE	locus of enterocyte effacement

LPS	lipopolysaccharide
LT	heat-labile toxin
mg	milligram
ml	millilitre
nm	nanometre
PBS	phosphate buffered saline
pmole	picomole
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide electrophoresis
SLT	shiga-like toxin
ST	heat stable enterotoxin
Stx	Shiga toxin
STEC	shiga toxin producing <i>E. coli</i>
TAE	tris-acetate-EDTA
TCA	trichloroacetic acid
TE	Tris-EDTA
μCi	microcuries
μg	micrograms
μl	microlitres
v	volume
w	weight
X-gal	5-chloro-4-bromo-3-indoyl- β -D-galactoside
XP	5-chloro-4-bromo-3-indoyl phosphate

Contents

Abstract	1
Chapter 1. General Introduction.	2
1.1 <i>Escherichia coli</i> and disease.	2
1.2 Historic overview of paediatric diarrhoea.	3
1.3 Confirmation of <i>E. coli</i> as the agent of summer diarrhoea	4
1.4 Identification of pathogenically distinct diarrhoeagenic <i>E. coli</i> groups.	6
1.4.1 Separation of EPEC from the other diarrhoeagenic <i>E. coli</i> .	6
1.4.2 Enterohaemorrhagic <i>E. coli</i> (EHEC).	9
1.4.3 Other diarrhoeagenic <i>E. coli</i> groups.	11
1.5 Histopathology of EPEC infection.	13
1.6 Pathogenic mechanism of EPEC.	15
1.6.1 Localised adherence.	15
1.6.2 EPEC induced signal transduction.	17
1.6.3 Intimate adherence.	20
1.6.4 EPEC invasion of epithelial cells.	22
1.6.5 Secretion of EPEC proteins.	23
1.6.6 Locus of enterocyte effacement.	25
1.6.7 Regulation of EPEC virulence genes.	26
1.7 Four stage model of EPEC pathogenesis.	28
1.8 Mechanism of EPEC-induced diarrhoea.	29
1.9 Epidemiology of EPEC disease.	30
1.10 Scope of this project.	32
Chapter 2. Materials and Methods.	33
2.1 Bacterial strains.	33

2.2 Plasmids.	34
2.3 Bacterial culture media.	35
2.4 Growth, storage and preservation of bacterial strains.	36
2.5 EPEC interactions with cultured human epithelial cell lines.	36
2.5.1 Maintenance of tissue culture cells.	36
2.5.2 HEp-2 adhesion assay.	38
2.5.3 Immunofluorescence microscopy.	38
2.5.3.1 Fluorescent actin staining (FAS) assay.	38
2.5.3.2 Detection of bundle-forming pili using fluorescent antibodies.	39
2.5.4 EPEC induced protein phosphorylation in cultured human epithelial cells.	39
2.5.5 Gentamicin protection (invasion) assay.	40
2.6 Autoaggregation assay.	41
2.7 ³⁵ S labelling of EPEC secreted proteins (EspA, EspB and EspC).	41
2.8 Conjugation.	42
2.9 DNA manipulations	42
2.9.1 Stock solutions for DNA work.	42
2.9.2 Restriction enzyme digestions and ligation of DNA.	43
2.9.3 Modification of DNA fragment termini generated by restriction enzyme digestion.	43
2.9.3.1 Filling in of 3' overhangs.	43
2.9.3.2 Removal of 5' overhangs.	43
2.9.4 Electrophoresis of DNA.	43
2.9.5 Isolation of DNA fragments from agarose gels using Polymer wool™.	44
2.9.6 DNA preparation.	44
2.9.6.1 Small scale plasmid DNA preparation (mini-prep).	44
2.9.6.2 Large scale CsCl preparation of plasmid DNA.	45
2.9.6.3 Small scale preparation of chromosomal DNA.	46

2.9.6.4 Large scale CsCl preparation of chromosomal DNA.	46
2.9.7 Transformation of <i>E. coli</i> with plasmid DNA.	47
2.9.7.1 Calcium chloride method.	47
2.9.7.2 Electroporation.	48
2.10 DNA labelling and detection.	48
2.10.1 Transfer of DNA from agarose gels to nylon membranes (Southern blot).	48
2.10.2 Non-radioactive detection of DNA.	49
2.10.2.1 Labelling of DNA probes.	49
2.10.2.2 Hybridisation of probes to Southern blots.	50
2.10.2.3 Detection of hybridised probe.	50
2.11 Screening of a multicopy plasmid library of EPEC chromosomal DNA.	51
2.11.1 Colony blotting.	51
2.11.2 Screening colony blots using $\alpha^{32}\text{P}$ -dCTP labelled probes.	52
2.11.2.1 Labelling of probes with $\alpha^{32}\text{P}$ -dCTP.	52
2.11.2.2 Hybridisation of $\alpha^{32}\text{P}$ -labelled probes to colony blots.	53
2.12 Oligonucleotide primers.	54
2.13 Polymerase chain reaction (PCR).	54
2.14 DNA sequencing.	55
2.14.1 Double stranded DNA sequencing from plasmids.	55
2.14.2 DNA sequencing from PCR products.	56
2.15 Computer analysis of DNA and protein sequences	56
2.16 Analysis of proteins.	57
2.16.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE).	57
2.16.2 Preparation of total protein extracts from <i>E. coli</i>	58
2.16.3 Preparation of outer membrane proteins from <i>E. coli</i> .	58
2.16.4 Transfer of proteins from SDS-PAGE gels to nitrocellulose membranes (Western blotting).	59
2.16.5 Antibody detection of bundle-forming pili proteins on	

Western blots.	59
2.17 Alkaline phosphatase assay.	60
Chapter 3. Isolation and characterisation of EPEC mutants deficient in induction of protein phosphorylation in cultured epithelial cells.	61
3.1 Introduction.	61
3.2 Results.	65
3.2.1 Construction of a bank of <i>TnphoA</i> mutants in strain CVD206.	65
3.2.1.1 Introduction of <i>F42ts114 lac zzf-1::TnphoA</i> into CVD206.	66
3.2.1.2 Isolation of <i>phoA</i> -positive <i>TnphoA</i> mutants from CVD206 <i>F42ts114 lac zzf-1::TnphoA</i> .	66
3.2.2 Screening of the CVD206:: <i>TnphoA</i> bank for mutants unable to induce protein phosphorylation in tissue culture cells.	68
3.2.3 Phenotypic analysis of the CVD206:: <i>TnphoA</i> mutants.	69
3.2.3.1 Autoaggregation assay.	69
3.2.3.2 Secretion of EPEC virulence proteins.	70
3.2.3.3 Transformation of the CVD206:: <i>TnphoA</i> mutants with pCVD438.	70
3.2.3.4 Localised adherence (LA) and the Fluorescent Actin Staining (FAS) assay.	71
3.2.3.5 Comparison of the invasive ability of the phosphorylation deficient CVD206:: <i>TnphoA</i> mutants.	71
3.2.3.6 Expression of bundle-forming pili (BFP).	72
3.2.4 Identification of multiple <i>TnphoA</i> insertions in the phosphorylation-deficient CVD206:: <i>TnphoA</i> mutants using Southern blot analysis.	72
3.2.4.1 Production of a <i>TnphoA</i> specific DNA probe by PCR.	72
3.2.4.2 Southern blot hybridisation analysis of the CVD206:: <i>TnphoA</i> mutants using a <i>TnphoA</i> specific probe.	73
3.2.5 Cloning and analysis of the DNA sequences flanking the	

CVD206:: <i>TnphoA</i> <i>phoA</i> -positive insertions.	74
3.2.5.1 Cloning of the regions upstream of the CVD206:: <i>TnphoA</i> insertions.	74
3.2.5.2 Restriction enzyme digestion analysis of <i>TnphoA</i> clones.	75
3.2.6 Analysis of DNA sequences flanking the cloned <i>TnphoA</i> insertions.	75
3.2.6.1 Analysis of DNA sequences from plasmids p13a, p137c and p181c.	76
3.2.6.2 Identification of open reading frames adjacent to the site of <i>TnphoA</i> insertion in mutants #13/#137 and #181.	77
3.2.6.3 Analysis of DNA sequence from plasmids p55a and p138b.	77
3.2.6.4 Identification of open reading frames adjacent to the <i>TnphoA</i> insertions in mutants #138 and #55.	78
3.2.7 Cloning the wildtype loci identified by the CVD206:: <i>TnphoA</i> mutants from EPEC strain E2348-69.	78
3.2.7.1 Design of locus-specific DNA probes.	79
3.2.7.2 Southern blot analysis of E2348-69 using probes 55a and 181c.	79
3.2.7.3 Construction of an E2348-69 <i>Bam</i> HI library in pBluescript.	80
3.2.7.4 Screening of the E2348-69 <i>Bam</i> HI library using probes 55a and 181c.	80
3.2.8 Further analysis of plasmids pRDH7 and pRDH8.	81
3.2.8.1 Restriction enzyme mapping of pRDH7 and pRDH8.	81
3.2.8.2 Southern blot analysis of EPEC strains using probes derived from pRDH7 and pRDH8.	81
3.2.9 Attempts to complement the mutations in the CVD206:: <i>TnphoA</i> mutants by introduction of plasmids pRDH7 or pRDH8.	82
3.3 Discussion	83
3.3.1 Problems encountered in creating a <i>TnphoA</i> library in EPEC.	83

3.3.2 A role for bundle-forming pili (BFP) in the induction of host cell protein phosphorylation ?	85
3.3.3 Future work.	88

Chapter 4. DNA sequencing and characterisation of the genes present in plasmids pRDH7 and pRDH8. 89

4.1 Introduction.	89
4.2 Results.	91
4.2.1 DNA sequencing of pRDH7 and pRDH8.	91
4.2.1.1 Construction of <i>TnphoA</i> insertions in plasmids pRDH7 and pRDH8.	91
4.2.1.2 DNA sequencing of <i>phoA</i> positive <i>TnphoA</i> insertions in pRDH7 and pRDH8.	92
4.2.1.3 Identification of <i>bfpA</i> in pRDH7.	92
4.2.1.4 DNA sequence analysis of <i>phoA</i> -positive <i>TnphoA</i> insertions in pRDH8.	93
4.2.1.5 Mapping of <i>TnphoA</i> insertions in pRDH7 and pRDH8.	94
4.2.1.6 The complete DNA sequence of pRDH7.	95
4.2.1.7 Completion of the DNA sequencing of pRDH8 using primer walking.	95
4.2.2 Analysis of the genes encoded in the <i>Bam</i> HI insert of pRDH7.	96
4.2.2.1 Identification of open reading frames in pRDH7	96
4.2.2.2 <i>bfpA</i> , <i>orf1</i> , <i>orf2</i> , <i>orf3</i> , <i>orf4</i> , and <i>orf5'</i> are organised as an operon.	97
4.2.2.3 <i>orf1</i> and <i>orf2</i> encode lipoproteins.	97
4.2.2.4 <i>orf3</i> and <i>orf4</i> .	98
4.2.2.5 <i>orf5'</i> encodes a member of the PulE/OutE/ExeE/XpsE /XcpR family of transport proteins.	98
4.2.2.6 Comparison of pRDH7 sequences to the <i>bfp</i> operon.	99

4.2.3 Analysis of the genes encoded by the <i>Bam</i> HI insert of pRDH8.	100
4.2.3.1 Identification of open reading frames in pRDH8.	100
4.2.3.2 <i>emtA</i> encodes a member of the lytic transglycosylase family.	101
4.2.3.3 The open reading frame <i>tag</i> encodes an integral inner membrane protein.	103
4.2.3.4 Identification of an <i>orf1'-orf2-ompX</i> operon.	104
4.2.3.5 <i>ompX</i> encodes a TonB-dependent outer membrane receptor protein homolog.	104
4.2.3.6 <i>orf1'</i> and <i>orf2</i> have homologues in <i>Haemophilus influenzae</i> which are located within operons of iron uptake gene homologues.	106
4.2.3.7 <i>orf3</i> and <i>orf4</i> .	107
4.2.3.8 Identification of the OmpX and EmtA proteins in outer membrane preparations.	108
4.2.4 The distribution of the <i>orf1'-orf2-ompX</i> operon in <i>E. coli</i> .	108
4.2.4.1 Comparison of pRDH8 sequence with the <i>E. coli</i> K-12 genome sequence.	108
4.2.4.2 Southern hybridisation analysis of the distribution of <i>orf1'-orf2-ompX</i> sequence in <i>E. coli</i> isolates.	109
4.2.5 Cloning and analysis of the region upstream of <i>ompX</i> which is deleted in <i>E. coli</i> K-12.	110
4.2.5.1 Cloning of <i>ompX</i> and upstream regions from CVD206:: <i>TnphoA</i> mutant #181.	111
4.2.5.2 Partial DNA sequencing and mapping of pRDH45.	112
4.2.5.3 Identification of the N-terminus of the Orf1 polypeptide and a further upstream gene in the operon.	113
4.3 Discussion.	114
4.3.1 pRDH7 encodes the proximal region of an operon for BFP biogenesis.	114

4.3.2 Analysis of the genes encoded on plasmid pRDH8.	115
4.3.2.1 A possible role for the lytic transglucosylase EmtA in transenvelope transport.	115
4.3.2.2 A novel siderophore uptake system in <i>E. coli</i> .	119
4.3.2.3 The distribution of the <i>orf1'-orf2-ompX</i> operon amongst pathogenic <i>E. coli</i> .	123
4.3.2.4 Future work: Analysis of the frequency of deletion of the <i>ompX</i> locus.	128
Chapter 5. Construction of non-polar mutants of the <i>emtA</i> and <i>ompX</i> genes in EPEC strains.	130
5.1 Introduction.	130
5.1.1 General introduction.	130
5.1.2 Construction of non-polar mutations.	132
5.1.3 Introduction of non-selectable mutations into <i>E. coli</i> using counters selectable markers.	133
5.2 Results	135
5.2.1 Strategy for the construction of <i>ompX</i> mutations.	135
5.2.1.1 Subcloning of a pRDH8 <i>Cla</i> I fragment containing the 3' end of <i>ompX</i> into pBluescript to generate pRDH35.	136
5.2.1.2 Construction of a frameshift mutation in <i>ompX</i> by the deletion of the unique <i>Sph</i> I site from pRDH35.	136
5.2.1.3 Cloning the <i>ompX Sph</i> I deletion mutation into delivery vector pRDH10.	137
5.2.1.4 Subcloning the <i>ompX/TnphoA</i> fragment from pRDH8-12 into pRDH10.	137
5.2.2 Introduction of the <i>ompX</i> mutations into strain CVD206.	139
5.2.2.1 Generation of merodiploids containing the <i>ompX</i> mutations in CVD206.	139

5.2.2.2 Isolation of resolved merodiploids using sucrose selection.	140
5.2.2.3 Confirmation of the CVD206 <i>ompX</i> mutants by Southern blotting analysis.	142
5.2.3 Analysis of the phenotypes of the CVD206 <i>ompX</i> mutants.	143
5.2.3.1 The <i>ompX</i> frameshift mutant RDH003 is wild-type for growth in low iron.	143
5.2.3.2 <i>ompX/phoA</i> expression in RDH004 is not induced by low iron concentrations.	143
5.2.4 Strategy for the construction of an <i>emtA</i> frameshift mutation.	144
5.2.4.1 Cloning a subgenic PCR fragment of <i>emtA</i> into pBluescript.	145
5.2.4.2 Construction of a frameshift mutation in <i>emtA</i> by deletion of the unique <i>HindIII</i> site in pRDH32.	146
5.2.4.3 Cloning of the <i>emtA</i> Δ <i>HindIII</i> fragment into pRDH10 to create recombinant plasmid pRDH33.	147
5.2.5 Introduction of an <i>emtA</i> frameshift mutation into E2348-69 Strep ^r by allelic exchange.	148
5.2.5.1 Generation of EPEC merodiploids containing an <i>emtA</i> frameshift mutation.	148
5.2.5.2 Isolation of resolved merodiploids using sucrose selection.	149
5.2.5.3 Confirmation of the <i>emtA</i> Δ <i>HindIII</i> mutation by Southern blot analysis and DNA sequencing.	150
5.2.6 Phenotypic analysis of EPEC strain RDH005.	151
5.2.6.1 Altered time course of BfpA expression in RDH005.	151
5.2.7 Investigation of a proposed dominant-negative phenotype for the <i>emtA/TnphoA</i> mutants CVD206:: <i>TnphoA</i> #13 and #137.	152
5.2.8 Phenotypic analysis of the IAP ⁺ strain RDH006.	153
5.3 Discussion.	155
5.3.1 A defined non-polar <i>emtA</i> mutant produces BFP.	158
5.3.2 Future work on the <i>emtA</i> locus.	161

5.3.3 Regulation of the <i>ompX</i> operon.	163
5.3.4 Future work: Identification of a siderophore for the OmpX receptor.	167
Chapter 6. General Discussion.	168
6.1 EPEC MLC phosphorylation-negative mutants.	169
6.2 Is there a role for OmpX in <i>E. coli</i> virulence?	170
6.3 Concluding remarks.	173
Appendix 1. The annotated sequence of pRDH7.	175
Appendix 2. The annotated sequence of pRDH8.	184
Appendix 3. Construction of pRDH10.	194
Literature cited.	195

Figures and Tables.

Chapter 1.

- Figure 1.1 Ultrastructure of EPEC infected human gut tissue.
- Figure 1.2 Schematic representation of the genes involved in EPEC pathogenesis.
- Figure 1.3 Four stage model of EPEC pathogenesis.
- Figure 1.4 Incidence of gastroenteritis diagnosed as *E. coli* in children under 5 years of age.

Chapter 3.

- Figure 3.1 Protein phosphorylation profile of EPEC infected HEp-2 cells.
- Figure 3.2 Diagrammatic representation of *TnphoA*.
- Figure 3.3 Secondary screening of CVD206::*TnphoA* mutants for induction of protein phosphorylation in human epithelial cultured cells.
- Figure 3.4 Growth curves of phosphorylation-deficient CVD206::*TnphoA* mutants.
- Figure 3.5 Autoaggregation phenotypes of the CVD206::*TnphoA* mutants.
- Figure 3.6 Secretion of EPEC virulence proteins.
- Figure 3.7 Actin rearrangement by the host cell phosphorylation deficient CVD206::*TnphoA* mutants.
- Figure 3.8 Invasive ability of phosphorylation-deficient CVD206::*TnphoA* mutants.
- Figure 3.9 Analysis of bundle-forming pili (BFP) production in phosphorylation-deficient CVD206::*TnphoA* mutants.
- Figure 3.10 Southern blot analysis of CVD206::*TnphoA* mutants using a *TnphoA* specific probe.
- Figure 3.11 Restriction enzyme digestion analysis of cloned *TnphoA* insertions.
- Figure 3.12 Diagrammatic representation of the *Bam*HI inserts of the *phoA*-positive pUC18 clones from the CVD206::*TnphoA* mutants.

- Figure 3.13 Identification of open reading frames in-frame with *phoA* in the CVD206::Tn*phoA* mutants #13, #137 and #181.
- Figure 3.14 Identification of open reading frames in-frame with *phoA* in the CVD206::Tn*phoA* mutants #55 and #138.
- Figure 3.15 Southern blot analysis of E2348-69 and DH5 α chromosomal DNA using probes specific to the CVD206::Tn*phoA* mutant loci.
- Figure 3.16 Map of plasmid pRDH7.
- Figure 3.17 Map of plasmid pRDH8.
- Figure 3.18 Southern blot analysis of EPEC strains using pRDH7 and pRDH8 specific probes.
- Figure 3.19 Protein phosphorylation profiles of Caco-2 cells infected with CVD206::Tn*phoA* mutants complemented with pRDH7 or pRDH8.

Chapter 4.

- Figure 4.1 Diagrammatic representation of the Tn*phoA* insertions used for transposon sequencing of plasmids pRDH7 and pRDH8.
- Figure 4.2 Nucleotide sequence alignment between pRDH8 sequences and the *E. coli* chromosomal sequence umudca.
- Figure 4.3 Summary of the strategy for mapping the sites of Tn*phoA* insertions in plasmids pRDH7 and pRDH8.
- Figure 4.4 Diagrammatic representation of the open reading frames identified in plasmid pRDH7.
- Table 4.1 Physical characteristics of the gene products of pRDH7.
- Figure 4.5 Amino acid alignment of pRDH7 Orf2 (BfpB) with MshD of *Vibrio cholerae* and PilN of *E. coli*.
- Figure 4.6 Amino acid alignment of pRDH7 Orf5 (BfpD) with TcpT of *Vibrio cholerae*.
- Figure 4.7 Diagrammatic representation of the open reading frames identified in pRDH8.

- Table 4.2** Physical characteristics of the gene products of pRDH8.
- Figure 4.8** Amino acid alignment of EmtA with MltC of *E. coli* and y761 of *Haemophilus influenzae*.
- Figure 4.9** The reaction catalysed by the lytic transglycosylases of *E. coli*.
- Figure 4.10A** Predicted transmembrane regions of the Tag protein.
- Figure 4.10B** Amino acid alignment of the Tag polypeptide with f82.
- Table 4.3** Outer membrane receptor proteins which show significant homology with OmpX.
- Figure 4.11** Consensus amino acid sequences found in known and predicted TonB-dependent receptors of *E. coli* and OmpX.
- Figure 4.12** Dendrogram representing the phylogenetic analysis of the nine known *E. coli* TonB-dependent outer membrane receptors and OmpX and f700.
- Figure 4.13A** Amino acid alignment of pRDH8 Orf1' and yc73 of *Haemophilus influenzae*.
- Figure 4.13B** Amino acid alignment of pRDH8 Orf2 with ModD of *Rhodobacter capsulatus* and ModD of *Haemophilus influenzae*.
- Figure 4.14** Amino acid alignment of pRDH8 Orf4 with MccF of *E. coli*.
- Figure 4.15** Identification of EmtA and OmpX in *E. coli* outer membranes.
- Figure 4.16A** Diagrammatic representation of the *emtA* loci in EPEC strain E2348-69 and *E. coli* K-12.
- Figure 4.16B** Alignment of the pRDH8 DNA sequence with the *E. coli* K-12 26.82 min chromosomal DNA sequence.
- Table 4.4** *E. coli* strains used to screen for *ompX*.
- Figure 4.17** Southern blot analysis of multiple *E. coli* virotypes using an *ompX* specific probe.
- Figure 4.18A** Restriction enzyme mapping of plasmid pRDH45.
- Figure 4.18B** Plasmid map of pRDH45.
- Figure 4.19** Identification of open reading frames in pRDH45.
- Figure 4.20A** Amino acid alignment of pRDH45/pRDH8 Orf1 with yc73 of *Haemophilus influenzae*.

Figure 4.20B Amino acid alignment of pRDH45 Orf5' with yc72 of *Haemophilus influenzae*.

Figure 4.21 General model of TonB-dependent ferrisiderophore uptake in *Escherichia coli*.

Figure 4.22 The major *E. coli* gene clusters involved in high affinity iron uptake.

Figure 4.23 The distribution of the LEE insertion site, the *shu* locus and the *ompX* locus amongst a series of phylogenetically aligned diarrhoeagenic *E. coli* clones (DEC).

Chapter 5.

Figure 5.1 Summary of non-polar mutant construction using a positive-selection allelic exchange vector.

Figure 5.2 Restriction enzyme map of the positive-selection allelic exchange suicide vector pRDH10.

Figure 5.3 Creation of a frameshift mutant in *ompX* by deletion of an *SphI* site.

Figure 5.4 Summary of the construction of pRDH38; a positive-selection suicide vector containing an *ompX* frameshift mutation.

Figure 5.5 Summary of the construction of pRDH39; a positive-selection suicide vector encoding an OmpX/PhoA fusion protein.

Figure 5.6 Increased autoaggregation phenotype (IAP) in EPEC.

Figure 5.7 Confirmation of the *ompX* mutations in RDH003 and RDH004 by Southern blot analysis.

Figure 5.8 Growth of strain RDH003 (*ompX*) in iron limiting conditions.

Figure 5.9 Investigation of *ompX* expression levels under iron limiting conditions.

Figure 5.10 Design of primers used in the PCR amplification of a subgenic fragment of the *emtA* gene and the creation of a frameshift mutation.

Figure 5.11 Summary of the construction of pRDH33; a positive-selection suicide vector containing an *emtA* frameshift mutation.

- Figure 5.12 Confirmation of the *emtA* frameshift mutation in RDH005 by Southern blot analysis.
- Figure 5.13 Localised adherence and AE lesion formation by RDH005 (*emtA*⁻).
- Figure 5.14 Characterisation of the adherence and invasion of RDH005 (*emtA*⁻) in HEp-2 cells.
- Figure 5.15 Production of BFP by RDH005 (*emtA*⁻).
- Figure 5.16 Western blot analysis of BfpA expression in RDH005 (*emtA*⁻).
- Figure 5.17 Production of filamenting cells in EPEC strains carrying pRDH8 *emtA*::Tn*phoA* plasmids.
- Figure 5.18 Growth characteristics of EPEC strain RDH006 (IAP⁺).
- Table 5.1 Autoaggregation phenotypes of E2348-69 IAP⁺ strains.

Abstract

Characterisation of Phosphorylation-Deficient Mutants of Enteropathogenic *Escherichia coli*.

Richard D. Haigh

Enteropathogenic *E. coli* (EPEC) is a major cause of chronic infantile diarrhoea which results in high levels of morbidity and mortality in the developing world. EPEC infection of cultured epithelial cells causes rearrangements of the cytoskeleton resulting in the effacement of microvilli and the formation of characteristic lesions involving the accretion of actin at the site of bacterial attachment. Attaching and effacing (AE) lesion formation involves the subversion of host cell signalling systems including tyrosine and serine/threonine phosphorylation of proteins, activation of phospholipase C- γ 1, and increases in the second messengers IP₃ and Ca²⁺.

In this study the transposon *TnphoA* was used to isolate EPEC mutants which were unable to induce the serine/threonine phosphorylation of host proteins. The five phosphorylation-negative mutants identified were all found additionally to be deficient in the assembly of the EPEC plasmid-encoded type IV bundle-forming pilus (BFP). Two of the mutations were localised to a gene encoding an outer membrane protein component of BFP. The other mutations defined two chromosomal genes *emtA* and *tag* which encode a novel endo-specific lytic transglycosylase and an uncharacterised inner membrane protein.

Based upon its known enzymatic activity it was proposed that EmtA was required to facilitate the assembly of BFP through the peptidoglycan cell wall. Construction of a non-polar mutation within the *emtA* gene of EPEC strain E2348-69 demonstrated only minor effects upon BFP biosynthesis. It is proposed that the phosphorylation-negative/BFP-negative phenotype observed in the original EPEC *emtA* mutant is due to polar effects of *TnphoA* insertion upon the expression of one of the adjacent uncharacterised genes.

Determination of the DNA sequence adjacent to the EPEC *emtA* locus identified a region which appears to have been deleted from *E. coli* K-12 strains. This locus contains genes encoding the TonB-dependent outer membrane receptor (OmpX) and inner membrane components of a novel siderophore uptake system. Analysis of the distribution of *ompX* amongst pathogenic *E. coli* strains demonstrated inconsistencies with the expected phylogeny.

Chapter 1. General Introduction

The work presented in this thesis concerns genetic analysis of the diarrhoeal pathogen enteropathogenic *Escherichia coli* (EPEC), which remains an important cause of morbidity and mortality worldwide. This introduction gives a brief historical overview of the identification of *E. coli* as an agent responsible for infantile diarrhoea, details the determination and characterisation of the different diarrhoeagenic *E. coli* virotypes and presents the current state of knowledge concerning both the pathogenic mechanism and the virulence determinants identified in EPEC.

1.1 *Escherichia coli* and disease.

The genus *Escherichia* is a typical member of the Enterobacteriaceae which lives in the bowel of humans and animals. *Escherichia coli* is a short straight Gram-negative rod, usually motile with peritrichous flagella, and often fimbriate. *E. coli* was originally described by Theodore Escherich in 1885 who isolated it as a natural component of the human intestinal flora and called it *Bacterium coli commune* (Escherich, 1885; 1886). With the long standing use of *E. coli* K-12 (Bachmann, 1996) as a model for the analysis of biochemical pathways (Tatum, 1945; Tatum and Lederberg, 1947) and bacterial genetics (Clowe and Hayes, 1968), and the recent genome sequencing (Blattner *et al.*, 1997; Berlyn, 1998; Rudd, 1998), *E. coli* is now probably the best studied organism on Earth. Whilst Escherich (1886) demonstrated that *E. coli* was pathogenic upon injection into rabbits he remained uncommitted as to its virulence in humans. Since then *E. coli* has been confirmed unequivocally as a cause of many diseases in both humans and domestic animals; it is now recognised as an important cause of urinary tract infection (UTI), wound infection, peritonitis, meningitis, pneumonia, septicaemia and gastrointestinal infections (Levine, 1985; Sussman, 1997; Wray and Woodward, 1997). Diseases due to *E. coli* can be classified into two groups; *specific* and *non-specific*. *Specific* infections involve the colonisation of mucosal surfaces, such as the urinary tract or the intestine, and the disease

symptoms produced are primarily limited to these regions. *Non-specific* infections are those not related to a site of mucosal colonisation and are due to events such as contamination of wounds or the peritoneal space during surgery, or result from the spread of bacteria from a specific infection, e.g. septicaemia arising from an urinary tract infection (Sussman, 1997).

1.2 Historic overview of paediatric diarrhoea.

Historically diarrhoea has been well documented as a significant cause of morbidity and mortality during early childhood (Creighton, 1975). Even today it is still estimated that 1.3 thousand million diarrhoeal episodes (at an average of 3.3 episodes per child) and 4 million deaths occur each year worldwide in the under-fives; furthermore, 80% of the deaths due to diarrhoea occur in the first two years of life (World Health Organisation, 1997; 1998). Whilst diarrhoea-associated childhood mortality is now primarily a problem in developing countries (Levine and Edelmann, 1984; Donnenberg, 1995), it is apparent from the historical records in the 18th century that a similar pattern of disease once existed in the developed nations, probably as a result of the massive increase in urbanisation which occurred during the industrial revolution (Creighton, 1975; Robins-Browne, 1987). The first systematic medical records were provided by the American Benjamin Rush (Rush, 1789) who noted the association of warm weather in the summer months with both the frequency and danger of childhood diarrhoea in the United States. Rush termed the condition "cholera infantum", though later studies in the 19th century, which also noted the seasonality of paediatric diarrhoea, used the euphemisms "disease of the season", "summer complaint" or "summer diarrhoea" (Topley and Wilson, 1936; Creighton, 1975; Robins-Browne, 1987). Whilst summer diarrhoea and speculation as to its cause were still accorded much importance into the start of the 20th century (Topley and Wilson, 1936), it is apparent from reports in both the United States and Great Britain that the diarrhoeal mortality rates amongst infants declined significantly during the 1920s and 1930s (Arnold, 1927; Smellie, 1939). It also became apparent that the seasonal nature of the disease had changed; the incidence of outbreaks was more evenly spread throughout

the year and the characteristic summer peak was no longer observed (Smellie, 1939). Despite the overall fall in infant mortality rates, medical interest in infantile diarrhoea was maintained during this period due to several reports concerning severe outbreaks of nosocomial neonatal enteritis (Barenberg *et al.*, 1936; Best, 1938; Greenberg and Wrocker, 1938). Though nosocomial outbreaks of infantile diarrhoea had been previously observed (Sandilands, 1909), these new reports provided thorough documentation of the phenomenon; these data included a study of 5,082 live-born babies involved in a series of outbreaks in New York City in which a morbidity rate of 14.8% and 356 deaths (mortality rate/case, 47.5%) were recorded (Best, 1938). Though the phenomenon of summer diarrhoea is no longer observed in developed countries such as the United States or Great Britain it still exists and remains a major cause of morbidity and mortality in temperate and sub-tropical developing countries (Edelman and Levine, 1983; Robins-Browne *et al.*, 1980).

1.3 Confirmation of *E. coli* as the agent of summer diarrhoea

Bray is rightly credited for the first demonstration of convincing evidence that *E. coli* was the cause of summer diarrhoea (Bray, 1945); however, several other studies preceded, and to some extent anticipated, his landmark publication. Lesage (1897) had shown that serum of patients convalescing from diarrhoea were able to agglutinate strains of *E. coli* isolated from other patients in the same epidemic; in contrast, *E. coli* strains from healthy people did not react. From these data Lesage postulated that there might be both pathogenic and non-pathogenic strains of *E. coli* (Lesage, 1987). Furthermore, between 1912 and 1935 a number of groups were to identify *E. coli* as an agent implicated in cases of gastroenteritis, and more specifically in infantile diarrhoea (Bahr, 1912; Adam, 1923; Davison, 1925; Adam, 1927; Goldshmidt, 1933; Dulaney and Michelson, 1935). At the same time it had been proposed that *E. coli* was a cause of diarrhoea in animals (Sojka, 1965); in support of this hypothesis was a series of experiments, conducted over many years, in which it was demonstrated unequivocally that a subgroup of *E. coli* strains could cause diarrhoea (scours) when fed to calves (Jensen, 1892; Christiansen, 1917).

The novel and convincing aspect of Bray's work, compared with these earlier data, was his demonstration that the *E. coli* strains which he had isolated from cases during an epidemic of summer diarrhoea were antigenically homogenous (Bray, 1945; Bray and Bevan, 1948). Bray's data were rapidly confirmed by Giles *et al.* (1949) and Taylor *et al.* (1949) who also isolated serologically distinct *E. coli* strains from the majority of the infants who were involved in a series of outbreaks of diarrhoea in Aberdeen and London, respectively. These studies were to be greatly assisted by the recently developed *E. coli* serotyping scheme (Kauffman, 1947). Kauffman's scheme, which is still the standard used today, divides *E. coli* initially into serogroups on the basis of their somatic (O) antigens and then further divides these to give individual serotypes using flagellar (H) or capsular (K) antigens. By 1950 Kauffman and DuPont (1949) had serotyped the strains from the outbreaks in Scotland and England and confirmed that these, along with strains from outbreaks in the United States, Denmark and the Netherlands, were all of serogroup O55 or a new serogroup O111. Within the next few years strains within the O55 and O111 serogroups were epidemiologically incriminated in outbreaks of infantile diarrhoea in numerous countries around the world (Smith, 1955; Herweg *et al.*, 1956; Levine and Edelman, 1983). Whilst the epidemiological data were compelling, the failure to identify characteristic histopathological lesions during autopsies (Neter, 1959), and the identification of strains of similar serotypes in healthy individuals (Payne and Cook, 1950; Cathie and MacFarlane, 1951; Shanks and Studzinski, 1952, Stevenson, 1952) resulted in some investigators being reluctant to accept these strains as pathogens. They proposed instead that the diarrhoea resulted from a different aetiology, and that the presence of the O55 or O111 strains was a secondary phenomenon which resulted from the overgrowth of specific *E. coli* serotypes due to the altered conditions in the intestine.

The doubts about the pathogenicity of *E. coli* O111 strains were resolved by a series of studies in which diarrhoea was induced by feeding 10^8 - 10^{10} O111 bacteria to a two-month-old child (Neter and Shumway, 1950) or adult volunteers (Kirby *et al.*, 1950;

Ferguson and June, 1952). Similar experiments were performed to prove the virulence of *E. coli* O55 strains (June *et al.*, 1953; Koya *et al.*, 1954a) and an *E. coli* O127 strain (Wentworth *et al.*, 1955). By contrast control subjects fed *E. coli* isolated from healthy individuals did not produce diarrhoeal symptoms (Koya *et al.*, 1954a; Koya *et al.*, 1954b). Concurrent with the volunteer studies reports were being received from many countries epidemiologically implicating additional *E. coli* serotypes as the cause of pediatric diarrhoea. By 1957 13 serogroups had been identified; these were O18a, O18b, O26, O44, O55, O86, O111, O114, O119, O125, O126, O127 and O128 (original references are given in Robins-Browne, 1987). Since that time only two further serogroups, O142 and O158, have been added to this list (Ørskov *et al.*; 1960; Rowe *et al.*, 1974; Rowe, 1979). This group of *E. coli* serotypes associated with infantile diarrhoea were finally given an identity when Neter *et al.* (1955) designated them "enteropathogenic" *E. coli* (EPEC).

1.4 Identification of pathogenically distinct diarrhoeagenic *E. coli* groups.

1.4.1 Separation of EPEC from the other diarrhoeagenic *E. coli*.

With the identification of diarrhoeagenic *E. coli* virulence determinants in the late 1960s and early 1970s it became apparent that there were several distinct groups of strains with different pathogenic mechanisms. A group of *E. coli* strains were identified which could cause intestinal hypersecretion by elaboration of enterotoxins (Gorbach *et al.*, 1971). These strains, termed enterotoxigenic *E. coli* (ETEC), were found to produce two enterotoxin types (Sack., 1980), the heat-labile toxin (LT) and/or two heat stable toxins (STa and STb) (reviewed recently in O'Brien and Holmes, 1996). LT is structurally and functionally similar to the *Vibrio cholerae* enterotoxin (CT) (Sixma *et al.*, 1993) and comprises a single 28 kDa A subunit and a pentamer ring of 11.5 kDa B subunits (Streatfield *et al.*, 1992). The B subunits bind to the ganglioside GM₁ (Teneberg *et al.*, 1994) and the A subunit is translocated through the cell to its target on the basolateral surface of the intestinal epithelial cells (Lencer *et al.*, 1995). The A subunit then ADP-ribosylates the GTP-binding protein G_s thereby

resulting in permanent activation of adenylate cyclase. The resultant increase in cAMP levels causes activation of cAMP-dependent protein kinases and the supranormal phosphorylation and activation of chloride channels in the apical cell membrane; the principal target is the cystic fibrosis transmembrane conductance regulator protein (CFTR). The increased ion content of the gut lumen then induces paracellular transport of water resulting in osmotic diarrhoea (Sears and Kaper, 1996).

In contrast to LT the STs are both small monomeric peptides. They both contain multiple cysteine-cysteine disulphide bonds, which are responsible for their extremely heat stable nature; however, they differ in both their structure and their mode of action. STa is an 18 or 19 amino acid peptide containing 6 cysteines which form 3 disulphide bonds. The toxin is initially translated as a 72 amino acid precursor which is then exported and processed by signal peptidase 1; the pro-toxin folds in the periplasm and is then cleaved to its mature form by a second unknown protease prior to its release by diffusion across the outer membrane (Rasheed *et al.*, 1990). The receptor for STa has been identified as guanylate cyclase C (GC-C; de Sauvage *et al.*, 1992; Vaandrager *et al.*, 1994), which is a membrane spanning protein located in the apical membrane of intestinal epithelial cells. Binding of STa to GC-C results in stimulation of GC activity and the increase of intracellular cGMP levels (Crane *et al.*, 1992; Mezoff *et al.*, 1992); this increase leads to stimulation of chloride secretion and/or inhibition of sodium chloride absorption resulting in net fluid secretion into the lumen (Sears and Kaper, 1996). The STb toxin is also translated as a precursor but this 71 amino acid polypeptide has no homology to STa. The precursor is processed to a mature 48 amino acid peptide containing 4 cysteines which form disulphide bonds (Dreyfus *et al.*, 1992; Arriaga *et al.*, 1995). It is unclear exactly how STb causes diarrhoea, it does not stimulate increases in intracellular cAMP or cGMP, and as yet no specific receptor has been reported (Chao and Dreyfus, 1997). However, STb has been shown to cause calcium influx (Dreyfus *et al.*, 1992) and to stimulate the secretion of bicarbonate from intestinal cells (Sears and Kaper, 1996).

ETEC are associated with two types of disease; endemic infantile diarrhoea in the developing world and traveller's diarrhoea. Although ETEC infections are most common in infants, immunologically naive adults are also susceptible (unlike the case for EPEC, as described below). It is therefore unsurprising that ETEC is the most common aetiological agent in cases of traveller's diarrhoea amongst visitors from developed countries when travelling in countries where ETEC is endemic (Black, 1990; Arduino *et al.*, 1993; DuPont and Ericsson, 1993; Mattila, 1994).

Coincident with the discovery of ETEC other researchers reported the discovery of *E. coli* strains which could cause a dysentery-like disease (Sakazakai *et al.*, 1967; DuPont *et al.*, 1971). These strains, which were invasive both in the guinea pig keratoconjunctivitis test (Serény, 1957) and in HEp-2 cells (Day *et al.*, 1981) were designated enteroinvasive *E. coli* (EIEC). EIEC strains are closely related both biochemically and genetically to *Shigella* spp. (Brenner *et al.*, 1973), and both species cause similar diseases (Kopecko *et al.*, 1985). The pathogenic mechanism proposed for both EIEC and *Shigella* spp. has five stages: epithelial cell penetration; lysis of the endocytic vacuole; intracellular bacterial multiplication; utilisation of cytoskeletal elements for intracellular movement; and finally intercellular spread by penetration into adjacent cells (Sansone, 1992; Goldberg and Sansone, 1993; Parsot and Sansone, 1996). The determinants required for cell attachment, invasiveness and cell to cell spread are all encoded on a 230 kb plasmid, designated pInv (Baudry *et al.*, 1987; Small and Falkow, 1988). These include the *mxi* and *spa* genes, which encode a type III secretion apparatus (Andrews *et al.*, 1991; Venkatesan *et al.*, 1992; Maurelli, 1994; Allaoui *et al.*, 1995), and the *ipa* genes which encode the secreted proteins IpaA-D. The proteins IpaB, IpaC and IpaD are all effectors of the invasion process (Hale *et al.*, 1985; Baudry *et al.*, 1987; Menard *et al.*, 1993; Menard *et al.*, 1996). Intracellular movement and intercellular spread of both the *Shigella* spp. and EIEC requires a further surface protein VirG (Goldberg *et al.*, 1993). VirG localises to one pole of the bacterial cell and forms the nucleation site for polymerisation of an actin tail which drives bacterial movement both through the cytoplasm and through the cell membrane

into adjacent cells (Sansonetti, 1992; Vasselon *et al.*, 1992; Adam *et al.*, 1995). EIEC strains are clearly distinguishable from classical EPEC strains both by their distinctive group of serotypes (O28ac, O128ac, O124, O136, O143, O144, O152, O164) and by the aetiology of the disease they cause (Robins-Browne, 1987).

When analysis of strains from the classical EPEC serotypes demonstrated that they were not invasive in the Serény test and they did not produce the known *E. coli* enterotoxins (Gross *et al.*, 1976), there was some doubt as to their pathogenic nature. There were two schools of thought: some researchers assumed that since *E. coli* enterotoxins are generally plasmid encoded the strains designated EPEC were merely ETEC strains which had lost their plasmids (South, 1971; Drachmann, 1974); however, other groups maintained that the EPEC strains were true pathogens but that their virulence determinants had yet to be determined (Gross *et al.*, 1976; Gurwith *et al.*, 1977). Unequivocal proof of the pathogenic nature of EPEC strains was provided by Levine *et al.* (1978) who demonstrated the induction of diarrhoea upon feeding EPEC strains to adult volunteers. 10 of 24 volunteers presented with diarrhoea when fed 10^6 - 10^{10} bacteria of two EPEC strains, E2348-69 (O127:H6) and E851/71 (O142:H6), which had been previously isolated from outbreaks of infantile gastroenteritis in the UK (Gross *et al.*, 1976). Subsequent testing confirmed that neither of these strains produced any of the *E. coli* LT, STa or STb toxins (Robins-Browne *et al.*, 1982; Long-Krug *et al.*, 1984).

1.4.2 Enterohaemorrhagic *E. coli* (EHEC).

In the early 1980s a fourth group of diarrhoeagenic *E. coli* was proposed. The identification of EHEC as a new *E. coli* pathogen was based upon the combination of data from two concurrent but separate epidemiological studies. Riley *et al.* (1983) isolated a rarely observed *E. coli* serotype, O157:H7, from two groups of patients involved in similar outbreaks of gastroenteritis. All of the patients presented with distinct clinical symptoms, i.e. abdominal pain, watery diarrhoea followed by bloody diarrhoea but with little or no fever; this illness was designated haemorrhagic colitis

(HC). At the same time Karmali *et al.* (1983b) reported their observation of a correlation between cases of sporadic haemolytic uraemic syndrome (HUS) and the presence of cytotoxin-producing *E. coli* in the stools of the patients. As HUS, which is characterised by haemolytic anaemia, thrombocytopenia and renal failure, is typically preceded by a bout of bloody diarrhoea which is indistinguishable from HC, it was concluded that the two illnesses might have a common cause.

For their cytotoxin studies Karmali *et al.* (1983b) used an assay previously used to demonstrate the irreversible cytopathic effect on cultured Vero cells of culture filtrates from some *E. coli* strains (Konowalchuck *et al.*, 1977). Konowalchuck *et al.* (1977) had called the cytopathic factor Vero cytotoxin and the strains which produced it were designated verotoxigenic *E. coli* (VTEC). Concurrently O'Brien *et al.*, (1977) had demonstrated that extracts from certain *E. coli* strains were cytotoxic for HeLa cells; however, they showed that this activity could be neutralised by pretreatment with antitoxin to *Shigella dysenteriae* 1 Shiga toxin (Stx). Therefore the factor responsible was called Shiga-like toxin (SLT) and the strains were designated SLT-producing *E. coli* (SLTEC). Subsequent work by O'Brien *et al.* (1983) demonstrated that the SLT and Vero cytotoxin were in fact the same factor; furthermore, they showed that the O157:H7 strain originally isolated by Riley *et al.* (1983) also produced this toxin. This observation led to the conclusion that production of this specific cytotoxin by *E. coli* strains was responsible for the symptoms observed in both HC and HUS (Karmali *et al.*, 1983a). Subsequent work has shown that SLT is virtually identical to Stx and that in *E. coli* the toxin is encoded on a bacteriophage (Scotland *et al.*, 1983b; Smith *et al.*, 1983; O'Brien *et al.*, 1984). Since then over 100 *E. coli* serotypes have been shown to produce Stx (Karmali, 1989) and this group has been renamed Stx-producing *E. coli* (STEC).

Interestingly it has been shown that both the O157:H7 strains and some other non-O157:H7 STEC (Johnson *et al.*, 1996; Tarr and Neill, 1996) cause lesions on tissue culture cells that are virtually identical to those produced by EPEC strains.

Furthermore it has been demonstrated that these strains, designated "typical EHEC" (Nataro and Kaper, 1998), contain homologues of the major EPEC virulence determinants (these EHEC virulence determinants will be discussed further in the relevant EPEC sections below).

1.4.3 Other diarrhoeagenic *E. coli* groups.

A number of *E. coli* strains have been isolated from patients with diarrhoea which are adherent to HEp-2 cells but which are not invasive and do not contain either the toxins described above or the EPEC virulence determinants discussed below. The best characterised groups of these strains are the enteroaggregative *E. coli* (EAggEC) and diffusely-adherent *E. coli* (DAEC), which were identified by their unusual binding patterns during early studies investigating the HEp-2 adherence properties of diarrhoeagenic strains proposed to be EPEC (Cravioto *et al.*, 1979; Scaletsky *et al.*, 1984; Nataro *et al.*, 1985b).

EAggEC bind to HEp-2 cells (and to some degree to the glass coverslips they are mounted on) forming a characteristic "stacked-brick" pattern termed aggregative adherence (AA). The AA phenotype has been correlated with possession of one of two aggregative adherence fimbriae (AAF); a minority of EAggEC strains express AAF/I, a bundle-forming fimbrial structure (Nataro *et al.*, 1992; Nataro *et al.*, 1993; Nataro *et al.*, 1994; Savarino *et al.*, 1994), whilst the majority express AAF/II whose gene cluster arrangement is typical of pili of the Dr family (Czeczulin *et al.*, 1997). Approximately 40% of EAggEC express the toxin EAST1, a 38 amino acid homologue of STa (Savarino *et al.*, 1991; Savarino *et al.*, 1993), though its role in pathogenesis is unclear. Other investigators have also described EAggEC cytotoxins, e.g. Eslava *et al.* (1993) reported a 108 kDa cytotoxin which caused mucosal lesions in rat ileum loops. It has been proposed (Nataro and Kaper, 1998) that this cytotoxin may be the same as the 120 kDa supernatant protein described by Baldwin *et al.* (1992) which causes increases in intracellular Ca²⁺ levels in HEp-2 cells. Histopathological studies using gnotobiotic piglets (Tzipori *et al.*, 1992), rabbit ileal loops (Vial *et al.*, 1988) or *in vitro*

organ culture (using paediatric small bowel mucosa; Hicks *et al.*, 1996) all show enhanced mucus production upon infection with EAggEC. Furthermore it has been observed in these studies that the bacteria are located within a thick mucoid gel or biolayer. Interestingly volunteers infected with EAggEC are also observed to produce a diarrhoea which is very mucoid (Nataro *et al.*, 1995). Histopathological studies of EAggEC infections show damage to mucosal surfaces; i.e. in the rat-ileal loop model Vial *et al.* (1988) observed shortening of villi and haemorrhagic necrosis of the villus tips; similar lesions were also observed, at autopsy, in the ilea of patients who died as a result of EAggEC-induced persistent diarrhoea (Eslava *et al.*, 1992). A model for EAggEC pathogenesis has been proposed where the bacteria bind to the mucus layer in the ileum and cause over-production of mucus to form a biolayer within which they can replicate; this is then followed by the elaboration of bacterial cytotoxins which cause damage to the mucosal surface resulting in the lesions previously described (Nataro and Kaper, 1998).

DAEC bind to HEp-2 cells as single bacteria in a diffusely adherent (DA) manner, similar to some EPEC strains, but they do not cause the characteristic EPEC lesions. It has been demonstrated that the DA phenotype is conferred by a fimbria, F1845 (Bilge *et al.*, 1989); however, until recently very little else was known about the pathogenesis of DAEC (Nataro and Kaper, 1998). New data now suggest that DAEC are able to induce cytoskeletal F-actin rearrangements in cultured human intestinal INT407 cells (Bernet-Camard *et al.*, 1996b) by piracy of the decay-accelerating factor signal transduction system (CD55; Davis *et al.*, 1988; Kammer *et al.*, 1988; Shenoy-Scaria, *et al.*, 1992; Shibuya *et al.*, 1992; Nicholson-Weller and Wang, 1994; Bernet-Camard *et al.*, 1996a). Furthermore, inhibitor studies by the same investigators (Peiffer *et al.*, 1998) suggest that the observed rearrangement of actin requires tyrosine kinase, phospholipase C γ , phosphatidylinositol 3-kinase, and PKC activities, as well as the release of intracellular Ca²⁺. As yet there is no model for DAEC pathogenesis.

There are also *E. coli* which have been implicated in infection but about which there is not enough epidemiological data to confirm a role in diarrhoeal disease. Amongst these are the cell detaching *E. coli* (CDEC) which were first observed due to their characteristic cytotoxic effect upon HEP-2 cells (Gunzberg *et al.*, 1993). *E. coli* identified as CDEC frequently encode haemolysin and some have also been shown to encode the cytotoxic necrotising factors CNF1 and CNF2 (de Rycke *et al.*, 1990; Oswald *et al.*, 1994), which are more commonly found in *E. coli* associated with extraintestinal infections (Caprioli *et al.*, 1987; Oswald *et al.*, 1994). There is also now evidence that there are diarrhoeagenic *E. coli* which do not fit easily into the strict groupings presented here and which can contain multiple virulence factors classically associated with quite different *E. coli* diarrhoeal diseases. This is supported by the observations of Hedberg *et al.* (1997) who reported the isolation of an *E. coli* strain (O39:NM) after an outbreak of diarrhoea in 1991 which was epidemiologically and clinically similar to ETEC infection. Intriguingly this strain did not produce LT, STa or STb, but on analysis was found to encode both the enteroaggregative *E. coli* toxin EAST1 and the EPEC genes required for enterocyte effacement.

1.5 Histopathology of EPEC infection.

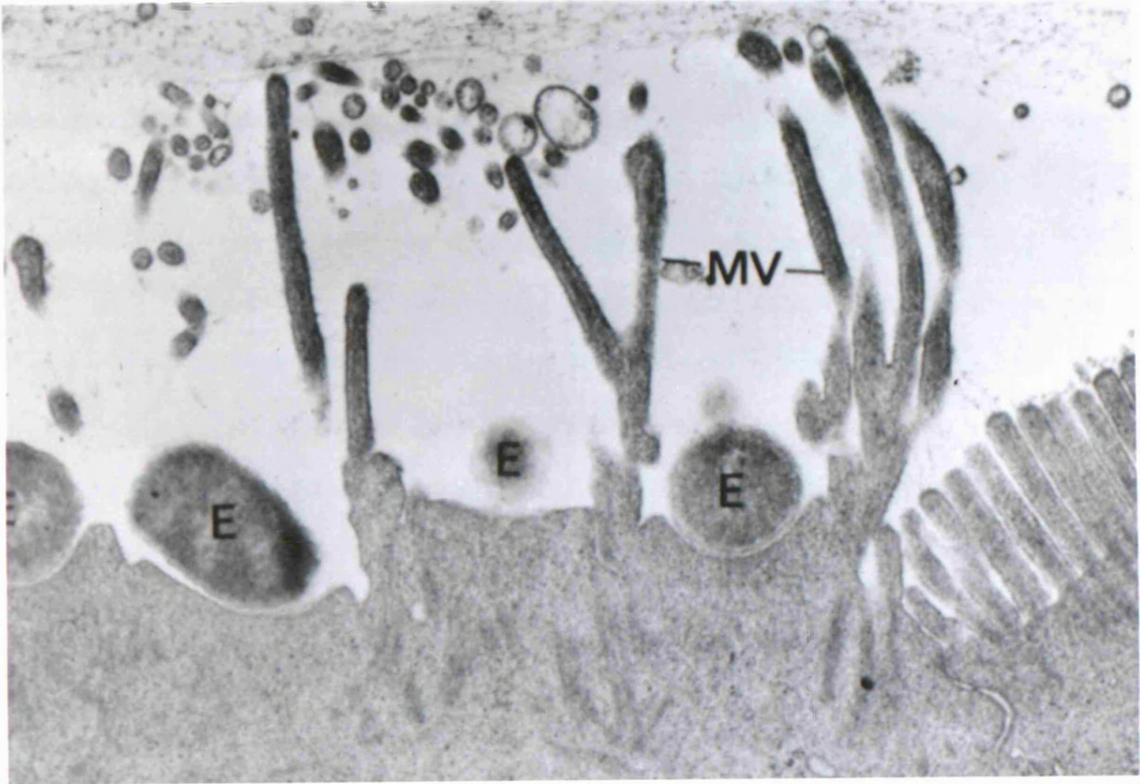
The first evidence of the effects EPEC infection upon gut ultrastructure was observed in an examination of small bowel tissue taken from gnotobiotic piglets which had been infected with *E. coli* O55:H7. Staley *et al.* (1969) described the lengthening and shedding of microvilli and the intimate attachment of bacteria to the cell surface. At the site of bacterial attachment they observed invagination and thickening of the plasma membrane and increased electron density in the cytoplasm. Polotsky *et al.* (1977) confirmed these ultrastructural changes using O111 and O26 EPEC serotypes in rabbit ileal loops. In a comprehensive study, Moon *et al.* (1983) proved the pathogenicity of several well characterised EPEC strains of human origin using both a piglet model and ligated intestinal loops in pigs and rabbits. In recognition of the characteristic pathological changes that these strains caused, i.e. effacement of the microvilli and intimate attachment of bacteria and cell to form a pedestal structure,

Moon *et al.* (1983) designated them attaching-effacing *E. coli*. The use of animal models was validated by two ultrastructural studies of the intestines of infants with EPEC infections (Ulshen and Rollo, 1980; Rothbaum *et al.*, 1982; Rothbaum *et al.*, 1983). In both cases examination of bowel biopsies demonstrated the characteristic attaching and effacing (AE) lesions seen previously.

AE lesion formation has also been demonstrated in *in vitro* EPEC infections of gut biopsies (figure 1.1) and with cultured human epithelial cells (Knutton *et al.*, 1987a; 1987b). The observation that the dense staining material seen in the AE lesion was composed of high concentrations of filamentous actin led to the development of the fluorescent-actin staining (FAS) test (Knutton *et al.*, 1988; 1989a). A fluorescein isothiocyanate (FITC)-labelled fungal toxin, phalloidin, which binds specifically to actin, was used to detect the dense concentrations of filamentous actin below bacteria attached to tissue culture cells. Development of the FAS test was to prove extremely useful; beside being a diagnostic test for EPEC (Knutton *et al.*, 1989b; 1991), it also provided a fast and easy method to screen for the EPEC genes encoding the virulence determinants required for AE lesion formation. In addition to filamentous actin AE lesions contain a number of other cytoskeletal components including α -actinin, talin, ezrin and myosin light chain (Finlay *et al.*, 1992).

Histopathological lesions similar to AE lesions have also been demonstrated in tissue culture cells infected with EHEC (Knutton *et al.*, 1989a; Ismaili *et al.*, 1995) and with diarrhoeagenic *E. coli* strains isolated from rabbits (RDEC; Cantey and Blake, 1977), dogs, pigs and calves (Drolet *et al.*, 1994; Fischer *et al.*, 1994; Zhu *et al.*, 1994). Formation of AE lesions is not limited to *E. coli* and has also been observed in strains of *Citrobacter rodentium* (formerly *freundii*), which cause murine colonic hyperplasia (Schauer and Falkow, 1993), and in a conserved subgroup of *Hafnia alvei* strains which are implicated in infantile diarrhoea (Albert *et al.*, 1991; 1992).

Figure 1.1 Ultrastructure of EPEC infected human gut tissue. Human duodenal biopsies infected with EPEC (E) and observed by transmission electron microscopy display characteristic lesion formation, i.e. effacement of the microvilli (mv), pedestal formation and accretion of densely staining actin below the attached bacteria (electron-micrograph kindly supplied by Dr Stuart Knutton, Birmingham; Knutton *et al.*, 1987).



1.6 Pathogenic mechanism of EPEC.

The formation of the characteristic EPEC AE lesions involves multiple steps and the coordinated expression of a number of bacterial virulence determinants which are encoded either on a large EPEC virulence plasmid or within a 35 kb chromosomal pathogenicity island called the locus of enterocyte effacement (LEE).

1.6.1 Localised adherence.

In contrast with other *E. coli* strains many EPEC are able to adhere efficiently to cultured epithelial cells. Cravioto *et al.* (1979) found that 80% of 51 strains belonging to classical EPEC serotypes were able to bind to HEp-2 cells whereas only 19% of other *E. coli* strains bound (both pathogenic and non-pathogenic). Other groups confirmed this phenotype in both HEp-2 (Clausen and Christie, 1982) and HeLa cells (Lacroix *et al.*, 1984); these groups also noticed that rather than completely covering the cells EPEC tended to adhere as discreet microcolonies. Scaletsky *et al.*, (1984; 1985) extended this work by demonstrating that individual EPEC strains bound to HeLa cells either as microcolonies, designated localised adherence (LA), or as single bacteria, designated diffusely attaching (DA). Expression of the HEp-2 adherence phenotype was shown to be associated with the carriage of large (50-70 MDa) genetically related plasmids termed EPEC adherence factor (EAF) plasmids (Baldini *et al.*, 1983; Nataro *et al.*, 1985b). Indeed when the HEp-2 adherent strain E2348-69 (O127:H6) was cured of its EAF plasmid it was no longer able to bind to HEp-2 cells (Baldini *et al.*, 1983); furthermore this strain was markedly less virulent and caused severely reduced disease in adult volunteers compared with its parent (Levine *et al.*, 1985). Nataro *et al.* (1985a) and Baldini *et al.* (1986) identified a 1 kb probe from the EAF plasmid pMAR2 which was specific for LA HEp-2 cell adhesiveness. It was observed that strains positive for the EAF probe (class I EPEC), typified by the serogroups O55, O111, O119, O127, O128 and O142, were associated with severe outbreaks and epidemic diarrhoea; in contrast EAF probe negative (class II EPEC), typified by serotypes O18, O44, O114, were rarely associated with outbreaks (Nataro *et al.*, 1985a; Levine, 1987). However it has been demonstrated, both by

epidemiological evidence (Levine and Edelman, 1984) and by volunteer studies using an O114:H2 strain (Levine *et al.*, 1985), that class II EPEC are diarrhoeagenic.

Despite speculation that the EAF adhesin responsible for LA was fimbrial in nature (Knutton *et al.*, 1987a) attempts to identify pili in EPEC had been unsuccessful (Scotland *et al.*, 1983a; Nataro *et al.*, 1987). Finally Giron *et al.* (1991) identified a 7 nm diameter pilus in EPEC strain B171 (O111:NM) which had been previously undetected because it was not expressed in normal laboratory media (Vuopio-Varkila and Skoolnik, 1991; Puente *et al.*, 1996). Because these fimbriae tended to align and aggregate to form rope-like structures they were named bundle-forming pili (BFP). The N-terminal sequence of the 19.5 kDa major subunit of purified BFP (Giron *et al.*, 1991) showed significant homology with the toxin-coregulated pilus (TCP) of *Vibrio cholerae* (Shaw and Taylor, 1990; Manning, 1997) and with other type IV pili (Strom and Lory, 1993). Antibodies raised against purified BFP significantly inhibited EPEC adherence to HEp-2 cells (Giron *et al.*, 1991). Simultaneously Donnenberg *et al.* (1990) identified a number of *TnphoA* mutants which were deficient in the LA phenotype. Several of these mutants defined a plasmid-encoded 579 bp open reading frame, *bfpA*, whose predicted gene product was homologous to the N-terminal sequence of the major subunit of BFP (Donnenberg *et al.*, 1992). BfpA contains a signal sequence typical of type IV fimbriae (Strom and Lory, 1992) and the mature 180 amino acid polypeptide has a predicted molecular weight of 18,730 Da. Additional *TnphoA* mutants identified the chromosomal gene *dsbA* which encodes a periplasmic protein required for disulphide bond formation in exported proteins. It has since been determined that formation of a disulphide bond in BfpA is critical for stability of the protein and therefore essential for pili biogenesis (Zhang and Donnenberg, 1996; Donnenberg *et al.*, 1997b).

Attempts to complement transposon mutations in *bfpA* with fragments of the EAF plasmid indicated that it probably formed part of a large operon (Donnenberg *et al.*, 1992). This was confirmed by the simultaneous identification by two independent

groups (Sohel *et al.*, 1996; Stone *et al.*, 1996) of an operon of 14 genes (**figure 1.2**) which was sufficient to encode expression of BFP on transfer to a *E. coli* K-12 strain (Stone *et al.*, 1996). Many of the *bfp* genes have homology with genes involved in the biosynthesis of type IV pili in other Gram negative bacterial species but a number appear to be specific to BFP. Intriguingly whilst the 14 gene *bfp* operon encoded all of the genes required to express BFP it could not confer the LA phenotype onto *E. coli* K-12 without the simultaneous introduction of a plasmid containing a second EAF plasmid fragment; this region had previously been identified to contain the *perABC* genes (Gomez-Duarte and Kaper, 1995) which encode a regulatory factor that is also required for full expression of chromosomally encoded EPEC virulence genes (see below).

Recent work using an EPEC strain containing a non-polar mutation in *bfpA* has demonstrated unequivocally that BFP are required for EPEC virulence in adult volunteers (Beiber *et al.*, 1998). Beiber *et al.* (1998) also found that a strain carrying a non-polar mutation in the *bfpF* gene, which is not required for BFP expression (Anantha *et al.*, 1998), was 200-fold less virulent than parent. Intriguingly it has been observed that the bacterial aggregates produced by the *bfpF* mutant upon growth under BFP-inducing conditions are stable and do not dissociate upon entry into stationary phase as is seen with the parent strain. It has therefore been proposed that BfpF is required for the shedding of BFP from the cell surface to allow dissociation of EPEC microcolonies and to facilitate the dispersal of bacteria in order to colonise additional intestinal sites (Beiber *et al.*, 1998).

1.6.2 EPEC induced signal transduction.

Adherence of EPEC to epithelial cells and the subsequent formation of AE lesions has been shown to involve the induction of a number of signal transduction pathways. Several groups have reported elevation of intracellular Ca^{2+} levels upon EPEC infection of host cells (Baldwin *et al.*, 1991; 1993; Dytoc *et al.*, 1994; Philpott *et al.*, 1996). Baldwin *et al.* (1991) proposed that the increase in intracellular Ca^{2+} levels

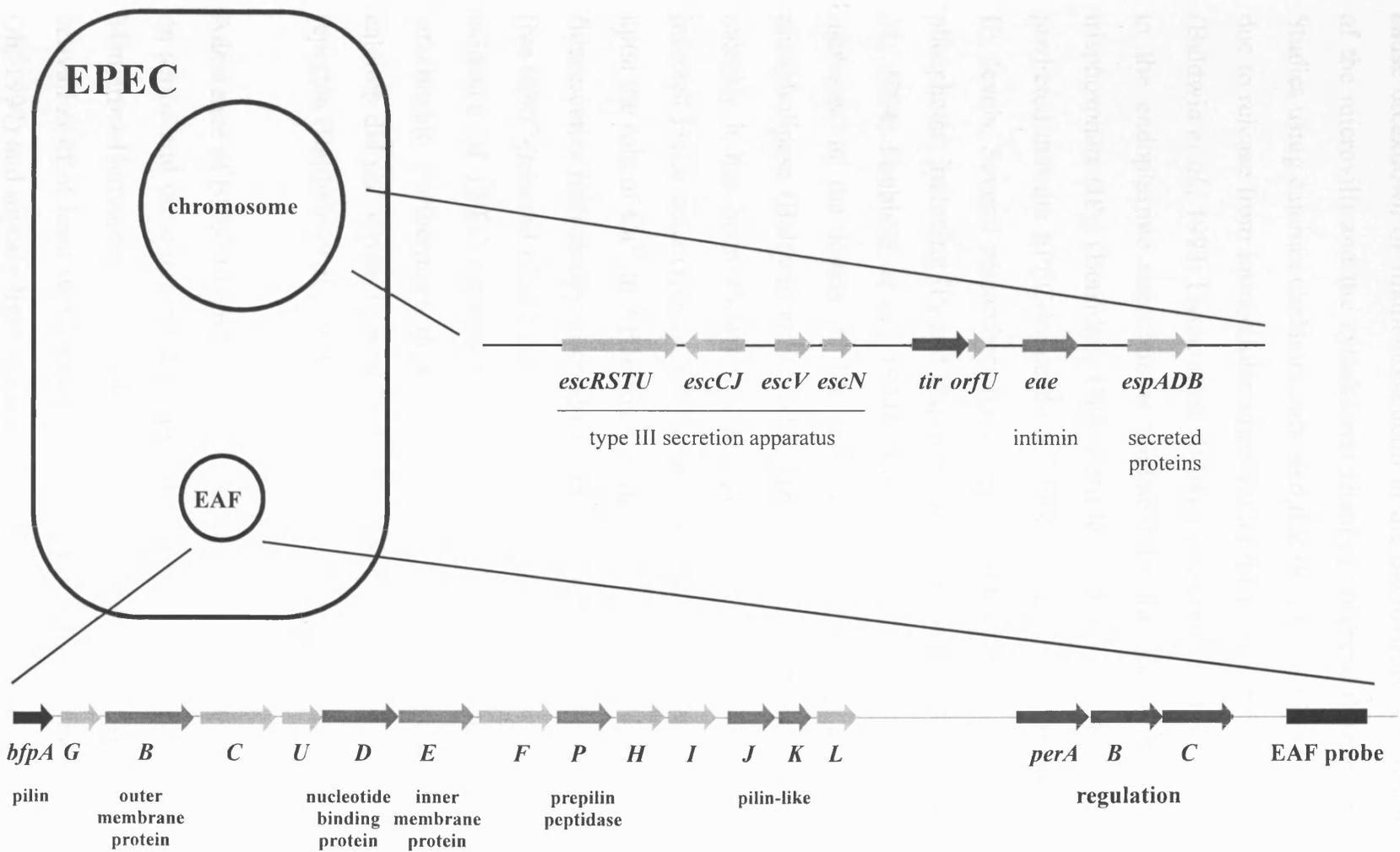


Figure 1.2 Schematic representation of the genes involved in EPEC pathogenesis. Chromosomal genes are clustered within the LEE, which encodes a type III secretion system, intimin and secreted effector proteins. The EAF plasmid encodes the BFP as well as the Per regulators. (Adapted from Nataro and Kaper, 1998).

would activate the calcium-dependent actin-severing protein gelsolin which would cause breakdown of filamentous actin in the microvillus core resulting in effacement of the microvilli and the cytoskeletal changes observed during AE lesion formation. Studies using calcium chelators indicated that the increase in intracellular Ca^{2+} must be due to release from intracellular stores rather than from an influx of extracellular Ca^{2+} (Baldwin *et al.*, 1993; Dytoc *et al.*, 1994). As mobilisation of intracellular Ca^{2+} stored in the endoplasmic reticulum is mediated by the second messenger 1,4,5-inositol trisphosphate (IP_3) (Berridge, 1984; Berridge and Irvine, 1984), Baldwin *et al.* (1991) predicted that the EPEC-induced Ca^{2+} fluxes would be accompanied by alterations in IP_3 levels. Several researchers have since reported increases in the levels of inositol phosphates, including IP_3 and IP_4 , in tissue culture cells infected with EPEC (Dytoc *et al.*, 1994; Foubister *et al.*, 1994b; Knutton *et al.*, 1994). It was proposed that the increases in the levels of IP_3 and Ca^{2+} were due to the activation of a host phospholipase (Baldwin *et al.*, 1991; Dytoc *et al.*, 1994; Foubister *et al.*, 1994b) and recently it has been shown that phospholipase C- γ 1 is indeed activated in EPEC infected HeLa cells (Kenny and Finlay, 1997). However recent data have cast doubt upon the role of Ca^{2+} in AE lesion formation. Bain *et al.* (1998) used calcium imaging fluorescence microscopy to conduct a comprehensive study of Ca^{2+} concentrations in live EPEC-infected HEP-2 cells; in contrast with previous studies, their work found no evidence of EPEC-induced increases in Ca^{2+} concentrations at sites of bacterial attachment. Furthermore these authors reported that the chelation of intracellular calcium did not appear to prevent the formation of AE lesions as had previously been reported (Baldwin *et al.*, 1991).

Adherence of EPEC to epithelial cells induces phosphorylation of several host proteins on serine and threonine residues, including myosin light chain (Baldwin *et al.*, 1990; Manjarrez-Hernandez *et al.*, 1992; 1996). This protein phosphorylation is due to the activation of at least two kinases, protein kinase C (Baldwin *et al.*, 1990; Crane and Oh, 1997) and myosin light chain kinase (Manjarrez-Hernandez *et al.*, 1996; Yuhan *et al.*, 1997). In addition adherence of EPEC induces the tyrosine phosphorylation of a

number of proteins in the host cell (Rosenshine *et al.*, 1992; Kenny and Finlay, 1997); the most prominent amongst these being a 90 kDa protein called Hp90 which is found in the host cell membrane (Rosenshine *et al.*, 1992). Hp90 and the other tyrosine phosphorylated proteins have been observed to localise directly under the site of bacterial attachment at the tip of the pedestal (Rosenshine *et al.*, 1996). Rosenshine *et al.* (1996) also demonstrated that the Hp90 protein serves as a receptor for binding of the EPEC adhesin intimin (see below). Recently the Hp90 protein has been identified as an EPEC secreted protein which has been designated Tir (translocated intimin receptor) (Kenny *et al.*, 1997b).

It has been previously reported that signal transduction, and specifically tyrosine phosphorylation, is required for the formation of AE lesions (Rosenshine *et al.*, 1996); however, with recent data it has become unclear if tyrosine phosphorylation of the Hp90/Tir receptor is essential for this process (Finlay *et al.*, 1998). Rabinowitz *et al.* (1996) have described a EPEC strain containing a mutation in the gene *sepZ* which is apparently wildtype for AE lesion formation yet is unable to induce tyrosine phosphorylation of Hp90 (Tir). In addition, whilst EHEC O26:H- infection of tissue culture cells has been shown to involve the transfer and subsequent tyrosine phosphorylation of an 80 kDa EHEC protein (EspE/Tir/Hp90) (Deibel *et al.*, 1998), it is clear that AE lesion formation by EHEC O157:H7 strains does not require the tyrosine phosphorylation of the Tir/Hp90 receptor (Ismaili *et al.*, 1998). It remains to be seen if tyrosine phosphorylation of Tir/Hp90 is required for EPEC or EHEC virulence *in vivo*.

In addition to the effects that EPEC have upon epithelial cells there is evidence that they can also both indirectly and directly affect the behaviour of cells of the immune system. Savkovic *et al.* (1996) demonstrated that in an *in vivo* coculture system containing polarised T84 cell monolayers and polymorphonuclear leukocytes (PMN), attachment of EPEC was sufficient to induce migration of the PMNs across the monolayer. The evidence suggests that binding of EPEC activates the eukaryotic NF-

κ B transcription factor which then upregulates interleukin-8 (IL-8) which is a PMN chemoattractant (Savkovic *et al.*, 1996; 1997). This activation is dependent upon known EPEC virulence determinants and was not observed in an EPEC mutant (*espB*) which adhered but which did not activate signal transduction pathways (Savkovic *et al.*, 1996). In contrast Klapproth *et al.* (1995) reported a heat and protease sensitive factor from EPEC and EHEC cell lysates which inhibited lymphocyte activation and lymphokine production, but which was still produced in EPEC strains containing mutations in known virulence determinants. As the factor inhibited the mitogen-stimulated expression of IL-2, IL-4, IL-5, and gamma interferon it could therefore be important in modifying gastrointestinal immune responses to EPEC and EHEC infection.

1.6.3 Intimate adherence.

The intimate adhesion of EPEC to the host cell membrane required for AE lesion formation is dependent upon a 94 kDa outer membrane protein called intimin (Jerse and Kaper, 1991). The gene for intimin, *eaeA* (*E. coli* attaching and effacing), was identified by Jerse *et al.* (1990) who used the FAS test to screen a bank of *TnphoA* mutants for those which were unable to form AE lesions. Though strains with mutations in *eaeA* cannot accrete actin to the site of bacterial attachment (Donnenberg and Kaper, 1991), they retain the ability to induce the signal transduction events described in section 1.6.2 (Rosenshine *et al.*, 1992; Foubister *et al.*, 1994b; Savkovic *et al.*, 1996; Stein *et al.*, 1996b). The intimin protein shares 31% identity and 50% similarity with the invasin protein of *Yersinia* spp. (Isberg *et al.*, 1987); however, in contrast with invasin, the expression of intimin in *E. coli* K-12 does not in itself confer the ability to adhere to, or invade, tissue culture cells (Jerse *et al.*, 1990).

An assessment of the importance of intimin in EPEC infection has been made using a randomised double-blind study (Donnenberg *et al.*, 1993a). Diarrhoea was observed in 11/11 adult human volunteers who ingested 2×10^{10} bacteria of a wildtype EPEC strain E2348-69 compared to 4/11 volunteers who ingested an equal dose of an isogenic *eaeA*

mutant strain, CVD206 (Donnenberg and Kaper, 1991). The number and volume of liquid stools was also significantly less in the recipients of the mutant strain. This study unambiguously indicated a role for intimin in EPEC virulence, but the residual diarrhoea seen in some volunteers who received the mutant demonstrated that other virulence determinants are also involved.

Homologues of the *eaeA* gene have been found in all EPEC, EHEC, *Citrobacter rodentium*, and *Hafnia alvei* strains which have been shown capable of forming AE lesions. Furthermore it has been demonstrated that expression of EPEC-derived intimin can functionally complement a *Citrobacter rodentium eaeA* mutant (Frankel *et al.*, 1996b). Comparison of the predicted amino acid sequences of the intimin proteins whose genes have currently been sequenced (Beebakhee *et al.*, 1992; Yu and Kaper, 1992; Schauer and Falkow, 1993; Frankel *et al.*, 1994; Agin *et al.*, 1996; Agin and Wolf, 1997) shows high conservation in the N-terminal region of the protein and much greater variability in the C-terminal region (Nataro and Kaper, 1998). As the C-terminal 150 amino acids of intimin have been shown to be the region involved in receptor binding (Frankel *et al.*, 1994; Frankel *et al.*, 1995), it has been proposed that the divergent nature of the C-terminal regions of the EPEC and EHEC intimins might explain the differences in their patterns of intestinal colonisation (Yu and Kaper, 1992; Nataro and Kaper, 1998).

It has been clearly demonstrated that the tyrosine-phosphorylated Tir protein (formerly Hp90; Kenny *et al.*, 1997b), which is transferred from the EPEC cell to the host, is the receptor for intimin (Rosenshine *et al.*, 1996); however, there is still some debate as to whether it is the only receptor. It has been reported (Frankel *et al.*, 1994; Frankel *et al.*, 1995) that intimin-maltose binding protein fusion proteins can bind to epithelial cells which have not been infected with EPEC; these cells cannot have Tir protein on their surface thus indicating the presence of another intimin receptor. Frankel *et al.* (1996a) have also reported that intimin can bind to the β_1 integrins; these cell surface proteins also serve as receptors for the invasin protein of *Yersinia* spp. (Leong *et al.*, 1990).

Based upon this evidence it seems probable that intimin has more than one receptor; however, it remains unclear if they are all equally important in EPEC virulence.

1.6.4 EPEC invasion of epithelial cells.

EPEC is not generally regarded as an invasive organism because it is negative in the Serény test (Levine *et al.*, 1978; Edelmann and Levine, 1983, Levine, 1987; Robins-Browne, 1987). However, intracellular bacteria have been observed in EPEC-infected tissue culture cells (Knutton *et al.*, 1987), in animal models of EPEC infection (Staley *et al.*, 1969; Moon *et al.*, 1983; Tzipori *et al.*, 1985) and in human biopsy specimens of EPEC-infected children (Ulshen and Rollo, 1980). Furthermore Donnenberg *et al.* (1989) demonstrated, using the gentamicin protection assay, that significant numbers of bacterial cells are internalised when EPEC are used to infect the non-phagocytic cell line HEp-2. As with EIEC and *Shigella* and *Yersinia* spp. the EPEC invasion process can be blocked by inhibitors of host cell microfilaments, but unlike these bacteria EPEC cell entry can also be blocked by inhibitors of host microtubules (Donnenberg *et al.*, 1990b; Francis *et al.*, 1991). EPEC also differ from recognised intracellular pathogens, such as *Salmonella*, *Shigella* and EIEC, in that the bacteria do not escape from their phagocytic vacuoles and do not multiply within the host cells (Donnenberg *et al.*, 1990b) suggesting that they are not adapted for intracellular survival.

Whether the internalisation of EPEC bacteria observed in tissue culture assays has any relevance in EPEC pathogenesis *in vivo* has yet to be determined; however, it has been observed that many of the genes required for AE lesion formation are also required for invasion (Kaper, 1994; Donnenberg *et al.*, 1996). Thus when Donnenberg *et al.* (1990a) screened a bank of EPEC Tn*phoA* mutants using a gentamicin protection assay to isolate those which were deficient in HEp-2 cell invasion, they isolated a large number of strains with mutations in genes which are involved in AE lesion formation. These strains included mutants deficient in BFP production, whose inserts were located in *bfpA* and *dsbA*, and mutants deficient in intimate adherence, whose inserts were located in *eaeA* and *espB* (Donnenberg *et al.*, 1993b; see below). A further class

of mutants, termed category four mutants (cfm), which were negative for both AE lesion formation and induction of signal transduction were subsequently found to be in genes encoding a type III secretion system (Jarvis *et al.*, 1995; see below).

1.6.5 Secretion of EPEC proteins.

It has been demonstrated repeatedly that classical EPEC strains do not express the *E. coli* enterotoxins LT, STa or STb (Levine *et al.*, 1978; Robins-Browne *et al.*, 1982; Levine, 1987). Therefore it was unexpected when it was reported that EPEC secrete a number of proteins when cultured in tissue culture media under an atmosphere of 5% CO₂ (Haigh *et al.*, 1995). Subsequent reports from two independent groups described the identification and N-terminal sequencing of four *E. coli* secreted proteins (designated Esp) (Jarvis *et al.*, 1995; Kenny and Finlay, 1995). For three of these proteins, EspA (25 kDa), EspB (37 kDa; previously identified as EaeB; Donnenberg *et al.*, 1993b) and EspD (40 kDa), it has been shown that they are required for induction of signal transduction in epithelial cells and that mutations in their respective genes abolish AE lesion formation (Donnenberg *et al.*, 1993b; Kenny *et al.*, 1996; Lai *et al.*, 1997). The genes for EspA, EspB and EspD are located downstream of *eaeA* within the LEE (**figure 1.2**). Secretion of proteins homologous to EspA, EspB and EspD has been reported in both O157:H7 and O26 EHEC strains (Ebel *et al.*, 1996; Jarvis and Kaper, 1996), and in DAEC strains (Beinke *et al.*, 1998). The fourth protein EspC (110 kDa) is not required for signal transduction or lesion formation and shows homology with the IgA protease family of autotransporter secreted proteins (Stein *et al.*, 1996a).

The polypeptides predicted from the DNA sequences of the *espA*, *espB* and *espD* genes do not contain conventional *E. coli* N-terminal signals (Donnenberg *et al.*, 1993b; Kenny *et al.*, 1996; Lai *et al.*, 1997); furthermore, N-terminal sequencing of the secreted proteins has demonstrated that they are exported to the medium intact (Kenny *et al.*, 1995). Jarvis *et al.* (1995) identified four chromosomal genes, initially named *sepABCD* (secretion of *E. coli* proteins), whose predicted protein products showed homology with components of type III secretion systems responsible for exporting

proteins without typical signal sequences in *Yersinia*, *Salmonella* and *Shigella*. A strain carrying a mutation in one of these genes, *sepB*, was found to be unable to secrete EspA, EspB and EspD but could secrete the 100 kDa EspC protein. In total nine genes with homology to type III secretion systems have been identified within the LEE region upstream of *eaeA* (Elliot *et al.*, 1998b; Nataro and Kaper, 1998; **figure 1.2**). The Tir protein also requires the type III secretion system for transfer to the host cell membrane (Kenny *et al.*, 1997b); however there is some indications that it may not be secreted in a manner identical to the Esp proteins (Finlay *et al.*, 1998).

Recent data from studies of type III secretion systems in other pathogenic bacteria (Lee, 1997) have shown that the function of these systems is to transport effector proteins directly into host cells. Kenny *et al.* (1995) initially observed that secreted EspB became protease resistant when associated with the epithelial host cell and they proposed that this might be due to insertion of EspB into the cell membrane or its entry into a cell compartment. This hypothesis has now been proven by Wolff *et al.* (1998) who have used immunoblotting of fractionated EPEC-infected whole cells and a calmodulin-dependent EspB-adenylate cyclase fusion to demonstrate that EspB is indeed transferred directly into the membrane and cytoplasm of HeLa or Caco-2 cells. Knutton *et al.* (1998) have confirmed these data and have shown in addition that translocation of EspB to the epithelial cell cytoplasm requires functional EspA. Two independent research groups have recently observed that EspA forms filamentous organelles on the surface of EPEC and EHEC O26 cells (Ebel *et al.*, 1998; Knutton *et al.*, 1998) and that at least for EPEC these filaments are seen to interact with the host epithelial cell surface (Knutton *et al.*, 1998). Based upon the similarity of the coiled-coil domain structure of EspA (Pallen *et al.*, 1997) to the flagellin monomer, it has been proposed that the EspA organelles observed could be hollow filaments, akin to flagella (Emerson *et al.*, 1970), through which effector proteins could be carried from EPEC to the host cell (Knutton *et al.*, 1998).

1.6.6 Locus of enterocyte effacement.

All of the genes which have been shown to be essential for the formation of the AE lesion have been localised to a single region of the EPEC chromosome called the locus of enterocyte effacement (LEE; McDaniel *et al.*, 1995). The LEE of EPEC strain E2348-69 O127:H6 has now been completely sequenced; it is 35 624 bp in length and encodes 41 predicted open reading frames (Elliot *et al.*, 1998b) including *eaeA*, *espABD*, *tir*, and the *sep* genes (McDaniel *et al.*, 1995; Kenny *et al.*, 1996; Lai *et al.*, 1997; **figure 1.2**). The sequence of the LEE from EHEC O157:H7 strain EDL933 has also been determined; it is larger than the LEE from E2348-69 (43,359 bp) and contains an additional 13 genes which have been putatively identified as a P4 family prophage (Perna *et al.*, 1998). Most of the other 41 EHEC LEE genes show greater than 95% identity with their EPEC homologues, however several genes unexpectedly show significantly lower levels of homology which cannot be explained as simple genetic divergence. The LEEs of E2348-69 and O157:H7 have been found to be inserted within the selenocysteine tRNA gene (*selC*) at a point equivalent to 82 min on the *E. coli* K-12 genome (McDaniel *et al.*, 1995). However, the LEEs of other serogroups of both EPEC and EHEC have been shown to be inserted elsewhere (Weiler *et al.*, 1997). The *selC* gene is also the site for insertion of a large 70 kb pathogenicity island (PAI) in uropathogenic *E. coli* strain 536 (Blum *et al.*, 1994). The complete sequence of the LEE has an average GC content of 38.4%, significantly lower than the *E. coli* K-12 chromosome (50.8%; Blattner *et al.*, 1997), and similar to many other pathogenicity islands, suggesting that the entire region has probably been recently acquired by horizontal transfer (Hacker *et al.*, 1997). The EPEC LEE is not only required for AE lesion formation but, as shown by transfer of the entire region to *E. coli* K-12 on a plasmid, it contains all of the genes required to confer the AE phenotype (McDaniel *et al.*, 1997).

To date genes in the LEE have been given names based upon the phenotypes of their mutants; i.e. *eae* (*E. coli* attaching and effacing), *esp* (*E. coli* secreted protein) and *sep* (secretion of *E. coli* proteins). However, since the completion of the sequence of the

LEE has allowed identification of gene homologues amongst virulence determinants from other bacterial species, a new standardised nomenclature has been proposed (Elliot *et al.*, 1998b) which is similar to those adopted for other type III secretion systems (Bogdanove *et al.*, 1996; Yahr *et al.*, 1997). The new system draws upon the similarities between components of the EPEC and *Yersinia* type III secretion systems; thus genes which are homologues of *Yersinia* secretion genes (*ysc*) have been redesignated *esc* (*E. coli* secretion; e.g. *sepA*, homologous to *yscV*, becomes *escV*). In addition the newly identified genes for chaperone proteins (Wainwright and Kaper, 1998) were designated *ces* (chaperone for *E. coli* secretion). However the genes for secreted proteins (*esp*) and *eaeA* remained unchanged and genes involved in secretion but without homologues in *Yersinia* will remain *sep*.

1.6.7 Regulation of EPEC virulence genes.

A common theme amongst bacterial virulence genes is that their expression is frequently transitory and that this expression is tightly regulated (Dorman, 1994; Finlay and Falkow, 1997). It is obvious that inappropriate expression of any gene (e.g. an enzyme in the absence of its substrate) is metabolically expensive for the cell. However, for virulence genes inappropriate expression could also potentially inhibit the pathogenic process or bring the bacteria prematurely to the attention of the host immune system (Straus, 1995). Pathogenic bacteria use a variety of environmental factors (e.g. temperature, osmolarity, pH, iron levels) to sense the microenvironment that they occupy within the host and to act as cues for regulation of virulence genes (Gross, 1993). It has been observed that in EPEC the expression of intimin (Knutton *et al.*, 1997) and BFP (Puente *et al.*, 1996), and the secretion of the Esps (Kenny *et al.*, 1997a) are all affected by environmental signals. BFP expression is very low in normal laboratory media (e.g. Luria broth), but upon transfer to a defined tissue culture medium (e.g. Dulbecco's Modified Eagle Medium, DMEM) it is rapidly upregulated (Vuupio-Varkila and Schoolnik, 1991; Puente *et al.*, 1996). Puente *et al.* (1996) reported that BFP are optimally expressed between 37-38°C and during the exponential phase of growth. Full BFP expression required the presence of Ca²⁺ and was

significantly inhibited by the presence of ammonium. Puente *et al.* (1996) proposed that the temperature, high Ca^{2+} concentrations and low ammonium concentrations required for BFP expression would be encountered in the small intestine resulting in the microcolony formation which has previously been observed in jejunum biopsies of EPEC-infected children (Rothbaum *et al.*, 1982; 1983). In contrast EPEC which passed into the large intestine would encounter significantly higher concentrations of ammonium (Sleisenger and Fordtran, 1993) and BFP expression would be down-regulated leading to the shedding of the bacteria in the faeces. It has also been demonstrated that the expression of the EPEC secreted proteins is significantly enhanced under conditions reminiscent of those found in the gastrointestinal tract (Kenny *et al.*, 1997a). Thus secretion of EspB *in vitro* was found to be maximal at 37°C, pH7 and at physiological osmolarity; in addition there is a requirement for the presence of sodium bicarbonate, Ca^{2+} and millimolar Fe^{3+} . In light of the similarities in the transfer of EspB directly into tissue culture cells with the host cell contact-dependent transfer of the *Yersinia* virulence proteins YopE, YopH and YpkA (Pettersson *et al.*, 1996; Cornelis and Wolf-Watz, 1997) it remains to be seen if such a regulatory step is also required for secretion of EPEC proteins *in vivo*. As well as the enhancement of gene expression it has been observed that some EPEC virulence factors; i.e. intimin and EspA filaments are down regulated as the formation of the AE lesion progresses (Knutton *et al.*, 1997; 1998). The exact basis for this down regulation is unknown but it has been proposed that it is due to a feedback signal from the host cell (Knutton *et al.*, 1998) and, at least for intimin, has been shown to be dependent upon the EPEC Per regulator (see below) (Knutton *et al.*, 1997).

Whilst many of the environmental signals which regulate EPEC virulence have been determined the identification of the bacterial proteins which are responsible for this regulation is still in its early stages. Gomez-Duarte and Kaper (1995) identified three EAF plasmid-encoded genes, *perABC* (also known as *bfpTVW*; Tobe *et al.*, 1996), which enhanced the expression of the *eaeA* gene. The gene product of *perA* is predicted to be a member of the AraC transcription factor superfamily (Ramos *et al.*,

1990; Dorman, 1992) which have been shown to be involved in virulence gene regulation in numerous bacterial species (Finlay and Falkow, 1997). It has also been demonstrated that the Per regulator is required for the maximal expression of *espB* (Gomez *et al.*, 1993) and the *bfp* operon (Puente *et al.*, 1996; Stone *et al.*, 1996; Tobe *et al.*, 1996). Significantly the recent study by Beiber *et al.* (1998) demonstrated that a *perA* (*bfpT*) mutant was unable to induce diarrhoea in adult volunteers.

Because of the complex nature of the gene expression required for AE lesion formation it seems probable that other regulatory factors will also be involved. Indeed it has recently been demonstrated that the *fis* gene, which encodes the DNA-binding global regulator protein FIS (Finkel and Johnson, 1993; Dorman, 1995; González-Gil *et al.*, 1996), is required for AE lesion formation (Martin Goldberg, personal communication). Furthermore Elliot *et al.* (1998b) have described the identification of a gene for an HNS-like protein, designated *orf1*, within the lefthand end of the LEE of strain E2348-69; this class of histone-like proteins are known to be involved in gene regulation both at the level of DNA supercoiling (Higgins *et al.*, 1988; Finlay and Falkow, 1997) and also by direct binding to AT rich DNA regions within promoters (Dersch *et al.*, 1993; Zuber *et al.*, 1994).

1.7 Four stage model of EPEC pathogenesis.

In 1992 Donnenberg and Kaper devised a three stage model of EPEC pathogenesis based upon the current data concerning the bacterial and host components required for AE lesion formation. Though much of Donnenberg and Kaper's model remains valid a great deal of new information has been amassed in the ensuing six years which has allowed a new four stage model to be proposed (Knutton *et al.*, 1998; **figure 1.3**). *Stage 1*; environmental conditions encountered upon entering the host induce expression of specific EPEC virulence determinants including intimin, BFP and EspA filaments. The initial attachment of bacteria to host cells may then occur via EspA filaments, BFP, or by another as yet unknown mechanism. *Stage 2*; initial attachment of the bacteria induces translocation of effector proteins, including at least Tir and

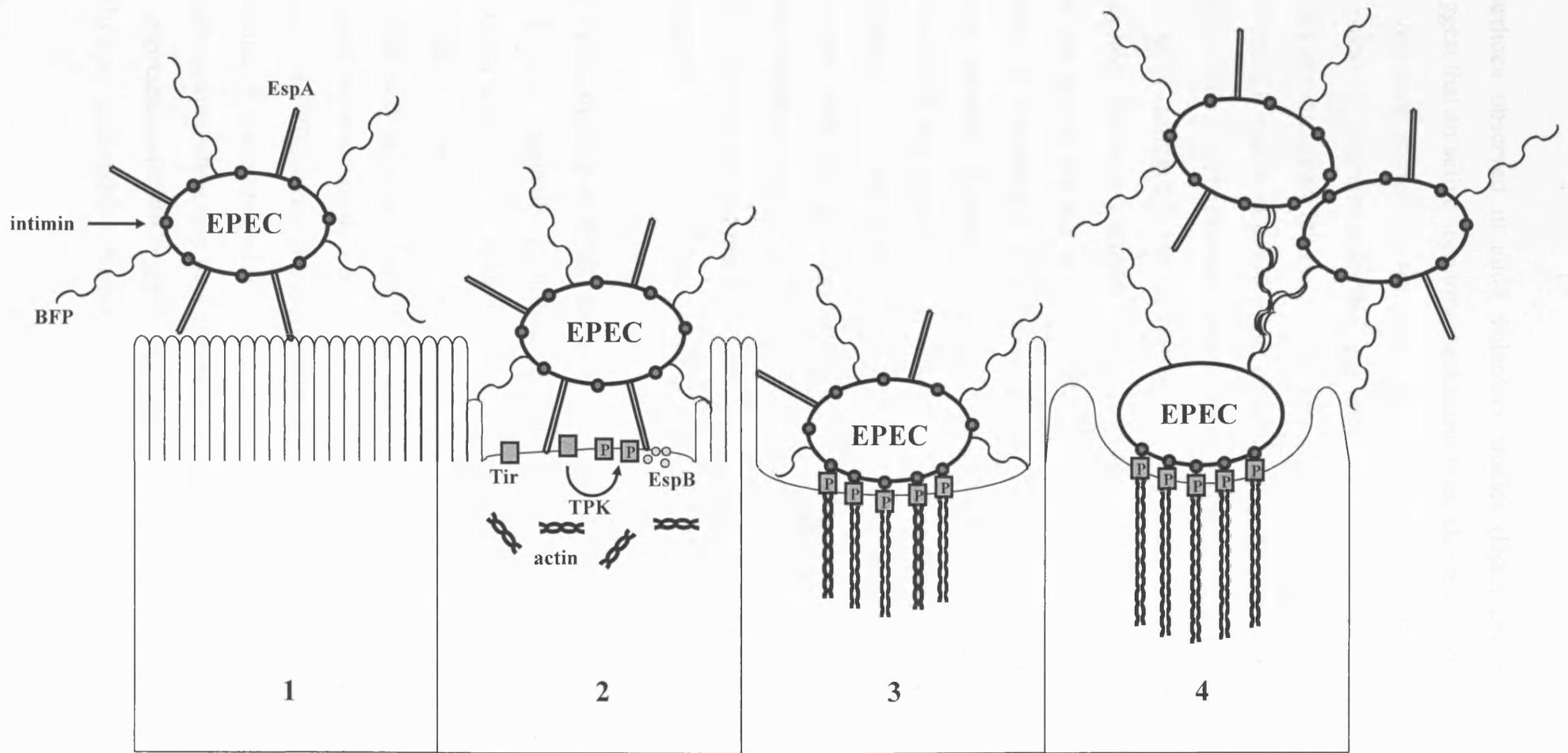
EspB, into the host cell via the type III secretion system and the EspA filaments. Signal transduction events then result in actin depolymerisation and effacement of the microvilli. Tyrosine phosphorylation of Tir occurs. *Stage 3*; binding of intimin to Tir brings the bacteria into intimate contact with the host cell. The intimin-Tir complex then forms the nucleation site for polymerisation of actin and other cytoskeletal elements. *Stage 4*; the expression of both intimin and EspA filaments is downregulated and they are gradually lost from the surface of the bacterial cell. The intense polymerisation of actin at the site of bacterial attachment forms the mature AE lesion often resulting in extended pedestal structures.

Although this new model of EPEC pathogenesis is in line with all of the current data it is unknown if the order of the events that it depicts is correct, and it is undoubtedly still incomplete; i.e. it does not address the roles of the other EPEC secreted proteins (EspD + others?). Furthermore whilst the expression of BFP has been shown to be essential for full virulence in humans (Beiber *et al.*, 1998) the long standing belief that they play a role in initial attachment of bacteria has recently been questioned. Hicks *et al.* (1998) demonstrated that, in contrast to the proven involvement of BFP in EPEC attachment to HeLa cells, they were not required for efficient initial attachment of bacteria during infections of *in vitro* organ culture of human small intestine biopsies. These data highlight the dangers inherent in the extrapolation of results derived from studies involving immortalised human epithelial cell lines directly to human disease.

1.8 Mechanism of EPEC-induced diarrhoea.

Diarrhoea essentially results from a net increase in the rate of fluid secretion compared to absorption from the lumen of the gut. Our current knowledge of the pathogenesis of EPEC allows for a number of possible mechanisms by which the responses observed in cultured human epithelial cells could lead to diarrhoea (Donnenberg *et al.*, 1997; Kaper, 1998). It is possible that the massive loss of the microvilli which results from EPEC infection (Moon *et al.*, 1983; Rothbaum *et al.*, 1982; Taylor *et al.*, 1986) could lead directly to diarrhoea due to malabsorption. However, the rapid onset of EPEC

Figure 1.3 Four stage model of EPEC pathogenesis. Growth in inducing conditions results in expression of adhesins (BFP and intimin) and the production of EspA filaments (stage 1). Initial attachment of EPEC via BFP or EspA filaments (or another mechanism ?) leads to translocation of EspB and Tir (plus other effector proteins ?) into the host cell. Signal transduction events promote actin disassembly, which results in the effacement of microvilli, and the induction of tyrosine kinase (TPK) which phosphorylates Tir (stage 2). Intimin binds to phosphorylated Tir and forms the nucleation site for polymerisation actin and other cytoskeletal proteins below the intimately attached bacteria (stage 3). EspA filaments and intimin are eliminated from the bacterial cell surface and further actin polymerisation forms the mature attaching and effacing lesion. Further bacteria may attach via BFP to form microcolonies (stage 4). (Adapted from Knutton *et al.*, 1998).



diarrhoea observed in adult volunteer studies (Donnenberg *et al.*, 1993a) would suggest that an active secretory mechanism must also be involved. In support of this it has recently been shown that the previously observed effects of EPEC infection upon ion transport in polarised epithelial cell monolayers (Knutton *et al.*, 1996; Stein *et al.*, 1996) are due to an increase in the secretion of Cl⁻ (Collington *et al.*, 1998). EPEC induced Cl⁻ secretion could occur either directly by PKC activation of ion channels, such as the cystic fibrosis transmembrane conductance regulator (CFTR) (Kaper, 1998), or indirectly via activation of the apical adenosine receptor by polymorphonuclear cells (Madara *et al.*, 1993) which are induced by EPEC to transmigrate into the lumen (Savkovic *et al.*, 1996). Increased secretion could also occur by a change in epithelial cell permeability due to relaxation of the cytoskeletal proteins which underlie the tight junctions. Tight junction permeability is known to be increased by the activation of myosin-light chain kinase (MLCK) (Hecht *et al.*, 1996), a process which has been observed during EPEC infection of both cultured human epithelial cells and gut biopsies (Baldwin, 1990; Manjarrez-Hernandez *et al.*, 1991; Manjarrez-Hernandez *et al.*, 1996; Yuhan *et al.*, 1997). (The topic of MLCK activation and epithelial cell permeability is covered in greater detail in the introduction to chapter 3).

1.9 Epidemiology of EPEC disease.

EPEC was originally identified as the cause of summer diarrhoea and remains primarily a disease of children under 2 years. Numerous studies have shown a strong correlation between the isolation of EPEC and diarrhoea in infants; this statistical correlation is especially high in the first 6 months of life (Levine and Edelman 1984). It is still unclear exactly why EPEC is restricted to infants; however, in the absence of reports of EPEC-induced traveller's diarrhoea in adults visiting areas where EPEC are endemic, it seems probable that it is not due to the acquisition of immunity. One possible explanation may be a physiological change after infancy, such as a decrease in the expression of EPEC-specific receptors (Nataro and Kaper, 1998). Transmission of EPEC, as with other diarrhoeagenic *E. coli*, is by the faecal-oral route, through

contaminated hands, food or even surfaces. The origin of infection is frequently an asymptomatic carrier and several studies have documented the spread of infection through hospitals and day care centres from a single individual (Edelman and Levine, 1984; Bower *et al.*, 1989; Wu and Peng, 1992). The numbers of EPEC bacteria required for transmission are unknown; it seems probable however that the figure will be much lower than the 10^8 to 10^{10} bacteria required to cause diarrhoea in adult human volunteer studies (Levine *et al.*, 1978). In several studies breastfeeding has been shown to protect from EPEC infection (Robins-Browne *et al.*, 1980; Blake *et al.*, 1993). This protection is mirrored by a significant inhibition of EPEC adhesion to HEp-2 cells *in vitro* upon addition of colostrum or breast milk (Cravioto *et al.*, 1991; Camara *et al.*, 1994)

Whilst EPEC still accounts for frequent outbreaks of infantile diarrhoea in paediatric wards or day care centres in the UK and United States (Rothbaum *et al.*, 1982; Paulozzi *et al.*, 1986; Bower *et al.*, 1989), the epidemics of high mortality summer diarrhoea with which it was once associated (Robins-Browne, 1987) no longer occur in developed countries. In a recent UK study of 167,630 cases of gastroenteritis in children under the age of 5 (conducted by the Communicable Disease Surveillance Centre, London) just 2.19% of all cases were diagnosed as *E. coli*. However, of the *E. coli* strains identified 44.3% belonged to the five classical EPEC serogroups O26, O55, O111, O126 and O128 (Crowley *et al.*, 1997; **figure 1.4**). It is unclear why the incidence of EPEC disease in developed countries has decreased since the 1940s and 1950s; however, as serological screening of *E. coli* is no longer routinely performed in either the UK or the United States, it is probable that the above figures underestimate the true extent of EPEC disease in these countries.

In contrast, in developing countries diarrhoea due to EPEC remains a major cause of infant morbidity and mortality. Numerous studies have demonstrated that the isolation of EPEC from infants with diarrhoea is significantly higher than that from healthy matched controls (Donnenberg, 1995; Levine and Edelman, 1984). Furthermore,

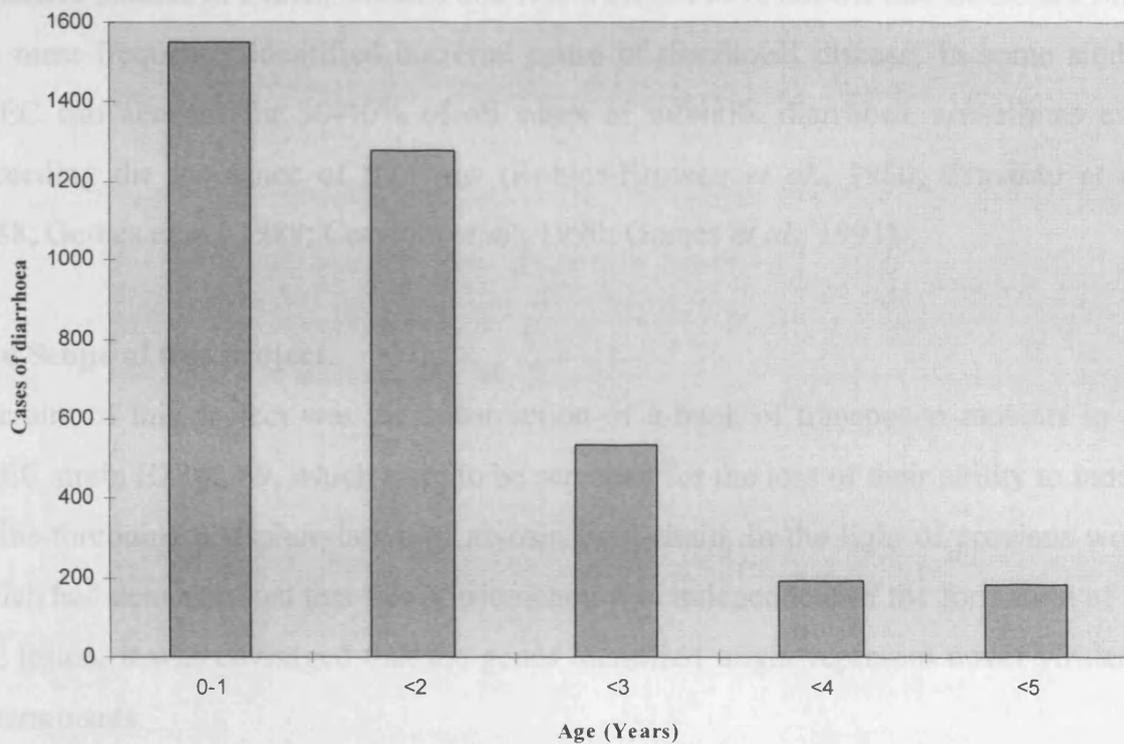


Figure 1.4 Incidence of gastroenteritis diagnosed as *E. coli* in children under 5 years of age. These figures are based upon data obtained from a study by the PHLS Communicable Disease Surveillance Centre (CDSC) for all confirmed incidences of gastroenteritis in children under 5 years of age in England and Wales for the period 1/1/1990 - 31/12/1994 (Crowley *et al.*, 1997). *E. coli* infections were identified as 2.19% (3660 cases) of all reported cases, with classical EPEC serogroups (O26, O55, O111, O126 and O128) and EHEC (VTEC O157:H7) representing 44.3% and 13.1% of this total, respectively.

extensive studies in Brazil, Mexico and South Africa have shown that EPEC are often the most frequently identified bacterial cause of diarrhoeal disease; in some studies EPEC can account for 30-40% of all cases of infantile diarrhoea, sometimes even exceeding the incidence of rotavirus (Robins-Browne *et al.*, 1980; Cravioto *et al.*, 1988; Gomes *et al.*, 1989; Cravioto *et al.*, 1990; Gomes *et al.*, 1991).

1.10 Scope of this project.

The aim of this project was the construction of a bank of transposon mutants in the EPEC strain E2348-69, which were to be screened for the loss of their ability to induce serine-threonine phosphorylation of myosin light chain. In the light of previous work, which had demonstrated that this phenomenon was independent of the formation of the AE lesion, it was envisaged that the genes identified might represent novel virulence determinants.

Chapter 2. Materials and Methods.

2.1 Bacterial strains. Details of the *E. coli* and EPEC strains used for cloning and phenotypic studies are summarised below. The *E. coli* strains used in the analysis of the distribution of *ompX* sequences are detailed in **Table 4.4**.

Strain	Relevant characteristics	Source or reference
E2348-69	O127:H6	Baldini <i>et al.</i> , 1983
E2348-69 Nal ^r	O127:H6, Nal ^r	J Kaper
E2348-69 Strep ^r	O127:H6, Strep ^r	S Clarke
MAR001	E2348-69, EAF ⁻ , Kan ^r	Knutton <i>et al.</i> , 1989
CVD206	E2348-69 Nal ^r , <i>eaeA</i>	Donnenberg and Kaper, 1991
CVD206:: Tn <i>phoA</i> #13	CVD206, <i>emtA</i> ::Tn <i>phoA</i>	Chapter 3
CVD206:: Tn <i>phoA</i> #55	CVD206, <i>bfpB</i> ::Tn <i>phoA</i>	Chapter 3
CVD206:: Tn <i>phoA</i> #137	CVD206, <i>emtA</i> ::Tn <i>phoA</i>	Chapter 3
CVD206:: Tn <i>phoA</i> #138	CVD206, <i>bfpB</i> ::Tn <i>phoA</i>	Chapter 3
CVD206:: Tn <i>phoA</i> #181	CVD206, <i>tag</i> ::Tn <i>phoA</i>	Chapter 3
RDH001	E2348-69 Nal ^r , IAP ⁺	Chapter 5
RDH003	CVD206, <i>ompX</i>	Chapter 5
RDH004	CVD206, <i>ompX</i> ::Tn <i>phoA</i>	Chapter 5
RDH005	E2348-69 Strep ^r , <i>emtA</i>	Chapter 5
RDH006	E2348-69 Strep ^r , IAP ⁺	Chapter 5
DH5 α	F ⁻ , Δ 80 <i>lacZ</i> M15, <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> (r _K ⁻ m _K ⁺), <i>relA1</i> , Δ (<i>lacIZYA-argF</i>), U169	Hanahan, 1983
SY327 λ <i>pir</i>	F ⁻ , <i>araD</i> , Δ (<i>lac</i> , <i>pro</i>), <i>argE</i> , <i>recA56</i> , <i>rif</i> , <i>nalA</i>	Miller and Mekalanos, 1988
SM10 λ <i>pir</i>	<i>thi1</i> , <i>thr1</i> , <i>leuB6</i> , <i>supE44</i> , <i>tonA21</i> , <i>lacY1</i> , <i>recA</i> ::RP4-2-Tc::Mu Km ^r	Simon <i>et al.</i> , 1983
CC118	<i>araD139</i> , Δ (<i>ara leu</i>)7697, Δ <i>lacX74</i> , <i>phoA20</i> , <i>galE</i> , <i>galK</i> , <i>thi</i> , <i>rpsE</i> , <i>rpoB</i> , <i>argE</i> (Am), <i>recA1</i> , Rif ^r	Manoil and Beckwith, 1985

table continued

HB101 F, *mcrB*, *mrr*, *hsdS20*(r_B⁻,m_B⁻), Hanahan, 1983
recA13, *supE44*, *ara14*, *galK2*, *lacY1*,
proA2, *rpsL20*(Smr), *xyl5*, *l-*, *leu*,
mtl1

2.2 Plasmids.

Plasmid	Relevant characteristics	Reference or source
pUC18	ColE1, <i>bla</i> , <i>lacI</i> , <i>lacZ</i>	Yannisch-Peron <i>et al.</i> , 1985
pBluescript SK-	ColE1, <i>bla</i> , <i>lacI</i> , <i>lacZ</i> , fl origin	Stratagene
pWKS30	pSC101 <i>ori</i> , <i>bla</i> , <i>lacZ</i>	Wang and Kushner, 1991
F42 <i>ts114 lacI3</i> zzf-2:: <i>TnphoA</i>	<i>ts</i> , <i>TnphoA</i>	Manoil and Beckwith, 1985
pGP704	<i>oriR6K</i> , <i>mobRP4</i> , <i>bla</i>	Miller and Mekalanos, 1988
pRT733	pGP704:: <i>TnphoA</i>	Taylor <i>et al.</i> , 1989
pCVD438	pACYC184 <i>eaeA</i>	Donnenberg and Kaper, 1991
pSW233	Tet ^r , Gm ^r , <i>sacRB</i> , <i>oriV</i> , <i>mobRP4</i>	Selbitschka <i>et al.</i> , 1993
pOUA14	<i>cat</i> , <i>bla</i> , <i>aphA-3</i>	Wang and Taylor, 1990
pMAR2	EAF plasmid from E2348-69	Baldini <i>et al.</i> , 1983
pJPN11	pMAR2::Kan ^r	Nataro <i>et al.</i> , 1987
p6.8.1	pMAR2 <i>bfpA</i> :: <i>TnphoA</i>	Donnenberg <i>et al.</i> , 1990
p13a	pUC18, 6.5 kb <i>Bam</i> HI <i>TnphoA</i> fragment from CVD206:: <i>TnphoA</i> #13	Chapter 3
p55a	pUC18, 7.4 kb <i>Bam</i> HI <i>TnphoA</i> fragment from CVD206:: <i>TnphoA</i> #55	Chapter 3
p137c	pUC18, 15 kb <i>Bam</i> HI <i>TnphoA</i> fragment from CVD206:: <i>TnphoA</i> #137	Chapter 3
p138b	pUC18, 6.7 kb <i>Bam</i> HI <i>TnphoA</i> fragment from CVD206:: <i>TnphoA</i> #138	Chapter 3
p181c	pUC18, 7.9 kb <i>Bam</i> HI <i>TnphoA</i> fragment from CVD206:: <i>TnphoA</i> #181	Chapter 3

table continued

pRDH7	pBluescript, 5.4 kb <i>Bam</i> HI fragment containing <i>bfp</i> region	Chapter 3
pRDH8	pBluescript, 6.3 kb <i>Bam</i> HI fragment containing <i>emtA</i> region	Chapter 3
pRDH8-12	pRDH8 <i>ompX</i> :: <i>TnphoA</i>	Chapter 4
pRDH10	<i>sacRB</i> , <i>mobRP4</i> , <i>ori</i> R6K, Cm ^r , Tet ^r	Chapter 5, Appendix 3
pRDH32	pBluescript, 310 bp <i>emtA</i> fragment	Chapter 5
pRDH32 Δ <i>Hind</i> III	pBluescript, 314 bp <i>emtA</i> Δ <i>Hind</i> III fragment	Chapter 5
pRDH33	pRDH10, 314 bp <i>emtA</i> Δ <i>Hind</i> III fragment	Chapter 5
pRDH35	pBluescript, 2.4 kb <i>orf2-ompX</i> fragment	Chapter 5
pRDH36	pRDH35 Δ <i>Sph</i> I	Chapter 5
pRDH38	pRDH10, 2.4 kb <i>orf2-ompX</i> fragment Δ <i>Sph</i> I	Chapter 5
pRDH39	pRDH8-12 14 kb subclone (<i>ompX</i> :: <i>TnphoA</i>) in pRDH10	Chapter 5
pRDH44	pWKS30, <i>TnphoA</i> downstream flanking DNA from strain CVD206 #181	Chapter 4
pRDH45	pBluescript, <i>TnphoA</i> downstream flanking DNA from strain CVD206 #181	Chapter 4

2.3 Bacterial culture media.

Luria Bertani broth (LUB) (Bertani, 1951; Lennox, 1955) consisted of the following; 1% (w/v) tryptone (Oxoid), 0.5% (w/v) yeast extract (Oxoid), and 0.5% (w/v) NaCl, adjusted to pH 7.0 with 1 N NaOH. Luria Bertani agar (LUA) consisted of LUB solidified with 1.5% agar (Difco). LUB and LUA were sterilised by autoclaving at 1 kg/cm² for 15 minutes and stored at room temperature. When required LUA was re-melted using a microwave oven, allowed to cool to 50°C before addition of antibiotics, then poured into sterile 90 mm plastic Petri dishes (approximately 20 ml per plate). Plates were dried by incubation at 37°C and then stored at 4°C.

Antibiotics were added to media, as appropriate, at the following final concentrations (unless otherwise indicated in the text); ampicillin (100 µg/ml), nalidixic acid (50 µg/ml), kanamycin (50 µg/ml), chloramphenicol (20 µg/ml), tetracycline (10 µg/ml), rifampicin (10 µg/ml) and streptomycin (100 µg/ml). The indicators 5-chloro-4-bromo-3-indoyl phosphate (XP; 40 µg/ml; Sigma) and 5-chloro-4-bromo-3-indoyl-β-D-galactoside (X-gal; 20 µg/ml; Sigma) were added to LUA plates to select for colonies expressing alkaline phosphatase or β-galactosidase respectively. Iron replete or iron restricted media were prepared by addition of 40 µM FeCl₂ or 300 µM 2,2'-dipyridyl (Sigma) respectively to LUB or LUA.

Sucrose broth was prepared initially as 1% (w/v) tryptone (Oxoid), 0.5% (w/v) yeast extract (Oxoid), adjusted to pH 7.0 with 1 N NaOH. Subsequent to autoclaving it was made up to 6% (w/v) sucrose using sterile 40% (w/v) sucrose stock solution. Sucrose agar was prepared by addition of 1.5% (w/v) agar to sucrose broth.

2.4 Growth, storage and preservation of bacterial strains.

Bacterial stock strains were routinely maintained by weekly replating upon LUA plates with appropriate antibiotics. Stock plates were replaced as required from frozen stocks maintained at -80°C (frozen stocks were prepared by pelleting bacteria from 1 ml of fresh overnight culture, resuspending in 1 ml of a cryoprotectant medium (25% (v/v) glycerol, 75% (v/v) LUB), and storing in plastic screw top cryo-tubes (Sarstedt).

2.5 EPEC interactions with cultured human epithelial cell lines.

2.5.1 Maintenance of tissue culture cells. The transformed cell lines HEp-2 (originally derived from human laryngeal carcinoma tissue; Toolan, 1954a; 1954b) and Caco-2 (of colonic carcinoma origin; Fogh *et al.*, 1977) were obtained from Dr Stuart Knutton, Institute of Child Health, Birmingham. Cell lines were routinely cultured as monolayers in 80 cm² tissue culture flasks (Nunc) at 37°C in Dulbecco's Modified

Eagle Medium, without pyruvate (DMEM; GibcoBRL), containing 10% (v/v) (20% for Caco-2 cells) heat inactivated, mycoplasma screened foetal calf serum (FCS; GibcoBRL) under an atmosphere of 5% CO₂ in air. Confluent monolayers were washed twice in 10 ml sterile phosphate buffered saline (PBS; 10mM phosphate containing 0.14 M NaCl, pH 7.4), then 2 ml of PBS and 1 ml of trypsin (0.25% in 5 mM KCl, 2.6 mM Na₂HCO₃, 115 mM NaCl, 5.5 mM Glucose, 0.05% Phenol Red; GibcoBRL) were added and monolayers incubated at 37°C until cells showed the appearance of rounding up (1-3 min). Cells were dislodged from the monolayer by striking the side of the flask against the palm of the hand. 10 ml of DMEM/FCS was added to the flask and the cells gently resuspended. This cell suspension was used to seed fresh flasks containing DMEM/FCS, at maximum split ratios of 1:10 for HEP-2 cells and 1:4 for Caco-2 cells. monolayers were grown in 5% CO₂ at 37°C until confluent, with changes of medium every 2 days. For coverslip cultures cell suspensions were seeded into bacteriological Petri dishes containing sterile 12 mm² glass coverslips in 25 ml of DMEM/FCS. Amounts of suspension added were designed to ensure the correct level of confluence i.e. 40-50% for adhesion assays or FAS tests. Once Caco-2 cells reach confluence they continue to differentiate in a manner characteristic of mature enterocytes, elaborating microvilli and forming domes; thus for the investigation of bacterial-cell interactions Caco-2 cells were used only after they had fully differentiated, usually a few days post-confluence (Pinto *et al.*, 1983). For storage of cell lines, trypsinised cell suspensions were gently pelleted by centrifugation at 1000 rpm for 2 min in 10 ml conical bottomed plastic tubes (Sterilin). Pellets were resuspended in 1 ml of storage medium (10% (v/v) FCS and 10% (v/v) DMSO in DMEM) by gentle agitation, transferred to 2 ml screw top plastic cryo-tubes (Sarstedt) and placed in an insulated freezing box. The temperature was reduced slowly (at approximately -1°C/min) to -80°C by placing the box initially at -20°C for 1 hour and then at -80°C for 1 hour. Tubes were finally stored under liquid nitrogen where cells can remain viable for up to 5 years. Cells were revived by removal of tubes from liquid nitrogen, and fast thawing at 37°C; they were then immediately seeded into flasks containing DMEM/FCS and incubated at 37°C under 5% CO₂.

2.5.2 HEp-2 adhesion assay. This is a variation of the original method used by Cravioto *et al.* (1979) for HEp-2 cell adhesion assays. Cover slips seeded with subconfluent Hep-2 cells were placed in 6 well tissue culture plates in 2 ml of DMEM, containing 0.5% (^{w/v}) D-mannose and preincubated at 37°C under 5% CO₂ for 1 hour. Individual wells were infected with 10⁸ bacteria (approximately 40 µl of an overnight LUB culture) and incubated at 37°C under 5% CO₂ for 3 hours. Non-adherent bacteria were removed by washing the monolayers 3 times in PBS, the cells fixed for 10 min with methanol and then stained with Giemsa stain (0.4%, Sigma). Stained coverslips were dried, mounted on glass slides and examined for adherent bacteria under a light microscope (1000x, oil immersion).

2.5.3 Immunofluorescence microscopy.

2.5.3.1 Fluorescent actin staining (FAS) assay. This is essentially the method as described by Knutton *et al.* (1989). Cover slips seeded with subconfluent HEp-2 cell monolayers were placed into 6 well tissue culture plates in 2 ml of DMEM, containing 0.5% (^{w/v}) D-mannose, and preincubated for 1 hour at 37°C under 5% CO₂. Each well was infected with 10⁸ bacteria (approximately 40 µl of an overnight LUB culture). Infected monolayers were incubated at 37°C under 5% CO₂ for 3 hours, washed 3 times in PBS to remove non-adherent bacteria, and then fixed in 3% formalin for 20 min. Cells were permeabilised with 0.1% Triton X-100 in PBS for 5 min, washed 3 times in PBS and then stained for 20 min in the dark with either 10 µg/ml fluorescein isothiocyanate-phalloidin or 0.1 µg/ml Texas Red isothiocyanate-phalloidin (FITC-phalloidin, TRITC-phalloidin; Sigma). Samples were washed 3 times in PBS and mounted in antifade (50% glycerol in PBS containing 1 mg/ml *p*-phenylenediamine). Infected monolayers were examined by incident light fluorescence and phase contrast microscopy using a Zeiss Axiophot fluorescence microscope.

2.5.3.2 Detection of bundle-forming pili (BFP) using fluorescent antibodies. HEp-2 cells were seeded onto coverslips, infected with EPEC, fixed and permeabilised essentially as described for the FAS assay above. After washing three times in PBS individual coverslips were treated with 40µl of rabbit anti-BFP antibody for 45 min (1/40 dilution in 3% BSA/PBS; Giron *et al.*, 1993; see below). The coverslips were then washed three times in PBS to remove non-adherent primary antibody and treated with 25µl of donkey anti-rabbit FITC conjugate for 45 min (1/50 dilution in 3% BSA/PBS; Jackson Laboratories). Finally the samples were washed three times in PBS and mounted on slides in antifade (as above). Infected monolayers were examined by incident light fluorescence and phase contrast microscopy using a Zeiss Axiophot fluorescence microscope.

Polyclonal antibody generated in rabbits against purified BFP from EPEC isolate B171 was generously provided by Dr Jorge A. Giron, Laboratorio de Patogenicidad, Ciudad Universitaria, Puebla, Mexico.

2.5.4 EPEC-induced protein phosphorylation in cultured human epithelial cells.

This is an adaptation of the method used to investigate protein kinase C activity in EPEC infected tissue culture cells (Baldwin, 1990; Baldwin *et al.*, 1990). Fully differentiated Caco-2 cells were used for the initial stages of this work because they show a much lower level of inherent protein kinase activity; however HEp-2 cells were used for secondary screening and subsequent testing due to the greater ease of culturing. Cultured cells were seeded in 96 or 24 well tissue culture plates and allowed to grow to confluence (and to differentiate in the case of Caco-2) in DMEM/20% FCS at 37°C in 5% CO₂ in air. Prior to the protein phosphorylation experiments differentiated cell monolayers were phosphate-starved overnight in phosphate-free Minimal Essential Medium (MEM; ICNFlow Laboratories) supplemented with 1% FCS. Starved monolayers were then incubated in MEM (500 µl in 24 well plates, 100 µl in 96 well plates) containing 0.1 mCi/ml of carrier free ³²P_i phosphate (10 mCi/ml; Amersham LIFE SCIENCE) for 4 hours to label the intracellular pool of ATP.

Individual wells were infected with overnight LUB cultures of *E. coli* (5 µl/well in 24 well plates, 1 µl/well in 96 well plates) and the plates incubated for a further 3 or 6 hours. To terminate phosphorylation reactions and remove unattached bacteria the cells were washed 3 times in ice cold PBS. The monolayers were resuspended in 2x SDS-PAGE loading buffer (200 µl/well for 24 well plates, 50 µl/well for 96 well plates; section 2.16.2) using a 1 ml Gilson pipette tip and were transferred to Eppendorf tubes. Samples were denatured by heating to 100°C for 10 min and then subjected to electrophoresis on 11% or 13% SDS polyacrylamide gels (section 2.16.1). Any remaining samples were stored at -20°C. After electrophoresis protein bands were visualised by staining with Coomassie blue (section 2.16.1). The gel was rinsed briefly in distilled water and dried onto 3MM paper (Whatman) under vacuum using a Biorad Model 583 Gel Dryer. Phosphorylated protein bands were visualised by autoradiography with Cronex[®] Medical X-ray film (Dupont) for 4 to 16 hours.

2.5.5 Gentamicin protection (invasion) assay. HEp-2 cells were seeded to 25-50% confluence in 24 well tissue culture plates and allowed to grow in DMEM/10% FCS until just confluent. Culture medium was replaced with 0.5 ml/well of DMEM containing 0.5% (^{w/v}) D-mannose and the plates preincubated for 1 hour at 37°C under 5% CO₂. Individual wells were infected with approximately 2×10^7 bacteria (5 µl of an overnight LUB culture) and then incubated at 37°C for 3 hours. For analysis of cell-associated bacteria the monolayers were taken at this point, whilst for analysis of invaded bacteria the monolayers were washed 3 times in PBS, the media replaced with DMEM containing 100 µg/ml gentamicin and incubation continued for a further 3 hours. After incubation monolayers were washed 3 times with PBS and the bacteria released by solubilisation of the HEp-2 cells by addition of 1 ml of 1% Triton X-100 in PBS and incubation for 20 min at room temperature. Bacterial numbers in samples were enumerated by serial dilution and plating onto LUA plates. Dilutions were performed in duplicate and plates with approximately 10^2 colonies were chosen for counting. All assays were performed with triplicate samples.

2.6 Autoaggregation assay.

Bacteria were inoculated (20 μ l of an overnight LUB bacterial culture per well) into individual wells of a 24 well tissue culture plate which contained either 1 ml of LUB or 1 ml of DMEM containing 1% mannose ($^w/v$), which had been preincubated at 37°C under 5% CO₂ for 90 min. Samples were incubated for a further 3 hours then examined using an inverted microscope and bacterial morphology recorded using photography.

2.7 ³⁵S labelling of EPEC secreted proteins (EspA, EspB, and EspC).

To each well of a six well tissue culture plate were added 2 ml of methionine-free MEM containing 2.5 μ l/ml of ³⁵S methionine (15 mCi/ml; Amersham LIFE SCIENCE) and the plate preincubated at 37°C under 5% CO₂ for 90 min. Each well was inoculated with 2 x 10⁸ bacteria (80 μ l of an overnight LUB culture) and the samples incubated at 37°C for a further 3 hours. 1.5 ml of each culture was transferred to an Eppendorf tube, the bacteria pelleted by centrifugation at 10,000 x g for 15 min and 1 ml of the supernatant transferred to a fresh tube. Each tube was made up to a final concentration of 10% TCA by addition of 110 μ l of ice cold 100% TCA ($^w/v$) and samples incubated on ice for 20 min. Precipitated proteins were pelleted by centrifugaion at 10,000 x g for 15 min, supernatants discarded and the pellet allowed to air dry. The dried pellet was neutralised by addition of a small volume (typically 2-3 μ l) of a saturated Tris solution and then resuspended in 2x SDS-PAGE sample buffer (section 2.16.2) to a final volume of 20 μ l. 10 μ l samples were denatured by boiling for 5 min and separated on 13% glycine-SDS-PAGE minigels. After electrophoresis the gel was fixed in 25% methanol ($^v/v$), 10% glacial acetic acid ($^v/v$) for 30 min, then soaked in 1 M sodium salicylate pH 6.0 for 1 hour to enhance the ³⁵S signal by fluorography (Chamberlain, 1979). Finally the gel was rinsed in distilled water and dried onto 3MM paper (Whatman) under vacuum using a Biorad Model 583 Gel Dryer. ³⁵S methionine-labelled secreted proteins were then visualised by autoradiography with Cronex[®] Medical X-ray film (Dupont).

2.8 Conjugation.

Overnight LUB cultures of donor and recipient strains were diluted 50-fold in LUB without antibiotics and grown with vigorous shaking until they reached an A_{600} of 0.6. A 1 ml aliquot of each strain was pelleted into the same Eppendorf tube and the supernatants discarded. The resultant pellet was resuspended in 100 μ l of LUB and gently layered onto a 25 mm nitrocellulose filter disk (0.45 μ m pore size; Millipore) already placed on a antibiotic free LUA plate. The conjugation mixture was incubated at 37°C for three hours and then the bacteria were washed free from the filter into 1 ml of LUB. Dilutions of the conjugation mixture were spread on LUA plates containing appropriate antibiotics to select for the transconjugants and incubated overnight at 37°C.

2.9 DNA manipulations.

2.9.1 Stock solutions for DNA work.

20x SSC	- 3 M NaCl, 0.3 M sodium citrate adjusted to pH 7.0 with NaOH.
TE	- 10 mM Tris-HCl, 1 mM EDTA pH 7.4.
Chloroform/isoamyl alcohol	- 96% chloroform (v/v), 4% isoamyl alcohol (v/v)
Phenol/chloroform	- 50% phenol (v/v), 50% chloroform/isoamyl alcohol (v/v), 0.05% (w/v) hydroxyquinoline.
Solution I	- 50 mM glucose, 10 mM EDTA, 25mM Tris-HCl pH 8.0.
Solution II	- 0.2M NaOH, 1% SDS.
Solution III	- 3M potassium acetate, adjusted to pH 5.0 with glacial acetic acid.
Depurinating solution	- 0.25 M HCl
Denaturing solution	- 1M NaOH, 1.5 M NaCl
Neutralising solution	- 1.5 M NaCl, 1M Tris-HCl pH 7.5

2.9.2 Restriction enzyme digestions and ligation of DNA. Ligations and restriction digestions were carried out using GibcoBRL enzymes and buffers according to the recommendations of the manufacturers, unless indicated otherwise in the text.

2.9.3 Modification of DNA fragment termini generated by restriction enzyme digestion. Following digestion with a large number of commonly used restriction enzymes DNA fragments are left with short single stranded 3' or 5' overhangs ("sticky ends"). For constructions which required the ligation of non-homologous sticky ends the 3' and 5' overhangs were blunted by action of the large fragment of *E. coli* DNA polymerase I (Klenow) prior to blunt ended ligation.

2.9.3.1 Filling in of 3' overhangs. In order to blunt 3' overhangs the 5'-3' polymerase activity of Klenow was utilised to fill in the missing nucleotides. 50-500 ng of DNA fragment were incubated with 1x ligase buffer (GibcoBRL), 1 mM dNTPs and 1 µl of Klenow (4 U/µl; GibcoBRL) at 25°C for 30 min. After filling in, the fragment was either gel purified, or extracted with phenol/choroform and chloroform/isoamyl alcohol, and then ethanol precipitated before ligation.

2.9.3.2 Removal of 5' overhangs. In order to blunt 5' overhangs the 3'-5' exodeoxyribonuclease activity of Klenow was used to remove the single stranded region. 50-500 ng of DNA fragment were incubated with 1x ligase buffer and 1 µl of Klenow at 37°C for 15 min to allow digestion of the 5' overhang. Then 1 mM dNTPs were added and the reaction was continued at 25°C for 30 min to allow filling in of any 3' overhangs caused by excessive exonuclease activity. The fragment was then either gel purified, or extracted with phenol/choroform and chloroform/isoamyl alcohol, and ethanol precipitated before ligation.

2.9.4 Electrophoresis of DNA. Restriction fragments or PCR products were analysed by agarose gel electrophoresis in 0.8% - 2.0% agarose (Boehringer Mannheim) gels containing 500 ng/ml ethidium bromide in 0.04 M Tris-acetate, 1 mM EDTA. DNA

bands were visualised by UV illumination and recorded by photography with Polaroid 695 film or by image capture using a UVP video camera and Mitsubishi Video Copy processor.

2.9.5 Isolation of DNA fragments from agarose gels using Polymer wool™. This method is a cheap, homemade, but very fast and effective alternative to the use of low melting point agarose or gel solubilisation DNA purification systems such as QIAEX (QIAGEN). A sieve tube was constructed by puncturing or melting a small hole in the bottom of a 0.5 ml Eppendorf tube and packing it (to a level equivalent to 50 µl of liquid) with Polymer wool™ (sold for use in aquarium filter systems; Interpet Ltd, Dorking, Surrey). DNA fragments of interest were cut from the agarose gel using a scalpel, trimmed of excess agarose and placed into the sieve tube. The tube was then placed within a larger 1.5 ml Eppendorf tube (which had had its lid removed to facilitate this) and centrifuged at 10,000 x g for 5 min. The liquid which had passed through the polymer wool into the larger Eppendorf was collected and transferred to a fresh tube; if significant lumps of agarose were still visible in the sieve tube it was recentrifuged and any further liquid added to the fresh tube. The recovered DNA solution was immediately extracted once with phenol/chloroform, then once with chloroform/isoamyl alcohol. Finally the DNA was precipitated by addition of two volumes of ethanol and pelleted by centrifugation at 10,000 x g for 10 min; the recovery of small amounts of DNA was significantly improved by addition of 1µl of 10 mg/ml tRNA to the ethanol precipitation to act as a carrier molecule. Pellets were washed in 70% ethanol, dried under vacuum and resuspended in TE or DW.

2.9.6 DNA preparation.

2.9.6.1 Small scale plasmid DNA preparation (mini-prep). Small amounts of plasmid DNA (10-20 µg for high copy number plasmids) sufficient for restriction enzyme analysis were prepared by an alkaline lysis method based on that of Birnboim and Doly (1979). Cells from 1.5 ml of an overnight LUB culture containing

appropriate antibiotics were harvested in Eppendorf tubes by centrifugation at 10,000 x g for 5 min. The cells were resuspended in 100 µl of solution I and then lysed by the addition of 200 µl of solution II and mixing by gentle inversion. 150 µl of solution III were added and the tube vortexed inverted for 10 seconds. The precipitate was pelleted by centrifugation at 10,000 x g for 5 min and the supernatant removed to a fresh Eppendorf tube. 500 µl of phenol/chloroform was added to the tube, the contents mixed by vortexing and the two layers separated by centrifugation at 10,000 x g for 5 min. The upper aqueous layer was removed to a fresh tube and similarly extracted with chloroform/isoamyl alcohol. The DNA was precipitated by addition of 800 µl of 95% ethanol and harvested by centrifugation at 10,000 x g for 15 min. The supernatant was removed by aspiration, the DNA pellet rinsed in 70% ethanol and then dried under vacuum. The DNA pellet was resuspended in 20 µl of sterile distilled water containing 0.1 mg/ml RNaseA.

2.9.6.2 Large scale CsCl preparation of plasmid DNA. This is a scaled up modification of the mini-prep DNA preparation including a CsCl gradient purification step which produces large quantities of DNA sufficiently pure for sequencing. 50 ml of an overnight LUB culture, grown with appropriate antibiotics, were harvested in Falcon tubes by centrifugation at 10,000 x g for 10 min. The pellet was resuspended in 5 ml of Solution I, 10 ml Solution II were added and cell lysis was achieved by repeated gentle inversion. 7.5 ml of Solution III were added, the mixture was shaken vigorously and the precipitate was pelleted by centrifugation at 10,000 x g for 10 min. The supernatant was transferred to a fresh tube, 30 ml of 95% ethanol added and the DNA/RNA precipitate harvested by centrifugation at 10,000 x g. The pellet was allowed to air dry for 20 min and then redissolved in sterile distilled water to a final volume of 1.6 ml. 1.76 g of CsCl and 66 µl of ethidium bromide (10 mg/ml) were added and mixed until the CsCl dissolved. The mixture was transferred to heat sealable bell-top centrifuge tubes (Beckman) and the CsCl gradient was formed by centrifugation at 100,000 rpm using a Beckman TL-100 benchtop ultracentrifuge for a minimum of 4 hours at 20°C. DNA bands were visualised, if necessary using UV, and

the lower band (contaminating chromosomal DNA is less dense than plasmid DNA in the presence of ethidium bromide) recovered to an Eppendorf tube using a needle and syringe. The ethidium bromide was removed by repeated extraction with water/CsCl-saturated isobutanol until there was no longer any pink colouration. An equal volume of sterile distilled water was added to the DNA solution and it was precipitated by addition of 0.6 volumes of isopropanol. The DNA was pelleted by centrifugation at 10,000 x g for 10 min, the pellet was washed twice with 70% ethanol, dried under vacuum and redissolved in 50 µl of TE.

2.9.6.3 Small scale preparation of chromosomal DNA. For the purposes of Southern blot analysis chromosomal DNA was prepared by an adaptation of the method of Chen and Kuo (1993). 1.5 ml of an overnight bacterial culture was harvested by centrifugation at 10,000 x g for 5 min. The pellet was resuspended in 200 µl of lysis buffer (40 mM Tris-acetate pH 7.8, 20 mM sodium acetate, 1 mM EDTA, 1% (^w/_v) SDS) by vigorous pipetting. To precipitate cell protein and debris 66 µl of 5 M NaCl were added and mixed well, and the viscous mixture was then centrifuged at 10,000 x g for 10 min at 4°C. The supernatant was removed to a fresh tube, an equal volume of phenol/chloroform added and mixed to a milky solution by gentle repeated inversion. The layers were separated by centrifugation for 3 min at 10,000 x g; the aqueous fraction was transferred to a fresh tube and similarly extracted with chloroform/isoamyl alcohol. DNA was precipitated with ethanol, washed twice with 70% ethanol, dried under vacuum and resuspended in 20 µl of TE containing 0.1 mg/ml RNaseA.

2.9.6.4 Large scale CsCl preparation of chromosomal DNA. This is an adaptation of the method of Sambrook *et al.* (1989), including an extra CsCl purification step, for the preparation of chromosomal DNA suitable for cloning or library construction. 50 ml of an overnight bacterial culture was harvested in polypropylene tubes by centrifugation for 10 min at 10,000 x g and the pellet resuspended in 10 ml of TES (10 mM Tris-HCl pH 7.6, 1 mM EDTA, 0.15 M NaCl). To lyse the cells 2 ml of 20% SDS and 100 µl of

proteinase K (20 mg/ml) were added, the solution mixed and then incubated at 37°C until significant clearing occurred (approximately 60 min). Protein was removed by extraction with an equal volume of phenol/chloroform and centrifugation for 15 min at 10,000 x g. The viscous aqueous phase was removed to a fresh tube using a plastic Pasteur pipette and the DNA precipitated by addition of a 1/10 volume of 3 M sodium acetate pH5.2 and 2 volumes of 95% ethanol. The precipitated DNA was spooled onto a glass Pasteur pipette "hook", and rinsed briefly in 70% ethanol, excess ethanol was allowed to drain, and the DNA was transferred to a tube containing 0.8 ml of TE and 0.8 ml of 3% N-lauryl sarcosine. Once the DNA had fully redissolved (this may take several hours), 1.76 g of CsCl and 66 µl of ethidium chloride were added and mixed until the CsCl has completely dissolved. The mixture was transferred to heat sealable bell-top centrifuge tubes (Beckman) and the CsCl gradient was formed by centrifugation at 100,000 rpm using a Beckman TL-100 benchtop ultracentrifuge for a minimum of 4 hours at 20°C. The large chromosomal band was recovered using a wide bore needle and a syringe and transferred to an Eppendorf tube. The ethidium bromide was removed by repeated extraction with water/CsCl-saturated isobutanol, the DNA diluted with an equal volume of sterile distilled water and precipitated by addition of 0.6 volumes of isopropanol. The DNA was pelleted by centrifugation at 10,000 x g for 10 min, the pellet washed twice with 70% ethanol, dried under vacuum and redissolved in 500 µl of TE.

2.9.7 Transformation of *E. coli* with plasmid DNA. Two different transformation methods were used depending upon the efficiency of transformation required. For simple transfer of plasmid DNA into *E. coli* K-12 strains the calcium chloride method was used, whereas for transformation of ligation products or transferring plasmid DNA into EPEC strains electroporation was the preferred method.

2.9.7.1 Calcium chloride method (Dagbert and Ehrlich, 1979). An overnight culture of *E. coli* was diluted 50-fold in LUB and grown with vigorous shaking at 37°C until they reached an A₆₀₀ of 0.6. 25 ml of culture were harvested by centrifugation at

10,000 x g for 10 min at 4°C and the pellet resuspended in 25 ml of ice-cold, sterile, 0.1M MgCl₂. The cells were reharvested and the bacterial pellet resuspended in 25 ml of ice-cold, sterile 0.1 M CaCl₂ and incubated on ice for 30 min. Finally the cells were reharvested and resuspended in 1 ml of ice-cold, sterile 0.1M CaCl₂ and maintained on ice. 100 µl aliquots of competent cells were mixed with plasmid DNA, incubated on ice for 5 min and then heat shocked by incubation in a 42°C water bath for 90 seconds. Transformed cells were recovered by addition of 1 ml of LUB and incubated at 37°C for 45 min, to allow expression of resistance determinants, before plating onto antibiotic containing LUA plates.

2.9.7.2 Electroporation. Plasmid DNA was introduced into recipient cells using high voltage electroporation by the method of Dower *et al.* (1988). Electrocompetent *E. coli* were prepared essentially as described by Sambrook *et al.* (1989). An overnight culture of *E. coli* was diluted 50-fold in LUB and grown with vigorous shaking until it reached an A₆₀₀ of 0.6. 100 ml of cells were harvested by centrifugation at 10,000 x g for 10 min at 4°C, then washed 3 times in 100 ml of ice-cold sterile distilled water. Finally the salt free bacterial pellet was resuspended in 200 µl of ice cold sterile distilled water and stored on ice. 40 µl aliquots of washed cells were transformed with salt free DNA in 0.2 cm gap electroporation cuvettes (Biorad) in a Biorad Gene Pulser™ set to 1.5 kV with a resistance of 1000 Ω and a capacitance of 25 µF. Transformed bacteria were recovered by immediate resuspension in 1 ml of SOC medium (2% tryptone, 0.5% yeast extract, 0.05% NaCl, 2.5 mM KCl, 10 mM MgCl₂, 20 mM glucose) and incubation at 37°C for 45 min, before plating onto antibiotic containing LUA plates.

2.10 DNA labelling and detection.

2.10.1 Transfer of DNA from agarose gels to nylon membranes (Southern blot).

Southern blotting of DNA to Hybond™-N membrane (Amersham LIFE SCIENCE) was performed essentially as described by Sambrook *et al.* (1989). Transferred DNA

was immobilised on the membrane using a UV StratalinkerTM (Stratagene) set at an energy of 70,000 μ Joules.

2.10.2 Non-radioactive detection of DNA. For detection of DNA bands on Southern blots the Fluorescein Gene ImagesTM labelling and detection system (Amersham LIFE SCIENCE) was used. The system uses fluorescein-11-dUTP to label the probe and after hybridisation and stringency washes the blot is incubated with an anti-fluorescein alkaline phosphatase conjugate which binds hybridised probe. After washes to remove unbound conjugate a chemiluminescent detection reagent is added; this provides both a light producing chemical (1,2-dioxetane phosphate which is cleaved by the alkaline phosphatase) and an enhancer which amplifies the produced light which can then be visualised by exposure of the blot to X-ray film.

2.10.2.1 Labelling of DNA probes. The standard labelling reaction as described here (using 50 ng of DNA) provides enough probe for a number of hybridisations (see Section 2.10.2.2). 50 ng of DNA, in a volume of approximately 20 μ l of water or TE, were denatured by boiling for 5 min, chilled on ice for 5 min, and then the labelling reaction was assembled by addition of reagents in the following order:

Nucleotide mix	10 μ l
Primers	5 μ l
Denatured DNA	50 ng
Klenow	1 μ l
H ₂ O	to 50 μ l

(Primers, nucleotide mix, and Klenow are supplied with the kit)

The reaction was mixed by gentle pipetting and then incubated at 37°C for a minimum of 1 hour (or incubated overnight at 25°C). The reaction was terminated by addition of EDTA to a final concentration of 20 mM. Probes were denatured by boiling for 5 min prior to addition to hybridisations. Surplus probe is stable at -20°C for up to 6 months.

2.10.2.2 Hybridisation of probes to Southern blots. All hybridisations were performed in glass hybridisation tubes using a Hybaid mini hybridisation oven. Blots were soaked in 5x SSC and rolled into nylon sheets (to allow ease of insertion and removal from the hybridisation tubes) then prehybridised in preheated hybridisation buffer (5x SSC, 0.1% (w/v) SDS, 5% (w/v) dextran sulphate, 1/20 vol of liquid block (supplied with kit); 0.3 ml/cm² of blot) at 65°C for a minimum of 1 hour. Denatured probe (10 µl for approx 100 cm² blot) was added directly to the hybridisation buffer and incubated at 65°C for a minimum of 6 hours and typically overnight. The hybridisation buffer was decanted and the blot subjected to stringency washes at 65°C; 1 x 15 min with preheated 1x SSC, 0.1% (w/v) SDS, and 1 x 15 min with preheated 0.5x SSC, 0.1% (w/v) SDS (wash volumes typically 2-5 ml/cm² membrane).

2.10.2.3 Detection of hybridised probe. Following the stringency washes, blots were washed briefly in an excess of diluent buffer (100 mM Tris-HCl, 300 mM NaCl pH 9.5, freshly autoclaved) and then incubated with gentle agitation at room temperature in a 1 in 10 dilution of liquid block (supplied with kit) in diluent buffer for 1 hour. After rinsing briefly in diluent buffer again, the blot was incubated in diluted anti-fluorescein-AP conjugate (conjugate diluted 5000-fold in 0.5% (w/v) BSA in diluent buffer; 0.3 ml/cm²) with gentle agitation at room temperature for 1 hour. Unbound conjugate was removed by washing for 3 x 10 min in 0.3% (v/v) Tween-20 (Sigma) in diluent buffer. Finally the blot was rinsed again in diluent buffer prior to development. The blot was placed sample side up onto a clean, over-sized piece of SaranWrap in a fume cupboard and sprayed with the dioxetane detection reagent (supplied with kit). The SaranWrap was immediately folded over the blot and the dioxitane spread evenly by gently wiping with a gloved hand. Finally the blot was wrapped in a fresh piece of SaranWrap and placed in an X-ray cassette with a piece of Cronex[®] Medical X-ray film (Dupont). Exposure time was dependent upon both target DNA and efficiency of probe labelling and varied from 10 min to 1 hour for low sensitivity to 4-16 hours for high sensitivity.

When blots were required for rehybridisation the probe was removed by boiling the blot in 0.1% (w/v) SDS for 10 min prior to prehybridisation.

2.11 Screening of a multicopy plasmid library of EPEC chromosomal DNA.

2.11.1 Colony blotting. This method is an adaptation of that used by Sambrook *et al.* (1989). DH5 α cells were electroporated with E2348-69/pBluescript *Bam*HI library DNA (Chapter 3, 3.2.7.2) and dilutions were spread onto ampicillin LUA plates to give approximately 500 colonies per plate. The plates were then incubated at 37°C until small colonies (1-2 mm diameter) were visible (typically 16 to 20 hours for DH5 α). 90 mm Hybond™-N+ (Amersham LIFE SCIENCES) membrane circles were laid gently onto the surface of the plates; both membranes and corresponding plates were numbered and marked (in pencil and indelible marker respectively) at 3 points on their circumference to allow subsequent orientation. After 1 min the membranes were lifted from the plates (which were reincubated to allow regrowth of the colonies) and placed colony side up onto a sheet of 3MM paper (Whatman) which had been soaked in denaturing solution. After 5 min the filters were transferred to a second sheet of 3MM paper, which had been soaked in neutralising solution. Finally, after a further 5 min, the filters were placed onto a dry sheet of 3MM paper and baked in an oven at 80°C for 1 hour to fix the DNA. After baking the filters were soaked in 5x SSC, 0.1% SDS for 10 min and the remaining colony debris was dislodged by gentle wiping with a clean paper towel prior to prehybridisation.

For secondary screening of potential positive clones individual colonies were patchplated directly onto 90 mm gridded Hybond™-N+ membranes which were laid on top of ampicillin LUA plates (grid side uppermost); the colony was also patchplated onto a corresponding master plate for reference. Plates were incubated overnight at 37°C prior to processing of the membranes as described above.

2.11.2 Screening colony blots using $\alpha^{32}\text{P}$ -dCTP labelled probes. For the detection of positives on colony blots $\alpha^{32}\text{P}$ -dCTP labelled probes were used in preference to the non-radioactive method of DNA detection due to the frequent tendency of the latter system to produce "spotting" which can appear as false positives.

2.11.2.1 Labelling of probes with $\alpha^{32}\text{P}$ -dCTP. The standard labelling reaction given here is designed for a single hybridisation; $\alpha^{32}\text{P}$ -dCTP labelled probe is not stable upon storage and should be used shortly after labelling. 50 ng of DNA were boiled for 3 min in a minimum volume of 10 μl and the labelling reaction set up as indicated below.

DNA	50 ng
OLB	5 μl
BSA	1 μl
$\alpha^{32}\text{P}$ dCTP	2.5 μl
Klenow	1 μl
H ₂ O	to 25 μl

The labelling reaction was incubated at 37°C for a minimum of 3 hours. Prior to addition to the hybridisation mixture the labelled probe was denatured by boiling for 5 min, diluted in 2 ml of hybridisation solution (see below) and passed through a 0.45 μm pore size filter.

Reaction components:

BSA (10 mg/ml, enzyme grade, New England Biolabs)

$\alpha^{32}\text{P}$ dCTP (10 $\mu\text{Ci}/\mu\text{l}$; Amersham LIFE SCIENCE)

Klenow (1 U/ μl ; GibcoBRL)

OLB is a mixture of A:B:C at 2:5:3, stored at -20°C.

Solution A.	625 μl	2M Tris HCl pH8.0
	25 μl	5M MgCl ₂
	350 μl	H ₂ O
	18 μl	β -mercaptoethanol
	5 μl	0.1 M dATP
	5 μl	0.1 M dTTP
	5 μl	0.1 M dGTP

Solution B. 2M HEPES, pH 6.6.

Solution C. Hexadeoxyribonucleotide primers (Pharmacia) resuspended in 3 mM Tris-HCl, 0.2 mM EDTA, pH 7.0, at 90 OD units/ml.

2.11.2.2 Hybridisation of $\alpha^{32}\text{P}$ -labelled probes to colony blots. All hybridisations were carried out in glass hybridisation tubes in a Hybaid mini hybridisation oven; when a number of blots were hybridised in the same bottle the filters were interleaved with nylon sheets to allow access of labelled probe to all surfaces. Blots were preincubated in hybridisation buffer (0.5 mM EDTA, 7% SDS, 0.5M NaHPO_4 , pH 7.2; 0.3 ml/cm² of blot) at 65°C for at least 1 hour prior to the addition of the denatured probe. Blots were hybridised for a minimum of 6 hours and typically overnight. The hybridisation buffer was decanted and the blots subjected to stringency washes at 65°C (approximately 2-5 ml/cm²); 2 x 15 min high stringency wash A (40 mM NaHPO_4 pH 7.2, 1% SDS) and 2 x 15 min high stringency wash B (0.1x SSC, 0.1% SDS). Finally the colony blots were wrapped in SaranWrap, and visualised by autoradiography by exposure to Cronex[®] Medical X-ray film (Dupont) at -80°C. Blots were exposed overnight (or longer for very weak signals), then developed and aligned with library master plates to identify positive clones.

2.12 Oligonucleotide primers.

All of the oligonucleotide primers used in this work were supplied by the Protein and Nucleic Acid Chemistry Laboratory (PNAAC), University of Leicester.

Primer	Purpose	Sequence
m13f	universal M13 forward sequencing primer	GTAAAACGACGGCCAGT
m13r	universal M13 reverse sequencing primer	AACAGCTATGACCATG
PA1	<i>TnphoA</i> sequencing 5' end; Figure 3.2	GTGAGCAGCCCGGTT
PB1	<i>TnphoA</i> sequencing 3' end; Figure 3.2	GTTAGGAGGTCACATG
kanF	PCR <i>Tn5</i> kanamycin gene; Figure 3.2	GGCACAACAGACAATCGG
kanR	PCR <i>Tn5</i> kanamycin gene; Figure 3.2	TGTCCTGATAGCGGTCCG
15b	pRDH8 sequencing; Figure 4.2	CAATCACTTTGCGCTCG
w9b	pRDH8 sequencing; Figure 4.2	TCAGGACTGGTCAGGCGC
w10b	pRDH8 sequencing; Figure 4.2	TGGGGATTTGAGGTCAGCG
w11a	pRDH8 sequencing; Figure 4.2	TACCTGCCCAAGAAATTGCG
w28a	pRDH8 sequencing; Figure 4.2	TCAACATTGGGGAATGGGG
w33b	pRDH8 sequencing; Figure 4.2	AGCGTGCAGTGGGGAGC
emtAF	PCR <i>emtA</i> 5' region; Figure 5.9	TTCCGGTCGACCGTGCGATGCAG
emtAR	PCR <i>emtA</i> 5' region; Figure 5.9	CCTTCGGATCCTCAATGCCTGC
orf1r	pRDH45 sequencing; Figure 4.18B	CCTGATTATTCAGTTTGC
rdh45a	pRDH45 sequencing; Figure 4.18B	AGGATTCTGTTGTTGTTACCG
ompxr	<i>ompx</i> mutant sequencing; section 5.2.1.2	GAATTAATCATCTTGCCG

2.13 Polymerase chain reaction (PCR).

This section describes a generic 10 μ l PCR reaction for the amplification of DNA fragments; all PCR reactions described in later chapters are based upon this with any alterations in reaction conditions (e.g. the $MgCl_2$ concentration of the PCR buffer) indicated in the text. A typical reaction consisted of the following; 1 μ l of DNA template (approximately 10 pg/ μ l), 1 μ l of 10x PCR buffer (500 mM KCl, 100 mM Tris-HCl pH 8.3, 1 mg/ml BSA (enzyme grade; New England Biolabs) and typically 100 mM $MgCl_2$), 1 μ l of 10 mM dNTPs (Ultra Pure; Pharmacia), 1 μ l each of Primer 1 and Primer 2 (1 pmole/ μ l), 5 μ l of distilled water, and 0.05 μ l of Taq DNA polymerase

(5 U/ μ l; Applied Biosystems). PCR was carried out using a Hybaid Omnigene thermal cycler; the cycling conditions for each primer pair are indicated separately in the results chapters.

For colony PCR the template DNA was prepared using the "boilate" method; single colonies were resuspended in 100 μ l of distilled water, heated at 100°C for 5 min and the cell debris then pelleted by centrifugation for 2 min. Typically 1 μ l of the boilate supernatant was sufficient as template in a PCR reaction as described above.

2.14 DNA sequencing.

All DNA sequencing in this work was performed using ABI PRISM™ technology (Perkin Elmer Applied Biosystems). Samples were analysed on an ABI PRISM™ 377 DNA sequencer by the Protein and Nucleic Acid Chemistry Laboratory (PNACL), Leicester University.

2.14.1 Double stranded DNA sequencing from plasmids. The initial DNA sequencing work in Chapter 3 was performed using the ABI PRISM™ dRhodamine Terminator Cycle Ready Reaction kit whilst the majority of the DNA sequencing of the pRDH7 and pRDH8 plasmids in Chapter 4 was performed using the improved (in terms of length of DNA sequence reads) ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit.

The ABI PRISM™ dRhodamine Terminator Cycle Ready Reaction kit was used according to the manufacturer's protocols. A typical reaction contained 500 ng of plasmid template (CsCl purified), 3.2 pmole of primer and 8 μ l of ABI PRISM™ dRhodamine ready reaction mix made up to 20 μ l with distilled water and then overlaid with mineral oil. The sequencing reaction was performed in a Hybaid Omnigene thermal cycler using a program of 25 cycles of 96°C for 30 sec, 50°C for 15 sec and 60°C for 4 min. After cycling the reaction mixture was transferred to a fresh

Eppendorf tube containing 2 μ l of 3 M sodium acetate (pH 4.6) and 50 μ l of ethanol, mixed by vortexing and incubated on ice for 10 min. The reaction products were recovered by centrifugation at 10,000 x g for 10 min, the pellet was washed with 500 μ l of 70% ethanol and then dried under vacuum. Samples were stored at -20°C prior to analysis.

The ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit was used in accordance with the protocols designed by the Advanced Centre for Genome Research (AGCT) at the University of Oklahoma (these can be accessed on their web page at www.genome.ou.edu/proto.html). A typical reaction consisted of 2-5 μ g of plasmid template (CsCl), 30 pmole of primer, and 8 μ l of ABI PRISM™ BIG DYES ready reaction mix made up to a final volume of 20 μ l with distilled water. The typical thermal cycling program consisted of one cycle of 95°C for 5 min, followed by 45 cycles of 95°C for 30 sec, 50°C for 30 sec and 60°C for 4 min. For sequencing reactions which gave poor DNA sequence it was possible to achieve improvements by increasing the number of cycles to 60, 80 or even 100. After the cycling reaction the products formed were recovered and purified as described above and stored at -20°C prior to analysis.

2.14.2 DNA sequencing from PCR products. PCR generated DNA fragments were purified using the QIAquick PCR purification kit (QIAGEN) to remove unincorporated primers and dNTP's. All PCR fragments were sequenced using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit; reaction mixtures and cycling conditions used were similar to those for double stranded plasmid DNA, except that only 200-500 ng of PCR fragment DNA were required per reaction.

2.15 Computer analysis of DNA and protein sequences.

DNA sequences were compiled and analysed using the SEQED, MAP, FRAMES, TRANSLATE, GAP and BESTFIT programs of the Wisconsin Molecular Biology

Package, Version 8.0 and Version 9.1-UNIX (Genetics Computer Group (GCG), Madison, Wisconsin).

Comparison of DNA or protein sequences with the GenBank databases were made using the programs BLAST (Altshul *et al.*, 1990) and Gapped-BLAST (Altshul *et al.*, 1997) which are available at the National Centre for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih). Variations of the BLAST programs allow the user to screen DNA sequences against a DNA database (BLASTN), protein sequences against a protein database (BLASTP), or protein sequences against a DNA database translated in all six frames to give protein sequence (TBLASTN). The quality of the match was estimated from the BLAST E value; E(xpect) value estimates the statistical significance of the match by specifying the number of matches within a given score that are expected in a search of a database of the given size.

Predictions of protein signal sequences cleavage sites and localisation of proteins in the cell were made using the SIGNALP and PSORT programs respectively (Nakai and Kanehisa, 1991) and identification of transmembrane regions was made using TMPRED; these programs are available on the ExpASY tools webpage (expasy.hcuge.ch/www/tools.htm).

The free energy of formation of stem-loops (ΔG) was calculated with the algorithm of Zucker and Stiegler (1981) using the online Vienna RNA Secondary Structure Prediction program available at www.tbi.univie.ac.at/cgi-bin/RNAfold.cgi (Hofacker *et al.*, 1994).

2.16 Analysis of proteins.

2.16.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Glycine-SDS-PAGE was performed by standard methods (Laemmli, 1970; Harlow and Lane, 1988). Tricine-SDS-PAGE was performed as described by Schägger and Von Jagow (1987).

Gels were cast and run using the Biorad Protean II and Protean II minigel systems. Separated proteins were visualised by Coomassie staining; the gel was incubated in Coomassie stain {10% (v/v) glacial acetic acid, 25% (v/v) methanol, 0.5 mg/ml Coomassie Brilliant Blue R-250 (Sigma)} for 2 hours with gentle agitation, then destained by repeated washes in 10% (v/v) glacial acetic acid, 25% (v/v) methanol.

2.16.2 Preparation of total protein extracts from *E. coli*. Bacterial cells were harvested by centrifugation at 10,000 x g for 10 min, and resuspended in 2x SDS-PAGE sample buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 2% SDS, 0.2% bromophenol blue, 20% glycerol) at a concentration of 0.1 OD units/10 µl (1 OD unit is equivalent to 1 ml of bacterial cells of OD₆₀₀ = 1.0). Samples were stored at -20°C prior to use, and denatured by boiling for 10 min before loading onto SDS-PAGE gels.

2.16.3 Preparation of outer membrane proteins from *E. coli*. Outer membrane proteins were prepared by differential solubility in Triton X-100 (Schnaitman, 1971; Wooldridge *et al.*, 1992). Samples were prepared from 1 ml of overnight bacterial culture or 10 ml of exponentially growing bacterial culture (A₆₀₀ of 0.6). Cells were harvested by centrifugation at 10,000 x g for 10 min, resuspended in 0.5 ml of envelope buffer (EB; 10 mM Tris-HCl, pH 7.5) and lysed by sonication. Unlysed cells were removed by centrifugation at 7,000 x g for 5 min, and crude total membranes were isolated by ultracentrifugation at 100,000 x g for 10 min in a Beckman TL-100 ultracentrifuge. Membrane pellets were resuspended in 1.0 ml of EB containing 2% (v/v) Triton X-100 (Sigma; EBT) and extracted at room temperature for 30 min. Triton insoluble pellets were collected by ultracentrifugation and re-extracted in 1.0 ml of EBT for a further 30 min at room temperature. Following ultracentrifugation outer membrane containing pellets were resuspended in 40 µl EBT and 40 µl 2x SDS-PAGE sample buffer (section 2.16.2). Samples were denatured by heating at 100°C for 5 min and 10 µl aliquots were analysed by SDS-PAGE and Coomassie staining (section 2.16.1). The major *E. coli* outer membrane proteins (OmpF/C and OmpA) were used

as internal standards to normalise the samples and a second SDS-PAGE gel was run with sample volumes corrected for protein concentration.

2.16.4 Transfer of proteins from SDS-PAGE gels to nitrocellulose membranes (Western blotting). Western blotting was performed essentially as described by Sambrook *et al.* (1989). Proteins were transferred from SDS-PAGE gels to nitrocellulose membranes (0.45 µm pore size; Schleicher and Schuell) using the Biorad Protean II Mini Blotter. Transfer buffer consisted of 25 mM Tris, 190 mM glycine, 10% (v/v) methanol. and transfer times were 90 min at a constant current of 250 mA. After transfer proteins were visualised and fixed to the nitrocellulose by staining with Ponceau S solution (0.5% (w/v) in 5% (w/v) TCA; BDH) for one minute, and then destained with PBS.

2.16.5 Antibody detection of bundle-forming pili proteins on Western blots. Protein blots were initially blocked overnight at 4°C in a solution of 5% BSA (Fraction V; Sigma), 0.5% Tween 20 in Tris-buffered saline (TBS; 135 mM NaCl, 2.7 mM KCl, 25 mM Tris-HCl pH 7.4). After a brief wash in TBS the blot was incubated with 1 ml of 1/1000 dilution of rabbit anti-BFP antibody (Giron *et al.*, 1993; section 2.5.3.2) in 0.5% BSA, 0.05% Tween 20 in TBS (TBS/BSAT) for 2 hours. Antibody incubations were performed using the "glass plate method" which allows the use of very small amounts of antibodies. A large glass plate, cleaned with methanol, was placed in a large sealable plastic box lined with dampened paper towels to prevent dehydration of the blot. The blot was placed sample side up on the glass plate; diluted antibody was pipetted directly onto its surface and spread evenly using a clean glass spreader. The blot was then inverted on the plate, the box sealed and the antibody left to incubate; improved results were obtained by respreading the antibody at intervals during the incubation. Blots were washed in TBS/BSAT (3 x 15 min washes) and then incubated in 1 ml of 1/5000 dilution of goat anti-rabbit IgG Peroxidase conjugate (Sigma Immuno Chemicals) diluted in TBS/BSAT for 1 hour. The blot was washed in TBS/BSAT (3 x 15 min) and then rinsed in TBS prior to development. Blots were

developed using the ECL system (Amersham LIFE SCIENCE) which utilises the oxidation of luminol by peroxidase to produce light which is then enhanced by second reagent. Equal volumes of Reagent 1 and Reagent 2 (supplied with the kit; 0.125 ml/cm² of blot) were mixed in a Sterilin tube and then poured onto the blot and incubated at room temperature for 5 min. The blot was drained, wrapped in SaranWrap and exposed to Cronex[®] Medical X-ray film (Dupont). Exposure times were approximately 10-20 min.

2.17 Alkaline phosphatase assay.

This is an adaptation of the method of Brickman and Beckwith (1975) and describes a standard assay for measuring alkaline phosphatase in bacteria growing shaking in broth. At regular intervals during growth the A₆₀₀ of the culture was measured and at each time point 2 x 1 ml samples were taken from the culture, transferred to Eppendorf tubes and the bacteria pelleted by centrifugation at 10,000 x g for 5 min. The supernatants were discarded and the two pellets resuspended together in 1 ml of 1 M Tris-HCl pH 8.0. The bacteria were preincubated in a static waterbath at 37°C for 5 min; 100 µl of *p*-nitrophenol phosphate (NPP; 4 mg/ml; Sigma) were added and the tube was re-incubated at 37°C until a yellow colour became visible. The reaction was stopped by addition of 100 µl of 1M K₂HPO₄ and the time since addition of NPP was noted. The cells were pelleted by centrifugation for 30 sec, the supernatants were transferred to plastic disposable cuvettes and A₄₂₀ and A₅₅₀ were measured (blank = 1 ml Tris-HCl pH 8.0 + 100µl NPP + 100 µl 1M K₂HPO₄). Alkaline phosphatase activity was calculated using the equation:

$$\text{Activity (Miller Units)} = 1000 \times \frac{A_{420} - (1.75 \times A_{550})}{\text{time (min)} \times \text{sample volume (ml)} \times A_{600}}$$

Chapter 3. Isolation and characterisation of EPEC mutants deficient in induction of protein phosphorylation in cultured human epithelial cells.

3.1 Introduction.

Despite the current extensive body of research detailing the effects of EPEC infection upon the host cell, the exact mechanism(s) by which EPEC-induced diarrhoea is caused is still unclear (Kaper, 1998). The rearrangement of cytoskeletal proteins in the formation of the attaching and effacing lesion and the tyrosine phosphorylation events observed do not provide complete explanations for the severe watery diarrhoea produced by EPEC (Sears and Kaper, 1996); however, recent investigations into EPEC induced protein kinase activity may offer explanations.

Baldwin *et al.* (1990) first demonstrated that EPEC infection of HEp-2 cell monolayers stimulated the serine-threonine phosphorylation of a number of host cell proteins, the most prominent amongst these being 21 kDa and 29 kDa (P29). A similar pattern of protein phosphorylation was also observed when either Caco-2 cells or human small intestinal mucosa were infected with EPEC (Manjarrez-Hernandez *et al.*, 1992) and induction of host cell protein phosphorylation was found to be a trait specific to those diarrhoeagenic *E. coli* strains which were associated with the formation of attaching and effacing lesions (Baldwin *et al.*, 1990). As a similar pattern of phosphorylated proteins could be induced by the treatment of HEp-2 cells with known protein kinase C (PKC) activators (such as 12-*O*-tetradodecanoylphorbol-13-acetate (TPA), phosphatidic acid or phospholipase-C) it was proposed that the EPEC-induced protein phosphorylation observed was the result of PKC activation (Baldwin *et al.*; 1990). It has recently been confirmed that infection of both HeLa and T84 cells with EPEC results in the elevation of membrane bound PKC activity and that this is dependent upon *eaeA*

gene function (Crane and Oh, 1997). Activation of PKC may play a direct role in the potentiation of diarrhoea; it has been shown that in intestinal cells elevation of PKC (usually by addition of TPA) can trigger secretion by the phosphorylation and activation of chloride channels (Fondacaro and Henderson, 1985; Picciotto *et al.*, 1992). PKC may also directly affect intestinal cell permeability as it has been reported to play a role in the mechanism by which the *zonula occludens* toxin of *Vibrio cholerae* causes alterations in tight junctions (Fasano *et al.*, 1995).

The prominent 21 kDa phosphoprotein observed in both EPEC-infected and TPA-induced HEp-2 cells was identified as myosin light chain (MLC) (Manjarrez-Hernandez *et al.*, 1991). MLC phosphorylation is known to be involved in the control of the actin-myosin interaction and is related to changes in actin organisation and cell shape in non-muscle cells (Bayley and Rees, 1986; Sellers, 1991). Experiments by Manjarrez-Hernandez *et al.* (1996) indicated that there were subtle differences between the phosphorylation of MLC seen in EPEC-infected HEp-2 cells and those induced with TPA; it was suggested that both TPA and EPEC were able to induce phosphorylation on MLC at a common site, which reduced its association with the cytoskeleton, but that only EPEC, via the Ca^{2+} /calmodulin-dependent MLC kinase (MLCK), was able to induce phosphorylation at a second site which was compatible with MLC-cytoskeletal association. Furthermore they proposed that these dual phosphorylation states of MLC might play a role in the polymerisation-depolymerisation of actin and therefore be involved in both the effacement of microvilli and subsequent actin accretion in EPEC lesion formation.

Adherence of EPEC bacteria to MDCK, Caco-2 or T84 cell monolayers has been reported to result in a decrease in the transepithelial electrical resistance (TEER) which provides an indicator of the barrier function of the monolayer (Canil *et al.*, 1993; Spitz *et al.*, 1995). Canil *et al.* (1993) proposed that the observed decrease in TEER in Caco-2 and MDCK cell monolayers was due to disruption of an intracellular pathway dependent upon AE lesion formation. More recently Spitz *et al.* (1995) have

demonstrated, using dual Na⁺ and mannitol flux measurements, that the decrease in TEER in T84 cells during EPEC infection is in fact due to a permeability defect in the tight junctions. Furthermore it was shown that this permeability defect was independent of the action of inhibitors of PKC or tyrosine kinase (Yuhan *et al.*, 1997), but that it could be prevented by sequestration of the intracellular calcium stores using dantrolene (Spitz *et al.*, 1995). Constitutive expression of a MLCK catalytic domain (tMK) in MDCK cells has been shown to result in enhanced MLC phosphorylation and to cause a disruption of the tight junction barrier (Hecht *et al.*, 1996). It was proposed that this enhanced phosphorylation of MLC could, by MLC interaction with the actinomyosin complex, induce the contraction of the peri-junctional ring of cytoskeletal proteins which underlie the intercellular tight junction thereby opening the junction and increasing cell permeability (Yuhan *et al.*, 1997). EPEC infection of MDCK cells expressing tMK (and therefore with enhanced MLC phosphorylation) failed to further increase paracellular permeability, suggesting that EPEC induced paracellular permeability also involves MLC phosphorylation; this hypothesis was confirmed by the complete inhibition of paracellular permeability in EPEC-infected cells by pretreatment with the MLCK-specific inhibitor ML-9 (Yuhan *et al.*, 1997). These data suggest that there is an alternative pathogenic mechanism induced in cell monolayers by EPEC infection which is independent of tyrosine phosphorylation, PKC activity, and attaching and effacing lesion formation, and which acts through MLCK to cause tight junction permeability; this mechanism may be responsible in part for the clinical symptoms of the disease.

The initial experiments reported in this thesis predated much of the recent work with MLCK and were based upon the observation that strain CVD206 (*eaeA*⁻), which was unable to induce the formation of attaching and effacing lesions, was still able to induce serine-threonine protein phosphorylation in HEp-2 or Caco-2 cell monolayers in a manner identical to wildtype EPEC (Haigh *et al.*, 1995; **Figure 3.1**). It was therefore proposed that screening for EPEC mutants which were inhibited in host cell protein phosphorylation would identify additional EPEC virulence determinants.

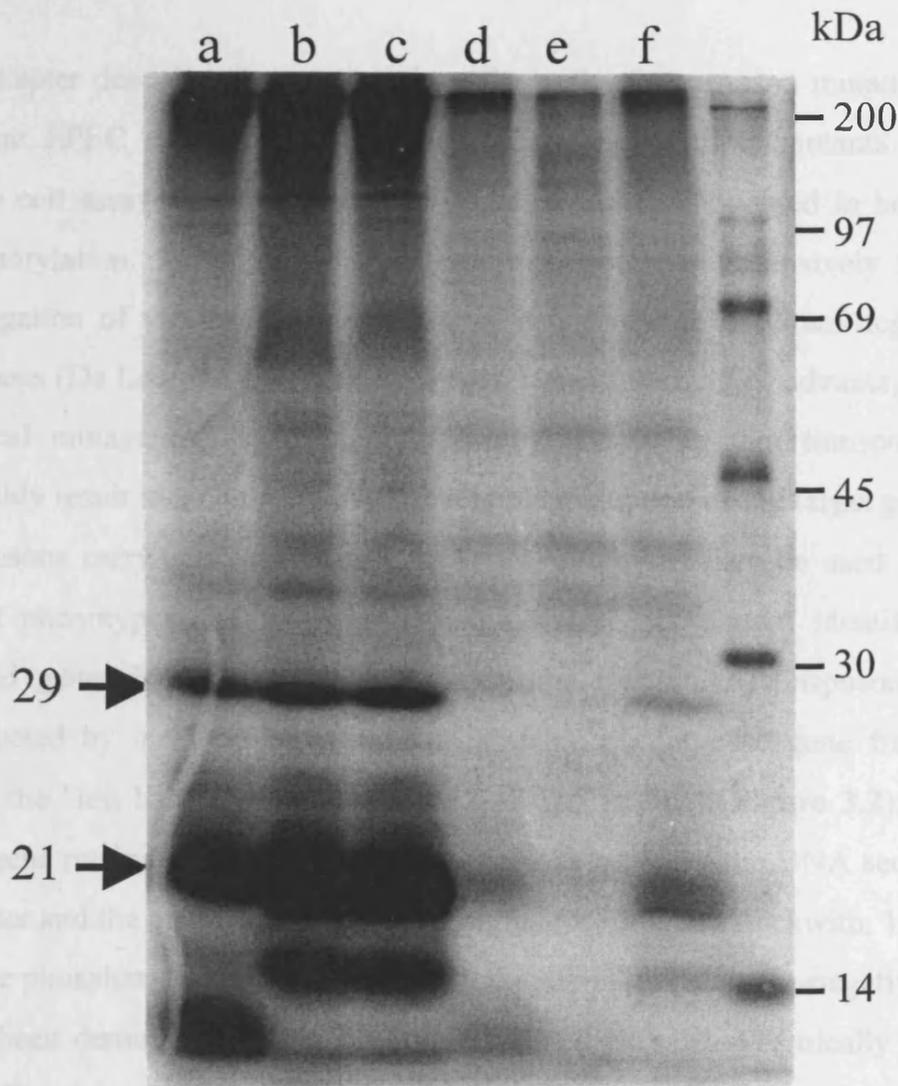
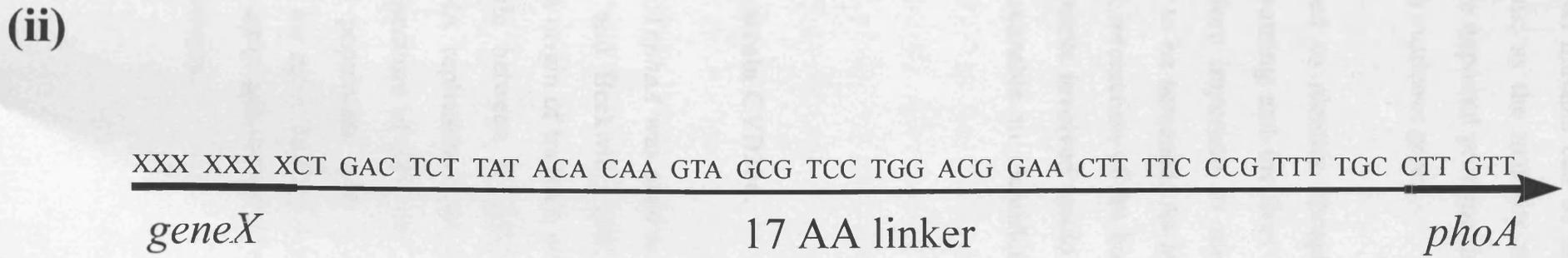
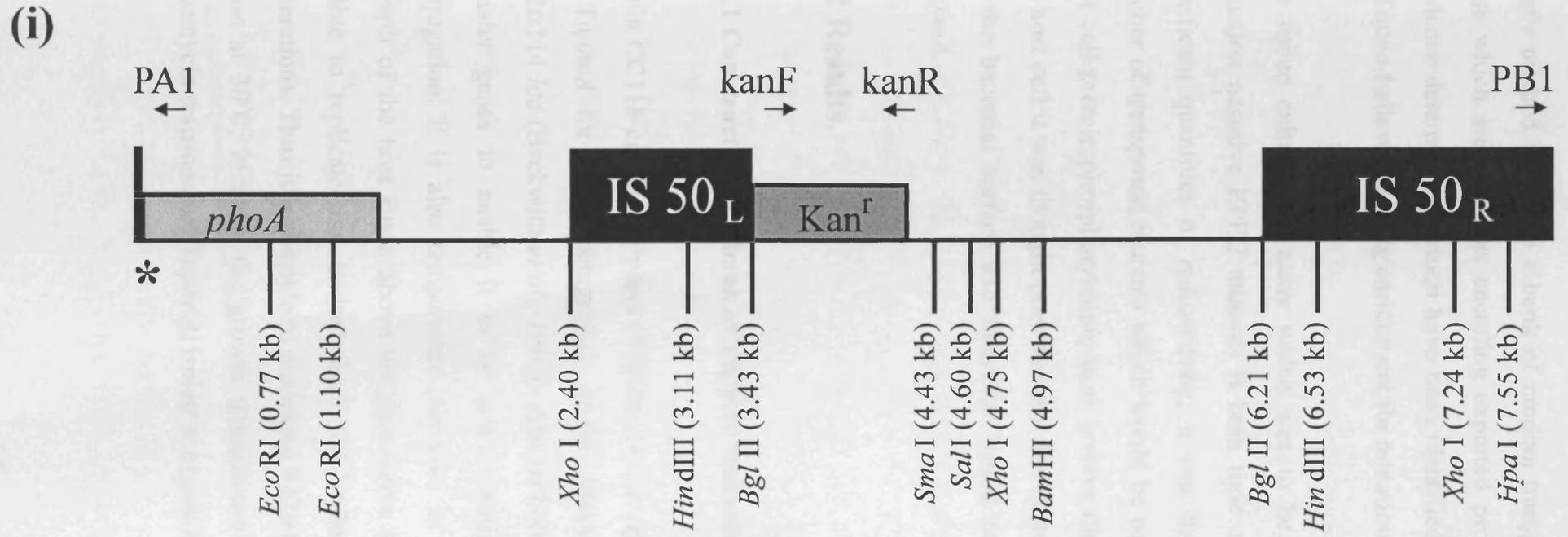


Figure 3.1 Protein phosphorylation profile of EPEC infected HEp-2 cells. Confluent HEp-2 cells grown in 24 well dishes were preincubated in MEM containing $^{32}\text{P}_i$ phosphate (0.1 mCi/ml) for 4 hours prior to infection with approximately 1×10^7 bacteria of strains: **a)** E2348-69, **b)** CVD206 (*eaeA*⁻), **c)** CVD206 pCVD438 (*eaeA*⁺), **d)** HB101 pJPN11, **e)** uninfected HEp-2 or **f)** MAR001(EAF⁻) and incubated for 3 hours at 37°C under 5% CO₂. Nonadherent bacteria were washed free and the HEp-2 cells lysed in 2x SDS-PAGE buffer. Samples were boiled for 10 min, the proteins separated by 15% SDS-PAGE and phosphorylated proteins determined by autoradiography (see also Materials and Methods, section 2.5.4); the major EPEC induced phosphorylated proteins, 21 kDa (myosin light chain) and 29 kDa are marked with arrows.

This chapter describes the construction of a bank of transposon mutants in the *eaeA*-deficient EPEC strain CVD206 and the screening of these mutants using a tissue culture cell assay to identify bacterial genes which are involved in host cell protein phosphorylation. Transposon mutagenesis has been used extensively for the genetic investigation of virulence determinants in a wide range of Gram-negative bacterial pathogens (De Lorenzo and Timmis, 1994). It has a number of advantages over UV or chemical mutagenesis systems, the primary one being that transposon insertions invariably result in a complete and irreversible disruption of the target gene. Also most transposons carry antibiotic resistance genes which can then be used as markers for mutant phenotypes and which can also assist in subsequent identification of the mutated genes by cloning or by chromosome mapping. Transposon *Tn_{phoA}* was constructed by insertion of a bacterial alkaline phosphatase gene fragment (*phoA*) within the "left hand" terminal sequence of Tn5 (IS50_L) (Figure 3.2); this modified *phoA* gene retains the enzyme's catalytic domain but lacks the DNA sequences for the promoter and the amino-terminal export signal (Manoil and Beckwith, 1985). Bacterial alkaline phosphatase is normally found in the periplasm of Gram-negative bacteria and it has been demonstrated that it must be located extracytoplasmically in order to be active (Brockman and Heppel, 1968). Random insertions of *Tn_{phoA}* into the bacterial chromosome will in some cases result in the creation of gene fusions with *phoA* (i.e. if transposon orientation is correct and the gene in-frame). If the gene concerned encodes a protein which is normally located in either the inner membrane, periplasm, or outer membrane then the fusion protein will also probably be exported and will therefore have alkaline phosphatase activity (Manoil and Beckwith, 1985). Such *phoA*-positive mutants are easily identified by production of a blue colouration when grown on LUA plates containing the alkaline phosphatase indicator substrate 5-chloro-4-bromo-3-indoyl phosphate (XP). *Tn_{phoA}* was initially designed as a probe for protein signal sequences in cloned or chromosomal genes, but it was rapidly realised that it could also be used as a method for the identification and characterisation of virulence determinants (Taylor *et al.*, 1989; Kaufman and Taylor, 1994). *Tn_{phoA}* provides a

Figure 3.2 Diagrammatic representation of *TnphoA*.

TnphoA (7.73 kbp) is based upon the transposon Tn5; it contains a promoterless truncated copy of the *phoA* gene, which has no signal sequence, inserted within the 5' region of IS50_L. A transposition event which results in insertion of the transposon in-frame into a gene will result in a PhoA fusion; if the gene encodes an exported product this will be detectable by increased alkaline phosphatase activity. (i) Restriction enzyme sites are indicated with the distance in kb from the 5' end of the transposon. The sequencing and PCR primers PA1, PB1, kanF and kanR are indicated as arrows. (ii) An enlargement of the 5' end of *TnphoA* (*) showing the 50 bp which encode the transposase recognition site from IS50_L and the 17 amino acid linker between an in-frame *TnphoA* insertion into *geneX* and the signal sequence-less *phoA* gene.



simple method by which a bank of random transposon mutants can be screened for those which are in genes encoding exported proteins; as the majority of bacterial virulence determinants which have been identified are exported proteins then the use of *TnphoA* allows a strong enrichment for mutations in virulence genes.

The tissue culture cell assay which was to be used to identify phosphorylation-induction negative EPEC mutants is both time consuming and involves the use of significant quantities of radioactivity; it was therefore important to minimise the number of transposon mutants which would be need to be screened. As induction of host cell protein phosphorylation must involve direct interaction of the bacteria with the host cell it was thought probable that the determinants involved would be located on the bacterial surface and therefore would be amenable to identification using *TnphoA*.

3.2 Results.

3.2.1 Construction of a bank of *TnphoA* mutants in strain CVD206.

Strain CC118 containing plasmid F42ts114 *lac zzf-1::TnphoA* was used as the source of *TnphoA* for these mutagenesis studies (Manoil and Beckwith, 1985). Plasmid F42ts114 *lac* (Beckwith *et al.*, 1966) contains both an origin of transfer, *oriT*, and the transfer genes to enable it to be self-transmissible between *E. coli* strains by conjugation. It is also temperature sensitive in DNA replication such that during growth of the host strain above the permissive temperature of 30°C the plasmid is unable to replicate and is lost from the bacterial population over a number of generations. Thus it is possible to conjugate F42ts114 *lac zzf-1::TnphoA* into an *E. coli* strain at 30°C, increase the growth temperature to 42°C and, by selection for the kanamycin resistance of *TnphoA*, isolate transposon mutants.

3.2.1.1 Introduction of F42ts114 *lac zzf-1::TnphoA* into CVD206. Conjugation between CC118 F42ts114 *lac zzf-1::TnphoA* and EPEC strain CVD206 was performed as described in Materials and Methods and CVD206 F42ts114 *lac zzf-1::TnphoA* transconjugants were selected by growth overnight at 30°C on LUA plates containing nalidixic acid, kanamycin and XP. A single *phoA*-negative transconjugant was selected for further analysis. An overnight LUB culture was serially diluted and aliquots containing approximately 10² to 10⁶ cfu were spread onto duplicate kanamycin, nalidixic acid, XP LUA plates; one series of plates was incubated at 30°C, the other at 42°C. It was apparent from the similar numbers of colonies on each pair of plates that, in contrast to its unstable phenotype in CC118, when transferred to CVD206 the F42ts114 *lac zzf-1::TnphoA* plasmid was no longer temperature sensitive for replication. The EAF plasmids of EPEC have been reported to possess IncFIIA-like replicons (Hales *et al.*, 1992). Therefore it is possible that the stabilising effect on plasmid replication observed in the CVD206 strain was due to the presence of a trans-acting factor encoded on the EAF plasmid (pMAR2) which was able to complement the temperature sensitive mutation. Despite the temperature resistant phenotype of the plasmid in CVD206, it was observed that strongly *phoA*-positive colonies were identifiable amongst the almost confluent growth on the lowest dilution plates, indicating that *TnphoA* transposition had occurred. Due to the conservative nature of the Tn5 transposition event when *TnphoA* transposes it destroys its host replicon. Thus the F42ts114 *lac zzf-1::TnphoA* plasmid would have been destroyed in the *TnphoA* transposition process and as it is a single copy plasmid would hopefully have been lost from the *phoA*-positive strains. It was therefore proposed that the *phoA*-positive colonies observed would be expected to contain single random transposon insertions, thus making them suitable for screening.

3.2.1.2 Isolation of *phoA*-positive *TnphoA* mutants from CVD206 F42ts114 *lac zzf-1::TnphoA*. Ten individual white colonies were selected from the original CVD206 F42ts114 *lac zzf-1::TnphoA* transconjugant plates to serve as stock for isolation of *phoA*-positive mutants. This multiplicity was designed to reduce the extent to which a

transposition event from plasmid to the chromosome which had already occurred in a stock strain might affect the randomness of the *TnphoA* bank. Overnight LUB cultures of each stock strain were diluted in LUB and 100 μ l aliquots of diluted culture, containing approximately 10^7 cfu, were spread onto multiple kanamycin, nalidixic acid, XP LUA plates, which were then incubated at 37°C until blue colonies were observed; typically the numbers of *phoA*-positive colonies were 10-20 per plate. In order to minimise the possibility of isolating sibling colonies a maximum of 4 blue colonies were selected from any single plate. Selected mutants were purified by streaking to single colonies on fresh kanamycin, nalidixic acid, XP LUA plates and then stored at -80°C.

In the original report on the use of *TnphoA* in EPEC (Jerse *et al.*, 1990) the authors used a bank of just 96 *phoA*-positive *TnphoA* mutants and were able to identify two separate transposon insertions in the *eaeA* gene. A rough calculation based on the probability of obtaining two independent insertions in the same gene from a bank of 96 mutants suggests that the number of EPEC genes encoding exported proteins which are both non-essential and which will form active *phoA* fusions could be as few as 50. However the total number of proteins present in the inner membrane, outer membrane and periplasm of *E. coli* B/r has been estimated to total in excess of 300 (Neidhardt *et al.*, 1990) and the number of genes which encode potential amino terminal signal sequences has been estimated at 575 (Blattner *et al.*, 1997); it must also be noted that this estimate was made in a laboratory *E. coli* strain and therefore does not include any exported virulence gene products. The possibility that 250 of the 300-plus proteins which are seen to be exported in *E. coli* strains could be essential for growth on a rich medium such as LUA, or that all of these proteins are unable to form stable *phoA* fusions, seemed most unlikely. It was concluded that the result obtained by Jerse *et al.* (1990) may just have been fortuitous and that any interpretations made from these data might therefore represent an underestimate of the true number of target genes. In light of these considerations a figure of 192 *phoA*-positive *TnphoA* mutants was chosen for

the initial study; this figure was double that used by Jerse *et al.* (1990) yet remained manageable to screen using the 96-well tissue culture cell phosphorylation assay.

3.2.2 Screening of the CVD206::*TnphoA* bank for mutants unable to induce protein phosphorylation in tissue culture cells.

The primary screen of the bank of 192 CVD206::*TnphoA* mutants was made using fully differentiated Caco-2 cells; this greatly reduced the background level of protein phosphorylation observed allowing negatives to more easily identified. The protein phosphorylation assays were performed essentially as described in Materials and Methods; Caco-2 cells were grown in 96-well tissue culture plates, preincubated with $^{32}\text{P}_i$, and then infected with the CVD206::*TnphoA* mutants for three hours. The Caco-2 cells were harvested, the proteins separated on 15% SDS-PAGE gels, and the phosphorylated proteins visualised by autoradiography. The initial screen identified eleven of the CVD206::*TnphoA* mutants which appeared to be unable to elicit a protein phosphorylation response in Caco-2 cells. These eleven potential protein phosphorylation deficient mutants were then rescreened in a scaled up phosphorylation assay using HEp-2 cells grown in 24-well tissue culture plates. Each mutant was assayed after infection for three hours and six hours and the samples analysed by SDS-PAGE and autoradiography (**Figure 3.3**). Five of the eleven CVD206::*TnphoA* mutants were confirmed to be completely unable to elicit host cell protein phosphorylation even after the extended infection time of six hours; these were CVD206::*TnphoA* #13, #55, #137, #138 and #181.

All of the protein phosphorylation deficient mutants had been observed to grow normally and reach similar cell densities in overnight culture in LUB prior to the phosphorylation assay (data not shown). However it was thought possible that the lack of host cell phosphorylation could be due to a simple metabolic effect which rendered the mutants unable to grow in the defined media (MEM) used in the assay. Growth curves were performed in MEM for the five mutants and a CVD206 control (**Figure 3.4**); these demonstrated that mutants #13, #55, #137 and #181 did indeed show a

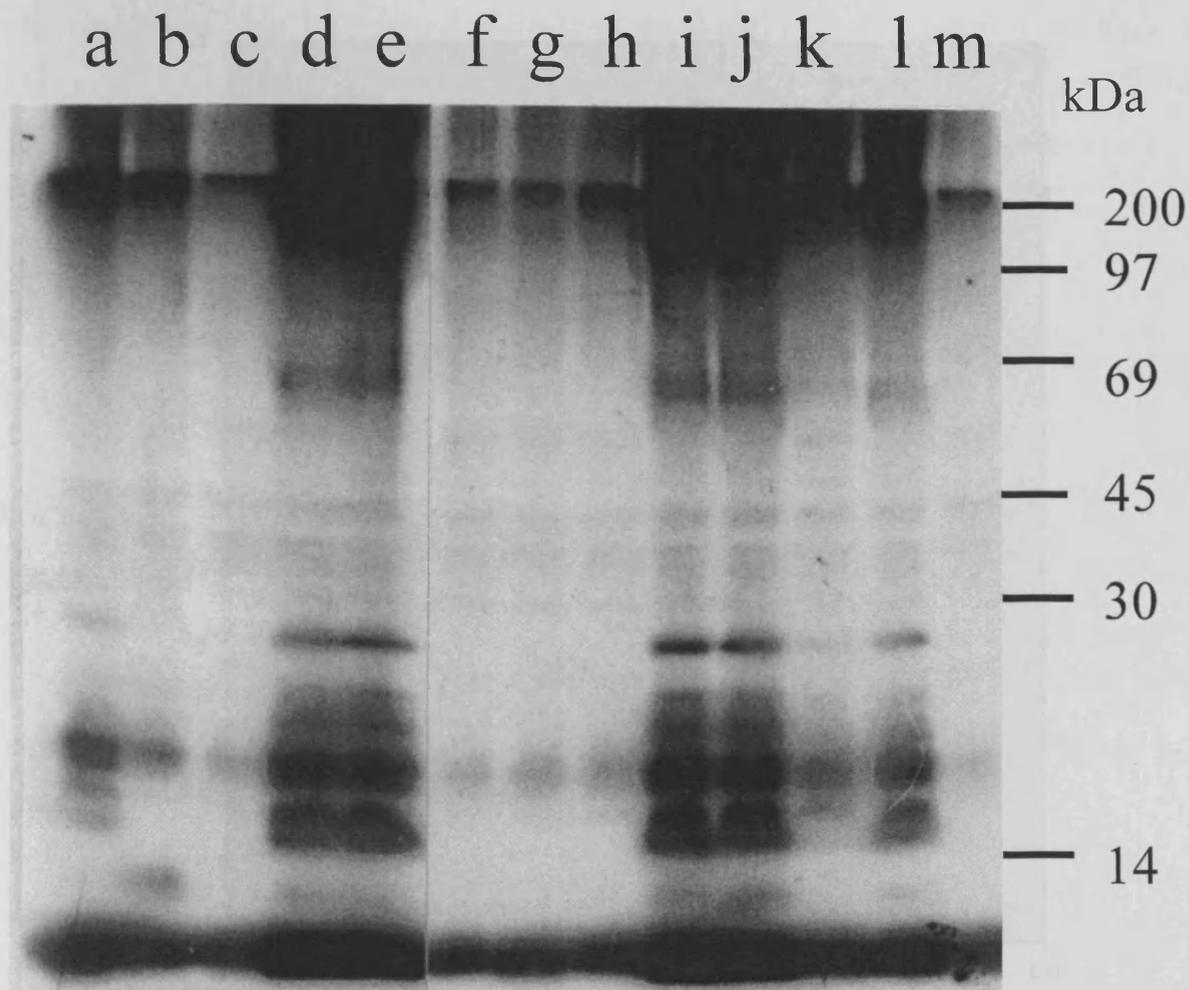


Figure 3.3 Secondary screening of CVD206::*TnphoA* mutants for induction of protein phosphorylation in human epithelial cultured cells. Confluent HEp-2 cells grown in 24 well dishes were preincubated in MEM containing $^{32}\text{P}_i$ phosphate (0.1 mCi/ml) for 4 hours prior to infection with approximately 1×10^7 bacteria and then incubated at 37°C in 5% CO_2 for 3 hours. Nonadherent bacteria were washed free and the HEp-2 cells lysed in 2x SDS-PAGE buffer. Samples were boiled for 10 min, the proteins separated by 15% SDS-PAGE and phosphorylated proteins determined by autoradiography (see also Materials and Methods, section 2.5.4). **a)** CVD206, **b)** HB101 pJPN11, **c)** #13 (**Negative**), **d)** #18, **e)** #19, **f)** #137 (**Negative**), **g)** #138 (**Negative**), **h)** #55 (**Negative**), **i)** #149, **j)** #125, **k)** #45, **l)** #162, **m)** #181 (**Negative**).

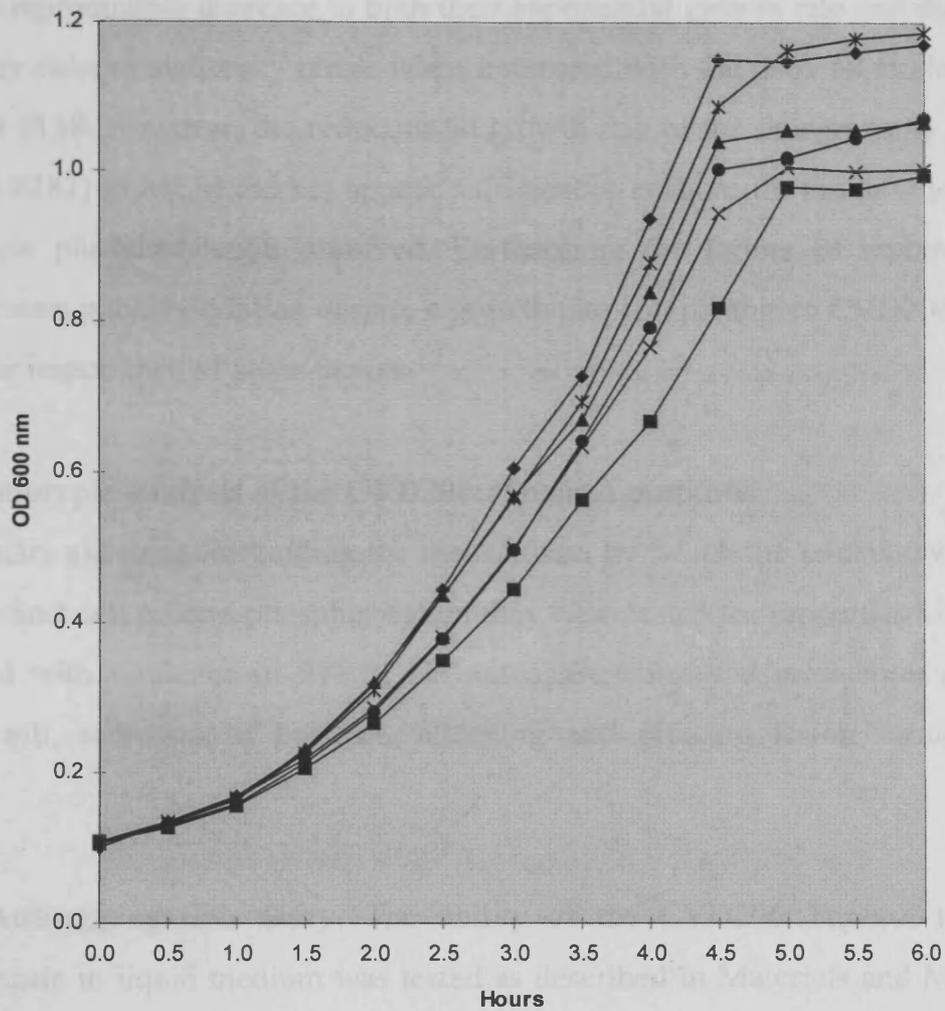


Figure 3.4 Growth curves of phosphorylation-deficient CVD206::TnphoA mutants. Overnight LUB cultures were diluted into 50 ml of MEM (ICNFlow Laboratories) to an OD₆₀₀ of 0.1 and grown shaking at 37°C; OD₆₀₀ measurements were made every 30 min until the cells entered stationary phase. ◆ CVD206, ■ CVD206::TnphoA #13, ▲ CVD206::TnphoA #55, X CVD206::TnphoA #137, * CVD206::TnphoA #138, ● CVD206::TnphoA #181. The above graph represents typical data from one of three separate repeats.

slight but reproducible decrease in both their exponential growth rate and the OD₆₀₀ at which they entered stationary phase when compared with the parental strain CVD206 or mutant #138. However, the reduction in growth rate of the four mutants (#13, #55, #137 and #181) in MEM did not appear sufficient to account for the total loss of host cell protein phosphorylation observed. Furthermore the failure of mutant #138 to induce protein phosphorylation despite a growth rate comparable to CVD206 indicated the greater importance of other factors.

3.2.3 Phenotypic analysis of the CVD206::*TnphoA* mutants.

As a primary aid to understanding the mechanisms by which the mutants were unable to induce host cell protein phosphorylation they were tested for properties known to be associated with virulence in EPEC, i.e. autoaggregation and production of bundle-forming pili, secretion of proteins, attaching and effacing lesion formation, and invasion.

3.2.3.1 Autoaggregation assay. The ability of the CVD206::*TnphoA* mutants to autoaggregate in liquid medium was tested as described in Materials and Methods. In CVD206 control cultures bacteria began to form small clumps within 30 min of inoculation into DMEM medium but remained dispersed even after 3 hours of incubation when grown in LUB (**Figure 3.5**). This result is consistent with those described previously for EPEC strain B171 (O111:NM) and CVD206 parental strain E2348-69 (O127:H6), and has been associated with the expression of the EPEC bundle-forming pilus (BFP; Vuopio-Varkila and Skoolnik, 1991). The five CVD206::*TnphoA* mutants were easily divided into one of two classes, each of which demonstrated phenotypes distinctly different from their parent. The first class, mutants #55 and #138, showed absolutely no aggregation in either DMEM or LUB, even after incubation for 6 hours (**Figure 3.5**). The second group, which included mutants #13, #137 and #181, formed bacterial clumps in DMEM in a manner apparently similar to CVD206, but they also formed aggregates when grown in LUB. The aggregates formed in LUB were visibly different from the compact spherical aggregates seen with DMEM cultures in

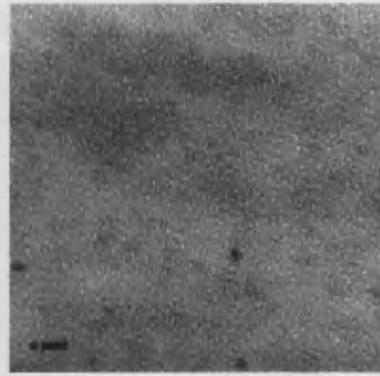
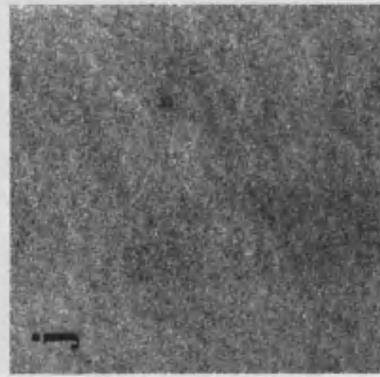
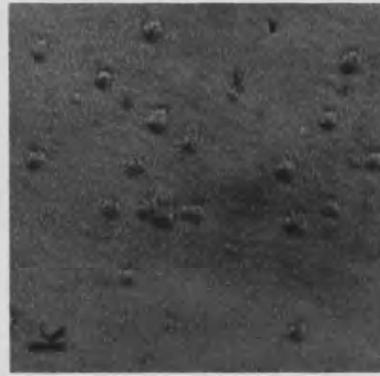
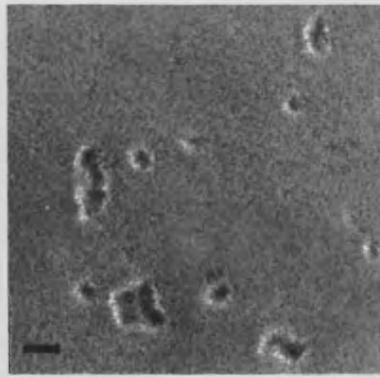
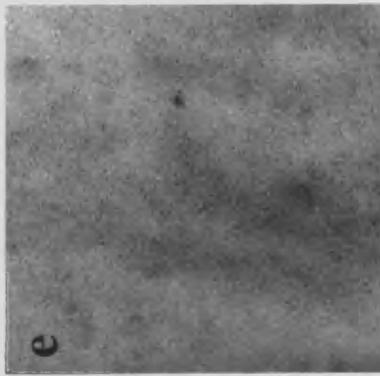
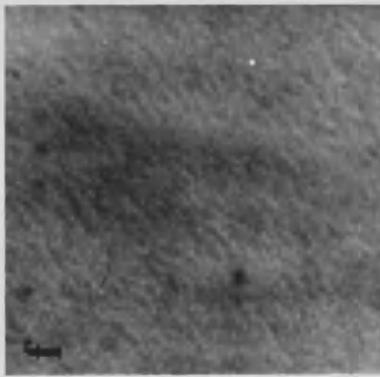
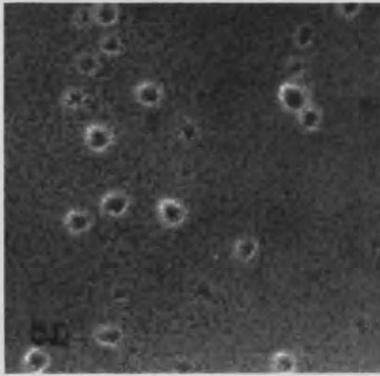
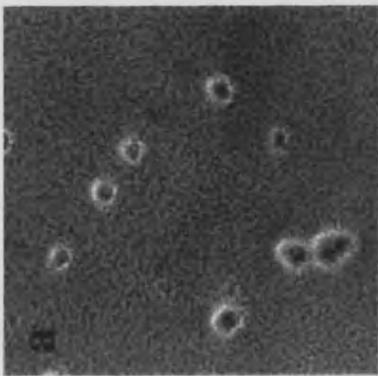
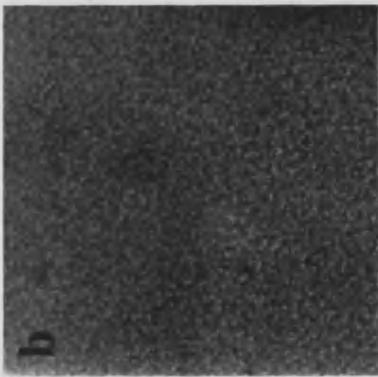
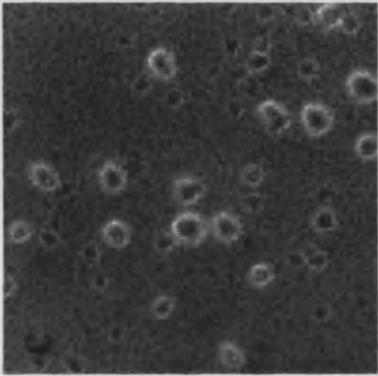
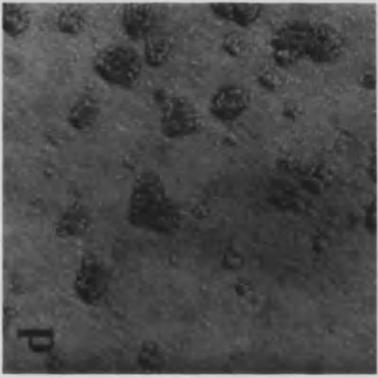
that they were typically much larger and more irregular in shape (**Figure 3.5**). Mutants #13, #137 and #181 also showed a secondary phenotype during the early stages of growth in DMEM (but not in LUB) which was not observed with CVD206 or with mutants #55 and #138. Both individual bacteria and bacterial aggregates were found to adhere to the base of the tissue culture well in a manner which was resistant to removal by either vigorous shaking or swirling of the plate; it was however possible to release bacteria by repeated pipetting with a 1 ml Gilson pipette. Free swimming bacteria were observed only 1-2 hours post-inoculation, at which point the entire base of the tissue culture well had been covered with a lawn of bacteria, therefore presumably preventing further adherence. Bacterial adherence to surfaces is usually associated with the expression of surface structures such as pili; CVD206 parental strain E2348-69 is reported to produce only two types of pili, BFP and the mannose-sensitive type 1 fimbriae (Elliot and Kaper, 1997). As similar levels of adherence were seen when the CVD206::*TnphoA* mutants #13, #137, and #181 were cultured in the presence and absence of 1% D-mannose it was concluded that the phenotype was not due to the altered expression of the type 1 fimbriae.

3.2.3.2 Secretion of EPEC virulence proteins. ³⁵S methionine-labelled secreted proteins were prepared from the supernatants of MEM cultures of CVD206 and the five phosphorylation deficient CVD206::*TnphoA* mutants using the protocol described in Materials and Methods. Analysis of the proteins by SDS-PAGE separation and autoradiography (**Figure 3.6**) demonstrated that each of the phosphorylation-deficient mutants were able to secrete EspA, EspB/D and EspC at levels comparable to those of CVD206.

3.2.3.3 Transformation of the CVD206::*TnphoA* mutants with pCVD438. Strain CVD206 lacks the *eaeA* gene. Therefore, in order to be able to test the effect of the *TnphoA* mutations upon attaching and effacing (AE) lesion formation or HEp-2 cell invasion it was first necessary to complement the *eaeA* mutation itself. Individual CVD206::*TnphoA* mutants were transformed by electroporation with the *eaeA*

Figure 3.5 Autoaggregation phenotypes of the CVD206::Tn*phoA* mutants. 20 μ l of overnight LUB bacterial cultures were inoculated into individual wells of a 24 well tissue culture plate which contained either 1 ml of LUB or 1 ml of DMEM containing 1% mannose (^w/_v) and which had been preincubated at 37°C under 5% CO₂ for 90 mins. Cultures were incubated for a further 3 hours, and then examined by phase contrast microscopy.

a) CVD206 DMEM, b) CVD206 LUB, c) #13 DMEM, d) #13 LUB, e) #55 DMEM, f) #55 LUB g), #137 DMEM, h) #137 LUB, i) #138 DMEM, j).#138 LUB, k) #181 DMEM, l) #181 LUB.



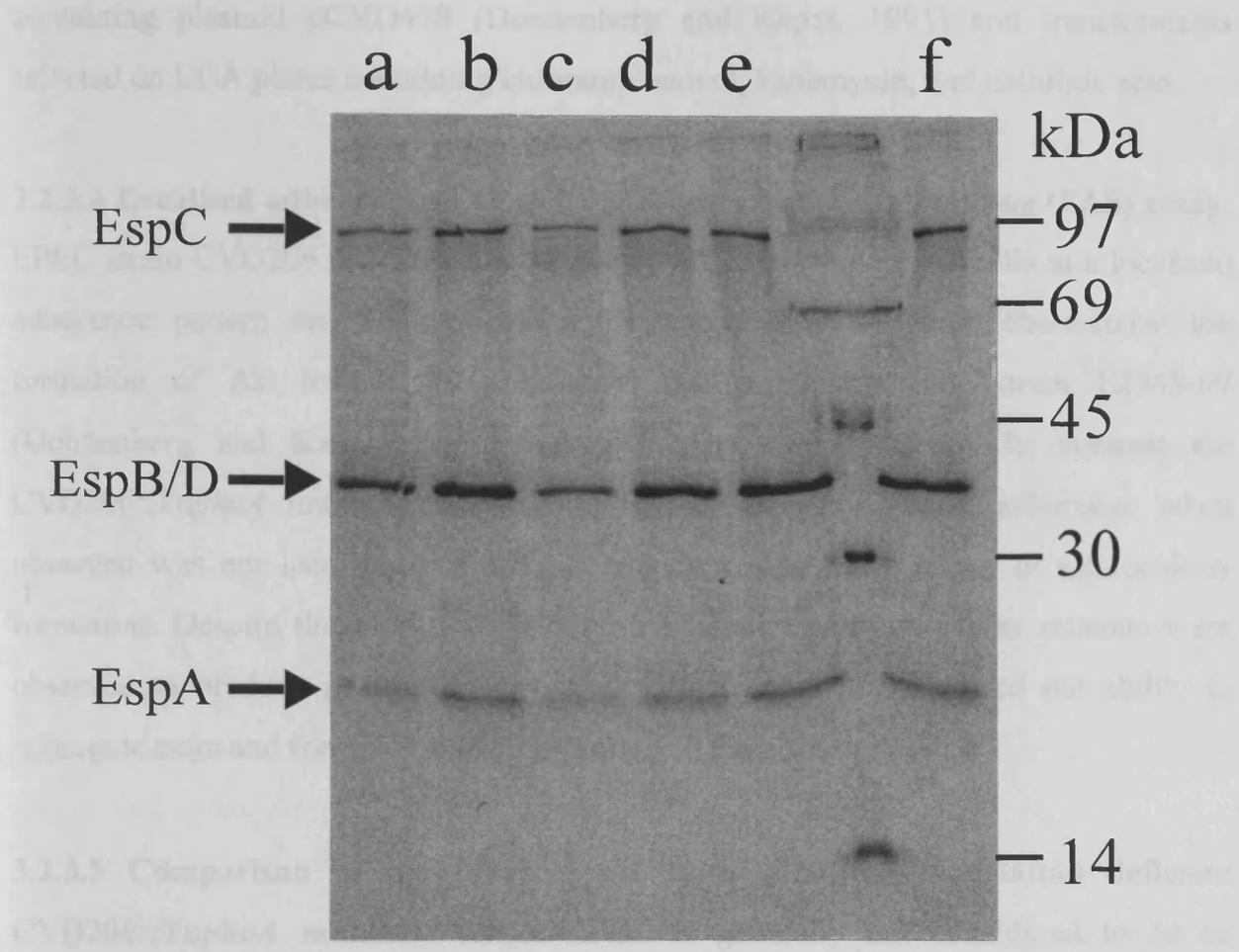


Figure 3.6 Secretion of EPEC virulence proteins. ³⁵S methionine labelled secreted proteins were prepared from EPEC strains grown in methionine-free MEM as described in Materials and Methods. Labelled proteins were separated on 13% SDS-PAGE and visualised by autoradiography.

The gel was loaded as follows: a) CVD206::*TnphoA* #13, b) #55, c) #137, d) #138, e) #181, f) CVD206.

containing plasmid pCVD438 (Donnenberg and Kaper, 1991) and transformants selected on LUA plates containing chloramphenicol, kanamycin, and nalidixic acid.

3.2.3.4 Localised adherence (LA) and the Fluorescent Actin Staining (FAS) assay.

EPEC strain CVD206 pCVD438 adheres very efficiently to HEp-2 cells in a localised adherence pattern and induces the actin rearrangements which characterise the formation of AE lesions in a manner identical to parental strain E2348-69 (Donnenberg and Kaper, 1991; **Figure 3.7**, panels a and b). In contrast the CVD206::*TnphoA* mutants adhered very poorly to HEp-2 cells; adherence when observed was not localised but diffuse, and there was no evidence of microcolony formation. Despite the very poor levels of adherence all five of the mutants were observed to produce positive FAS tests indicating that they retained the ability to aggregate actin and form AE lesions (**Figure 3.7**, panels c to l).

3.2.3.5 Comparison of the invasive ability of the phosphorylation deficient

CVD206::*TnphoA* mutants. Whilst EPEC is generally not considered to be an invasive organism, it has been demonstrated that significant numbers of viable bacteria can still be recovered after gentamicin treatment of EPEC infected HEp-2 cells (Donnenberg *et al.*, 1989) and that this gentamicin protection is dependent on the presence of the *eaeA* gene. Strains CVD206 (*eaeA*⁻), CVD206 pCVD438 (*eaeA*⁺) and the five CVD206::*TnphoA* phosphorylation mutants (carrying pCVD438) were assayed for their invasive ability in HEp-2 cells using the gentamicin protection assay as described in Materials and Methods (**Figure 3.8**). As previously reported (Donnenberg and Kaper, 1991) the strain CVD206 showed a 50-fold decrease in the number of gentamicin-protected bacteria recovered when compared to the *eaeA* positive strain CVD206 pCVD438. The effect of the *TnphoA* mutations upon the CVD206::*TnphoA* pCVD438 strains was to vastly reduce the number of gentamicin-protected bacteria recovered; though figures varied from assay to assay, typically the mutants were 100-1000 fold less invasive than the CVD206 pCVD438 control.

3.2.3.6 Expression of bundle-forming pili (BFP). The negative autoaggregation phenotype observed with CVD206::*TnphoA* mutants #55 and #138 strongly suggested that these mutants were unable to produce BFP. Western blot analysis using an anti-BFP antibody (Giron *et al.*, 1993) demonstrated that the mature BfpA protein could be identified in whole cell extracts from all five of the CVD206::*TnphoA* mutants (data not shown). Pili production was assayed on bacteria attached to HEp-2 cells using the rabbit anti-BFP antibody and a fluorescent donkey anti-rabbit FITC conjugate antibody as described in Materials and Methods section 2.5.3.2. With control strains CVD206 and CVD206 pCVD438 BFP were seen associated with both individual adherent bacteria and microcolonies; in contrast none of the adherent bacteria from the any of the five CVD206::*TnphoA* mutants appeared to produce pili (**Figure 3.9**). This lack of BFP production in the mutants accounted well for the poor adherence and invasion levels observed; however it was not clear what was responsible for the autoaggregation phenotype observed with mutants #13, #137 and #181.

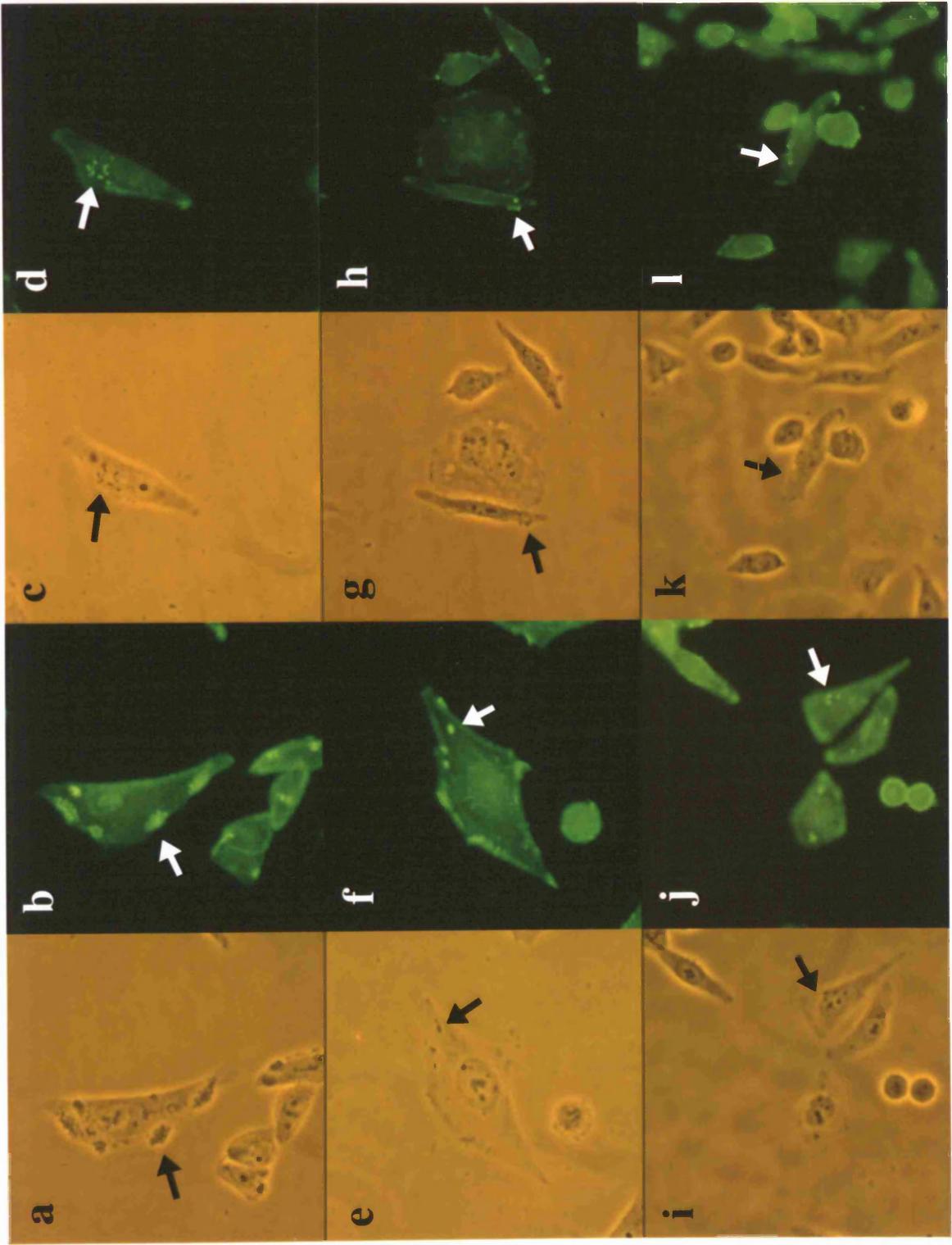
3.2.4 Identification of multiple *TnphoA* insertions in the phosphorylation-deficient CVD206::*TnphoA* mutants using Southern blot analysis.

It is possible after transposon mutagenesis for a mutant to carry more than one copy of the transposon usually due to secondary transposition events. This can present problems as it may be impossible in the case of two previously uncharacterised genes to identify immediately which is responsible for the phenotype observed. To confirm that the CVD206 mutants were the result of single *TnphoA* insertions each of the strains were screened by Southern hybridisation with a transposon specific DNA probe designed against the kanamycin resistance gene.

3.2.4.1 Production of a *TnphoA* specific DNA probe by PCR. PCR primers kanF and kanR (section 2.12 and **Figure 3.2**) were designed to amplify a 600 bp DNA fragment internal to the kanamycin resistance gene of *TnphoA*. Template DNA was prepared from a strain of CC118 containing plasmid F42*ts114 lac zzf-1::TnphoA* using the boilate method (section 2.13) and 10 µl of template were used in a total PCR

Figure 3.7 Actin rearrangement by the host cell phosphorylation deficient CVD206::*TnphoA* mutants. CVD206 pCVD438 (*eaeA*⁺) and the CVD206::*TnphoA* mutants each containing the plasmid pCVD438 were assayed for the formation of attaching and effacing lesions using the FAS test as described in Materials and Methods. Attached bacteria were identified using phase contrast microscopy (panels a, c, e, g, i, and k) and lesion formation was identified by bacterial-associated regions of bright fluorescence (aggregated actin stained with FITC-phalloidin) using incident light fluorescence microscopy (panels b, d, f, h, j, and l).

CVD206 pCVD438 (panels a and b), CVD206::*TnphoA* #13 pCVD438 (panels c and d), CVD206::*TnphoA* #55 pCVD438 (panels e and f), CVD206::*TnphoA* #137 pCVD438 (panels g and h), CVD206::*TnphoA* #138 pCVD438 (panels i and j), CVD206::*TnphoA* #181 pCVD438 (panels k and l). Adherent bacteria and lesions are indicated with arrows.



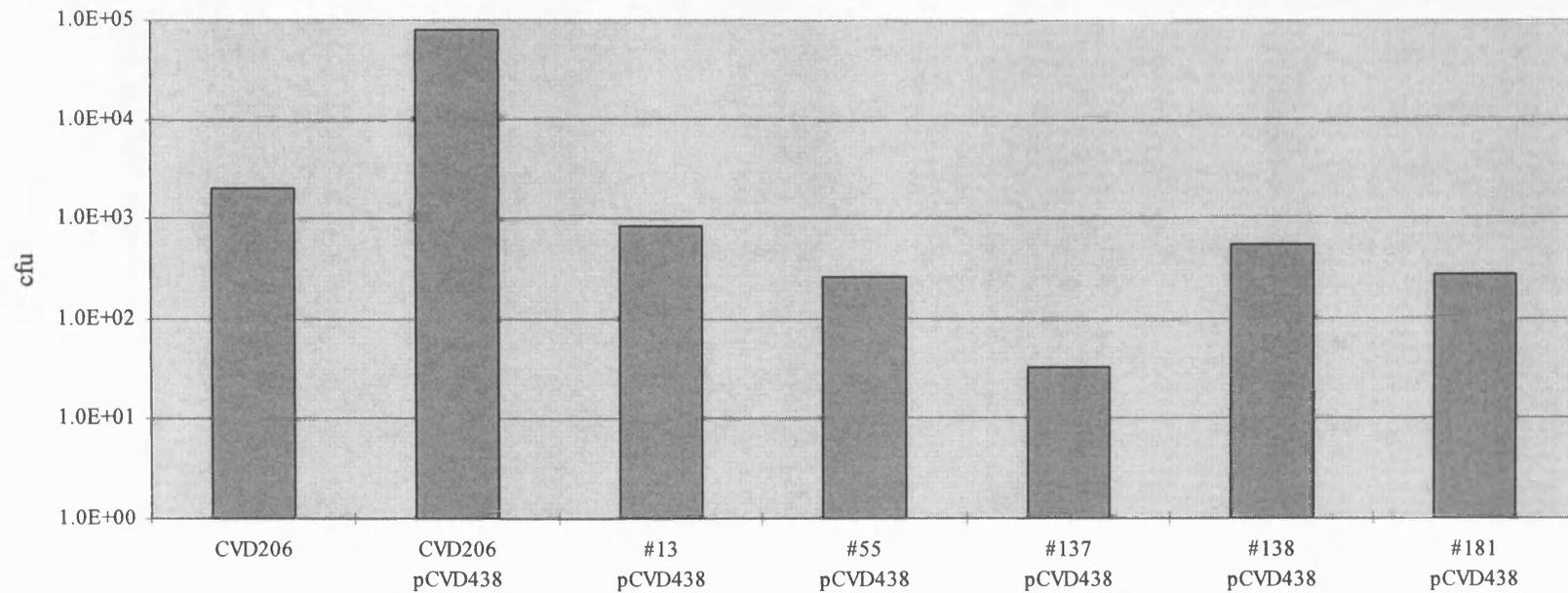
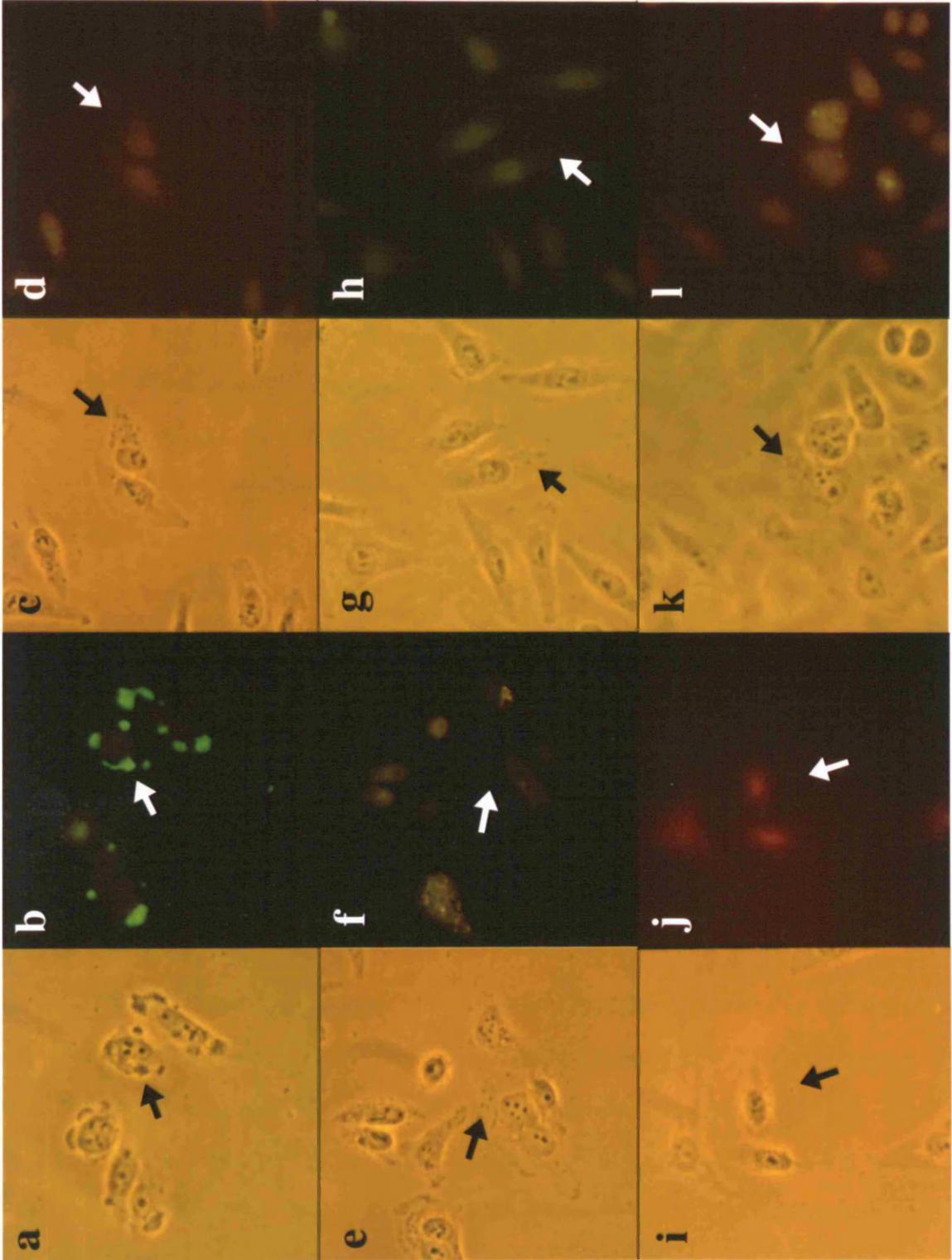


Figure 3.8 Invasive ability of phosphorylation-deficient CVD206::TnphoA mutants. Approximately 2×10^7 bacteria were added to each well of a 24 well tissue culture plate containing a confluent HEp-2 cell monolayer. Cells were incubated at 37°C under 5% CO₂ for 3 hours. Nonadherent bacteria were removed by washing in PBS and the monolayer overlaid with DMEM containing 100 µg/ml gentamicin and incubated for 2 hours. Gentamicin-protected bacteria were recovered from the monolayer by lysing the cells in 1% Triton X-100 and enumerated by serial dilution. Data presented are the average of three wells.

Figure 3.9 Analysis of bundle-forming pili (BFP) production in the phosphorylation-deficient CVD206::TnphoA mutants. CVD206 pCVD438 (*eaeA*⁺) and pCVD438 containing transformants of each of the CVD206::TnphoA mutants were allowed to attach to HEp-2 cell and then assayed for the production of BFP as described in Materials and Methods (section 2.5.3.2). Attached bacteria were identified using phase contrast microscopy (panels a, c, e, g, i, and k) and BFP were visualised as bacterial-associated fluorescent strands (detected with rabbit anti-BFP and donkey anti-rabbit FITC conjugate antibodies) using incident light fluorescence microscopy (panels b, d, f, h, j, and l).

CVD206 pCVD438 (panels a and b), CVD206::TnphoA #13 pCVD438 (panels c and d), CVD206::TnphoA #55 pCVD438 (panels e and f), CVD206::TnphoA #137 pCVD438 (panels g and h), CVD206::TnphoA #138 pCVD438 (panels i and j), CVD206::TnphoA #181 pCVD438 (panels k and l). Adherent bacteria and pili are indicated with arrows.



reaction of 100 μ l. PCR reaction conditions were optimised for the primer pair and found to require a final $MgCl_2$ concentration of 4 mM using a PCR program consisting of: 1 cycle of 95°C for 5 min; 35 cycles of 95°C for 30 sec, 50°C for 15 sec and 72°C for 90 sec; and 1 cycle of 95°C for 30 sec, 50°C for 15 sec and 72°C for 3 min. The resultant PCR product was analysed by 1.5% agarose gel electrophoresis and found to produce a single DNA band of approximately 600 bp. The band was excised from the gel and purified using Polymer wool, as described in Materials and Methods, prior to labelling using the Fluorescein Gene ImagesTM system (section 2.10.2.1).

3.2.4.2 Southern blot hybridisation analysis of the CVD206::Tn*phoA* mutants using a Tn*phoA* specific probe. Chromosomal DNA was prepared from each of the phosphorylation-deficient CVD206::Tn*phoA* mutants using the CsCl protocol described in Materials and Methods. The DNA was digested with *Bam*HI and the fragments separated by 1% agarose gel electrophoresis prior to transfer to Hybond-NTM membrane by Southern blotting. The resultant Southern blot was then probed at high stringency with the Tn*phoA* specific Kan probe described above. It had been expected that each mutant strain would carry a single copy of the Tn*phoA* transposon, but the Southern blot analysis clearly demonstrated that they all carried multiple insertions; three of the five mutants (#55, #137, and #181) had two copies of Tn*phoA* whilst the other two mutants (#13 and #138) contained three copies (**Figure 3.10**). It was immediately clear by the pattern of the Tn*phoA* specific bands that each of the mutants were independently derived, but they each appeared to have a large band of similar size. The probability of three independent transposon insertions having occurred was extremely low and therefore a more feasible explanation was sought for the results observed. It had been assumed that the single copy F42*ts114 lac zzz-1::TnphoA* plasmid would be lost due to the conservative nature of the Tn*phoA* transposition event. However it was considered possible that transposition had occurred during replication when two copies of the plasmid existed and that the Tn*phoA* mutant might still retain the plasmid. Chromosomal DNA was prepared from strain CC118 containing F42*ts114 lac zzz-1::TnphoA* and the size of the Tn*phoA*-containing *Bam*HI

fragment was determined by Southern hybridisation as above (**Figure 3.10, lane a**). The F42ts114 *lac zzf-1::TnphoA* derived fragment was found to be of a similar size to the large common band observed in each of the mutants. It was therefore proposed that the plasmid encoded copy of *TnphoA* was responsible for one of the *TnphoA* specific bands observed in each mutant, and that only mutants #13 and #138 were really the result of double insertions into the chromosome. The *TnphoA* insertion in F42ts114 *lac zzf-1::TnphoA* is *phoA*-negative and would therefore present no problem during the cloning and characterisation of the flanking DNA regions in mutants #55, #137 and #181. The secondary insertions in #13 and #138 however were problematic as it was now impossible to determine if the *phoA*-positive insertion (assuming that both insertions were not *phoA*-positive) was responsible for the phenotype observed.

3.2.5 Cloning and analysis of the DNA sequences flanking the CVD206::*TnphoA* *phoA*-positive insertions.

The kanamycin resistance gene in *TnphoA* can be used as a direct selection for cloning DNA fragments which contain both transposon sequence and the chromosomal sequences flanking the transposon insertion site. A number of common restriction enzymes which are present in the multiple cloning sites of the plasmids pUC18/19 or pBluescript do not cut within the *TnphoA* sequence (e.g. *KpnI*, *SstI*, *EcoRV*) and can therefore be used to clone the whole transposon. Cloning of large DNA fragments (*TnphoA* excluding the flanking regions is 8 kb) into multicopy plasmids can result in plasmid stability problems and the occurrence of deletions. Therefore it was proposed that the initial attempts at cloning the *TnphoA* flanking regions would involve use of the internal *BamHI* site (**Figure 3.2**) to clone a fragment which contained the "left hand" end of *TnphoA* (including the kanamycin resistance gene) and the DNA upstream.

3.2.5.1 Cloning of the regions upstream of the CVD206::*TnphoA* insertions. 5 µg of CsCl purified chromosomal DNA were digested to completion with *BamHI* and ligated overnight at 16°C with 0.5 µg of *BamHI* digested pUC18 DNA. The ligations

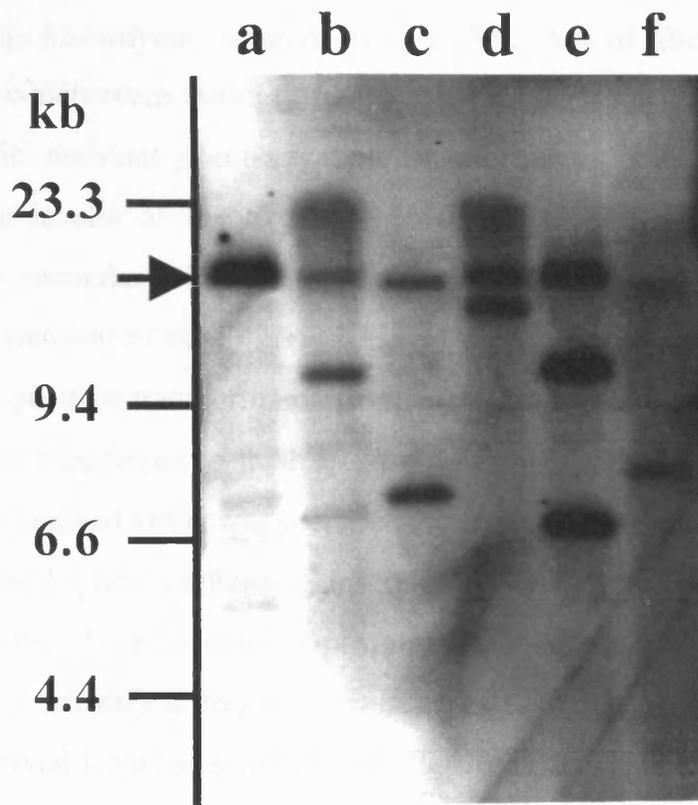


Figure 3.10 Southern blot analysis of CVD206::TnphoA mutants using a TnphoA specific probe. CsCl prepared chromosomal DNA was digested with *Bam*HI, separated by 1% agarose gel electrophoresis and transferred to nylon membrane. The blot was then hybridised with the 600 bp TnphoA specific Kan probe (the PCR product of kanF and kanR; section 3.2.4.1). The gel was loaded as follows: **a)** CC118 F42ts114 *lacI3 zzf-2::TnphoA*, **b)** CVD206::TnphoA #13, **c)** CVD206::TnphoA #55, **d)** CVD206::TnphoA #137, **e)** CVD206::TnphoA #138 and **f)** CVD206::TnphoA #181. The arrow indicates a common band in all five of the mutants which co-migrates with the TnphoA derived *Bam*HI band from F42ts114 *lacI3 zzf-2::TnphoA*.

were transformed into DH5 α by electroporation and transformants selected on LUA containing kanamycin, ampicillin and XP. All of the transformations produced colonies which were strongly *phoA*-positive, but each of the mutants also produced kanamycin resistant *phoA*-negative transformants. This result was not unexpected given the results of the Southern blot analysis performed in section 3.2.4 which indicated secondary Tn*phoA* insertions in mutants #13 and #138 as well as the apparent survival of the F42*ts114 lac zzzf-1::TnphoA* plasmid in all of the strains. Only the *phoA*-positive transformants from each plate were selected for further analysis and these were transferred to fresh LUA plates containing kanamycin, ampicillin and XP. Miniprep plasmid DNA was prepared from each of the clones, digested with *Bam*HI to release the Tn*phoA* containing inserts and the digests analysed on 1% agarose gels. Many of the clones contained plasmids with multiple *Bam*HI fragments, but it was possible to identify a fragment present in each plasmid which was common to all others derived from the same mutant. Representative clones, which contained plasmids with a single *Bam*HI insert, were chosen for each mutant; these were p13a, p55a, p137c, p138b and p181c.

3.2.5.2 Restriction enzyme digestion analysis of Tn*phoA* clones. Plasmid DNA was prepared from each of the Tn*phoA* containing clones using the CsCl method and the sizes of the *Bam*HI inserts were determined; p13a (6.5 kb), p55a (7.4 kb), p137c (15 kb), p138b (6.7 kb) and p181c (7.9 kb) (Figure 3.11A). The plasmids were also mapped with the restriction enzymes *Hind*III and *Sma*I, these enzymes each cut once in the partial Tn*phoA* fragment and once in the pUC18 MCS, thus allowing the orientation of the *Bam*HI insert within the vector (with respect to Tn*phoA*) to be determined (Figure 3.12); it was found that the plasmids p13a, p55a, p137c and p181c all contained an additional *Hind*III site in the DNA upstream of *phoA*.

3.2.6 Analysis of DNA sequences flanking the cloned Tn*phoA* insertions.

The DNA sequences directly adjacent to the Tn*phoA* insertion sites and directly adjacent to the pUC18 multiple cloning site were determined for each of the cloned

transposon insertions using CsCl prepared plasmid DNA with the ABI PRISM™ dRhodamine Terminator Cycle Sequencing system as described in Materials and Methods. The primers used were PA1 (section 2.12), which was designed to anneal to a sequence approximately 60 bp within the 5' end of *TnphoA*, thus allowing determination of the *phoA* fusion site and the DNA sequence directly upstream (Figure 3.2), and the universal primers m13f and m13r, which anneal to sequences in the *lacZ* gene either side of the multiple cloning sites of both pUC18/19 and pBluescript thus allowing determination of the sequences of DNA fragments cloned into these vectors (section 2.12). Each sequencing reaction was performed in duplicate; comparison of the pairs of sequences using the program BESTFIT (section 2.15) indicated that they each gave approximately 300 bp of reliable DNA sequence.

3.2.6.1 Analysis of DNA sequences from plasmids p13a, p137c and p181c. The DNA sequence determined from plasmids p13a, p137c and p181c clearly indicated that the *TnphoA* insertions in their respective parent strains were clustered at a single locus. Sequencing of the plasmids p13a and p137c using primer PA1 demonstrated that the sites of the *phoA*-positive *TnphoA* insertions in mutants #13 and #137 were identical to the base pair; it also identified the exact position of the *Hind*III site previously mapped in these clones (Figure 3.13A). The DNA sequences obtained from plasmids p181c and p13a using the m13r primer were also identical indicating that the *TnphoA* insertions in mutants #13 and #181 were both downstream of the same *Bam*HI site and therefore at the same locus. However the sequence from p181c determined using PA1 was different from that of both p13a and p137c indicating that the actual site of the *TnphoA* insertion in mutant #181 was different (Figure 3.12A). Using m13r to sequence the region upstream of *TnphoA* in plasmid p137c gave the unexpected result that this region was different from that in p13a and p181c. This lack of homology, combined with the unexpected difference in the size of the p13a and p137c cloned *Bam*HI fragments (Figure 3.9A), could only be explained if it was assumed that the *Bam*HI site upstream of the *TnphoA* insertions in mutants #13 and #181 was absent in

mutant #137 (**Figure 3.12A**). All of the DNA sequences obtained were compared against the GenBank sequence database using the BLASTN program (Altschul *et al.*, 1990; Altschul *et al.*, 1997; section 2.15), but they failed to detect any significant homology.

3.2.6.2 Identification of open reading frames adjacent to the site of *TnphoA* insertion in mutants #13/#137 and #181. The DNA sequences obtained from plasmids p13a, p137c and p181c using primer PA1 were analysed using the MAP program of the Wisconsin GCG Molecular Biology software package to identify potential open reading frames. Sequence p13a/p137c PA1 was found to contain an open reading frame which was in-frame with the *phoA* gene and which extended for the complete length of the determined sequence (**Figure 3.13A**). Sequence p181a PA1, however, was found to encode a stop codon approximately 200 bp upstream of, and in frame with, the *phoA* fusion site indicating that the amino terminus of the fusion protein must be encoded within that region. There were three potential ATG start codons in-frame with the *phoA* of *TnphoA* within the 200 bp of sequence, but only one had a good consensus ribosome binding site associated with it (GAGGG; Miller, 1992); this start site gave a partial polypeptide of 52 aa (**Figure 3.13B**). Analysis of the predicted PhoA fusion protein using the signal sequence determination program PSORT (section 2.15) predicted an integral inner membrane localisation which was in accordance with the *phoA*-positive nature of the p181c plasmid. The predicted polypeptides identified from both the p13a/p137c PA1 and p181c PA1 DNA sequences were used to screen the Genbank database with the BLASTP program (Altschul *et al.*, 1990; Altschul *et al.*, 1997; section 2.15), but they failed to identify any significant matches.

3.2.6.3 Analysis of DNA sequence from plasmids p55a and p138b. The DNA sequences obtained from plasmids p55a and p138b indicated that the *TnphoA* insertions cloned from mutants #55 and #138 were at the same locus. Sequences determined for 55a and 138b using primer PA1 showed no homology indicating that

the sites of *TnphoA* insertion were different, but the sequences from the M13 primers (m13f for p138b, m13r for p55a) were identical demonstrating that the *Bam*HI site and regions upstream of the transposon insertions were common to both. The sequence of p55a PA1 also identified the exact location of the *Hind*III site mapped in section 3.2.5.1 (Figure 3.14A). The DNA sequences obtained with both the PA1 and the m13 primers were used to screen the Genbank database with BLASTN, but they failed to identify any significant matches.

3.2.6.4 Identification of open reading frames adjacent to the *TnphoA* insertions in mutants #138 and #55. The sequences obtained from plasmids 55a and 138b using primer PA1 were analysed using the MAP program of the Wisconsin GCG Molecular Biology software package to identify potential open reading frames. Both sequences encoded potential polypeptides which were predicted to form in-frame fusions to the *phoA* gene of *TnphoA* which ran for the entire length of the sequences determined (Figure 3.14). The predicted polypeptides identified from both the 55a and 138b PA1 sequences were used to screen the Genbank database with the BLASTP program, but they failed to identify any significant matches.

3.2.7 Cloning the wildtype loci identified by the CVD206::*TnphoA* mutants from EPEC strain E2348-69.

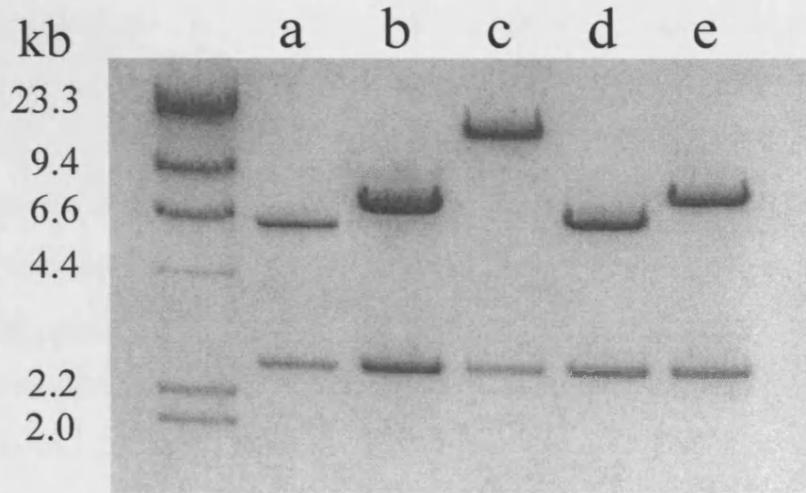
It was initially proposed to clone the equivalent wildtype regions of the two loci identified by the CVD206::*TnphoA* mutants #13/#137/#181 and #55/#138 as fragments from a partially digested *Sau*3A E2348-69 chromosomal library. However preliminary Southern blot analysis using the proposed DNA probes (see below section 3.2.7.2; Figure 3.15) indicated that both of the loci were encoded on *Bam*HI fragments of approximately 5-6 kb, and it was therefore determined to clone these fragments instead. There were two advantages in cloning the loci as *Bam*HI fragments: (i) construction and screening of a library from a complete restriction enzyme digest is much simpler than using a *Sau*3A partial digest library, and (ii) the upstream *Bam*HI sites had already been mapped in relation to the *TnphoA* insertions and it was known

Figure 3.11 Restriction enzyme digestion analysis of cloned *TnphoA* insertions.

(A) The 5' end of *TnphoA* and upstream DNA sequences from each of the host cell phosphorylation-deficient CVD206::*TnphoA* mutants were cloned into pUC18 as *Bam*HI fragments using the *TnphoA* kanamycin resistance gene for selection. A *phoA*-positive representative clone was chosen for each mutant and CsCl plasmid DNA prepared. *Bam*HI digestions of each clone were separated by 1% agarose gel electrophoresis; the gel loading was as follows (estimated *TnphoA* insert size): a) p13a (6.5 kb), b) p55a (7.4 kb), c) p137c (15 kb), d) p138b 6.7 kb) and e) p181c (7.9 kb).

(B) Clones 55a and 181c were used to prepare DNA probes specific to the loci into which the *TnphoA* transposons in mutants #55/#138 and #13/#137/#181 were inserted. *Bam*HI and *Hind*III digestions of CsCl prepared plasmid DNA were separated by 1% agarose gel electrophoresis and DNA fragments isolated using polymer wool as described in Materials and Methods. Gel loading: a) p55a *Bam*HI b) p55a *Bam*HI *Hind*III, c) p181c *Bam*HI, d) p181c *Bam*HI *Hind*III. Arrows indicate the *Bam*HI-*Hind*III fragments used as probes (see also Figure 3.12).

A



B

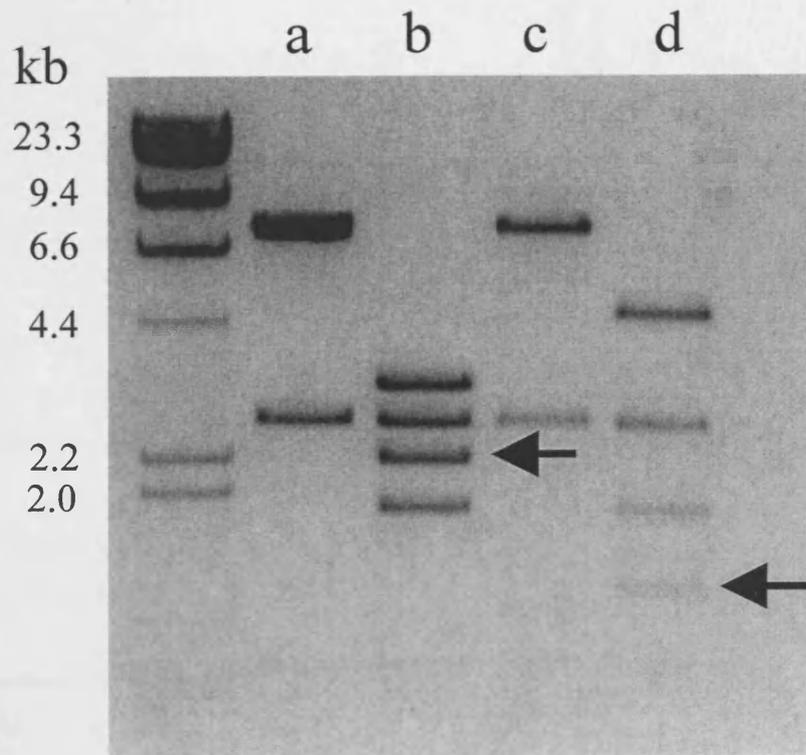
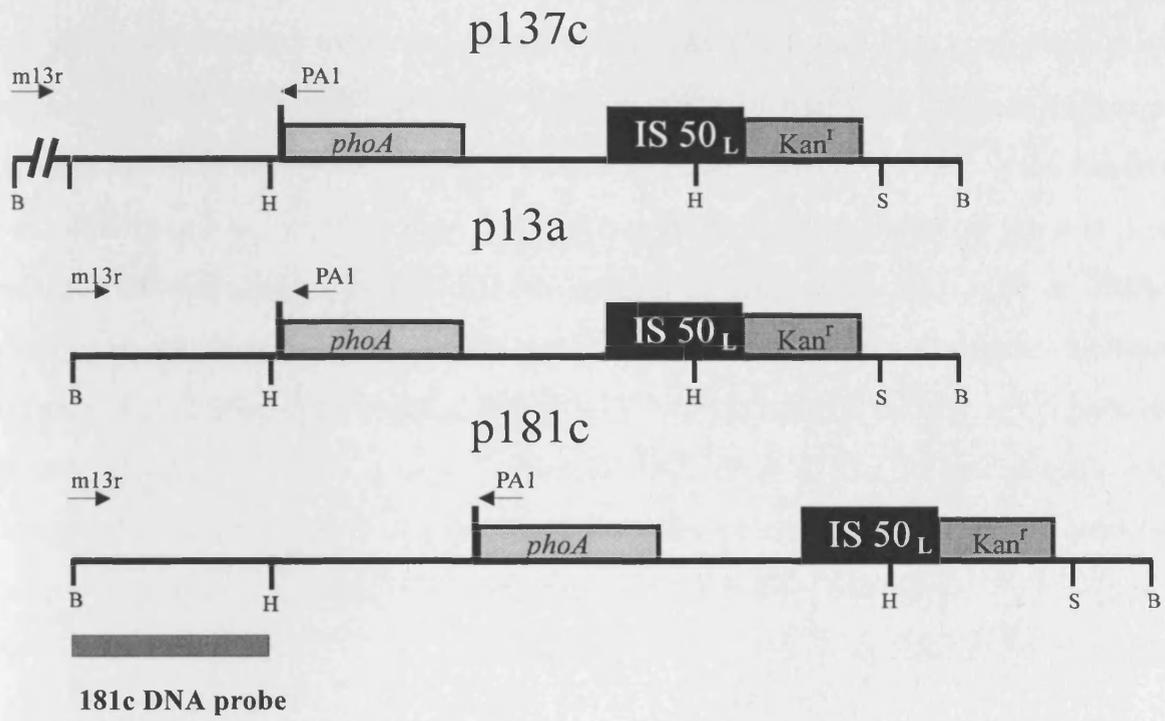


Figure 3.12 Diagrammatic representation of the *Bam*HI inserts of the *phoA*-positive pUC18 clones from the CVD206::*TnphoA* mutants. The sites of the transposon insertions in the CVD206::*TnphoA* mutants were mapped to two separate loci on the basis of DNA sequencing of their pUC18 clones using the primers PA1, m13f and m13r. The *Bam*HI fragments from each of the clones are shown grouped by locus (A) p13a/p137c/p181c and (B) p55a/p138b and aligned to their common upstream *Bam*HI sites (excepting p137c where the *Bam*HI immediately upstream of *TnphoA* is apparently missing); indicated are the *phoA* gene, the kanamycin resistance gene and the IS50_L region of *TnphoA*. The loci specific DNA probes used to screen an E2348-69 chromosomal library are indicated under each group of clones as a shaded bar.

(B - *Bam*HI, H - *Hind*III, S - *Sma*I)

A



B

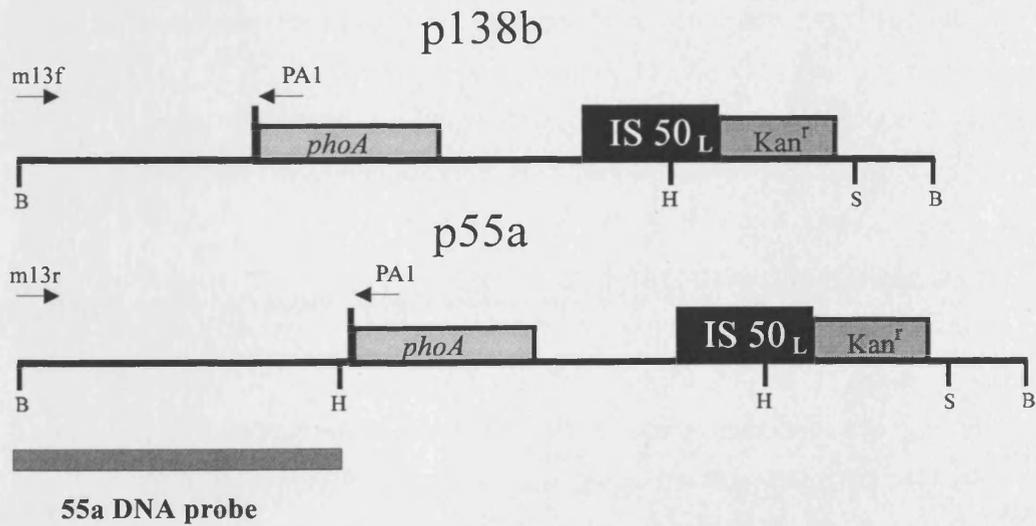


Figure 3.13 Identification of open reading frames in-frame with *phoA* in the CVD206::Tn*phoA* mutants #13, #137 and #181. Plasmids p13a, p137c and p181c (see also **Figures 3.11** and **3.12**) were DNA sequenced using primer PA1 to identify the region directly upstream of their Tn*phoA* insertion. The PA1 DNA sequences (approximately 300 bp) are shown in reverse, i.e. in the direction of transcription of *phoA*, and the deduced amino acid sequence of the reading frame in-frame with *phoA* is given in bold. (A) p13a and p137c gave identical DNA sequences with PA1 and identified an open reading frame which ran throughout the 300 bp sequenced. Indicated is the precise location of the *Hind*III site mapped in **Figure 3.12A**. (B) The DNA sequencing of p181c with PA1 identified a stop codon (TGA, *) 200 bp upstream of the Tn*phoA* insertion and in-frame with *phoA*. Analysis of the sequence identified a potential start codon (ATG) and Shine-Delgarno sequence (SD) which would give a truncated polypeptide of 61 amino acids.

A

GGATAGAAGAAGTGA AATTGAGATGGTTTGCCTTTTGGATTGTGTTATTAGCGGGTTGTT
 1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
 CCTATCTTCTTCACTTTAACTCTACCAAACGGAAAACTAACACAATAATCGCCCAACAA

I E E V K L R W F A F L I V L L A G C S
 CATCAAAGCATGACTATACGAACCCGCCGTGGAACGCGAAAGTCCGGTGCAACGTGCGA
 61 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
 GTAGTTTCGTACTGATATGCTTGGGGCGCACCTTGCCTTTCAAGGCCACGTTGCACGCT

S K H D Y T N P P W N A K V P V Q R A M
 TGCAGTGGATGCCAATAAGCCAGAAAGCCGGTGCAGCCTGGGGCGTCGATCCACAATTGA
 121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
 ACGTCACCTACGGTTATTTCGGTCTTTTCGGCCACGTCGGACCCCGCAGCTAGGTGTTAACT

Q W M P I S Q K A G A A W G V D P Q L I
 TCACGGCGATTATCGCTATCGAATCGGGTGGTAATCCGAACGCGGTGAGTAAATCGAATG
 181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
 AGTGCCGCTAATAGCGATAGCTTAGCCACCATTAGGCTTGCGCCACTCATTTAGCTTAC

T A I I A I E S G G N P N A V S K S N A

HindIII **-> TnphoA**
 CCATTGGTTTGGATGCAGATAAAGCTTCAACCTCCGGACGTGATGCTGACTCTTATAACAC
 241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
 GGTAACCAAACCTACGTCTATTTTCGAAGTTGGAGGCTGCACTACGACTGAGAATATGTG

I G L M Q I K A S T S G R D A D S Y

B

TAAATGCGTGAATATTTTTTACATGTCGTTCTTAAATCAGCCGACATACGCCAGCATCGA
 1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
 TTTACGCACTTATAAAAAATGTACAGCAAGAATTTAGTCCGGCTGTATGCGGGTTCGTAGCT

K C V N I F Y M S F L N Q P T Y A Q H R
 TCCCGCCCTTGCCTATACTTAGAGCGTTGACGTAAGCATTTCTTGCCTCCGATTCATCGA
 61 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
 AGGGCGGGAACGGATATGAATCTCGCAACTGCATTCGTAAGAACGCAGGCTAAGTAGCT

S R P C L Y L E R * R K H F L R P I H R
SD
 AACGAGGGCCTGAACATGCGGAATTATTGCCTGGATTATTTTGGCCTGATAGCCGGCATT
 121 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
 TTGCTCCGGACTTGTACCCTTAATAACGGACCTAATAAAAACCGGACTATCGGCCGTAA

N E G L N M G I I A W I I F G L I A G I
 ATCGCCAAGCTAATCATGCCGGGGCGTGATGGTGGTGGATTTTTTCTGACCTGTATTCTT
 181 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
 TAGCGGTTTCGATTAGTACGGCCCCGCACTACCACCCTAAAAGGACTGGACATAAGAA

I A K L I M P G R D G G G F F L T C I L
-> TnphoA
 GGGATAGTCGGTGCGGTGGTTCGGCGGCTGGCTGGCGACCATGTTTGGCACTGACTCTTAT
 241 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
 CCCTATCAGCCACGCCACCAGCCCGGACCGCTGGTACAAACCGTACTGAGAATA

G I V G A V V G G W L A T M F G T D S Y

Figure 3.14 Identification of open reading frames in-frame with *phoA* in the CVD206::*TnphoA* mutants #55 and #138. Plasmids p55a and p138b (see also Figures 3.11 and 3.12) were DNA sequenced using primer PA1 to identify the region directly upstream of their *TnphoA* insertion. The PA1 DNA sequences (approximately 300 bp) are shown in reverse i.e. in the direction of transcription of *phoA*, and the deduced amino acid sequence of the reading frame in-frame with *phoA* is given in bold. The DNA sequences of both p55a (A) and p138b (B) identified open reading frames in-frame with *phoA* which ran throughout the 300 bp sequenced. Indicated in (A) is the precise location of the *Hind*III site mapped in **Figure 3.12B**.

A

```

GAGTTGAAATTCGATTTCAGATGTGGATTTCTGGAAAGATATTGAGAAGCTCGATAAAACTG
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
CTCAACTTTAAGCTAAGTCTACACCTAAAGACCTTTCTATAACTCTTGAGCTATTTTGAC

E L K F D S D V D F W K D I E N S I K L -

ATACTGGGCAGCGATGGTTCGTATTCAATATCGACAAGTACATCGTCTGTTATTGTGAGA
61 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
TATGACCCGTCGCTACCAAGCATAAGTTATAGCTGTTTCATGTAGCAGACAATAACAGTCT

I L G S D G S Y S I S T S T S S V I V R -

ACATCTTCTCCAAACATGAAGAAGATAAACGAATATATTAATACACTGAATGCACAGCTT
121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
TGTAGAAGAGGTTTGTACTTCTTCTATTTGCTTATATAAATTATGTGACTTACGTGTGAA

T S S P N M K K I N E Y I N T L N A Q L -

GAAAGACAGGTTACCATTGATGTGCGGATCTATAATGTTACAACAAGTGCACAGCAGTGT
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
CTTTCTGTCCAATGGTAACTACAGCGCTAGATATTACAATGTTGTTGACTGTCGCTACTA

E R Q V T I D V A I Y N V T T T D S S D -

HindIII                               -> TnphoA
CTGGCAATGTCTCTGGAAGCTTGCTAAAGCATAACGGAGGTGTTCTGACTCTTATACA
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
GACCGTTACAGAGACCTTCGAAACGATTTTCGTATTGCCTCCACAAGGACTGAGAATATGT

L A M S L E A L L K H N G G V P D S Y T -

```

B

```

TTGGCGTAATCGATAAAAAATATTCTGCATGCTGACACCTCATTGCTGAAGTCGAAGAATA
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
AACCGCATTAGCTATTTTTATAAGACGTACGACTGTGGAGTAACGACTTCAGCTTCTTAT

G V I D K N I L H A D T S L L K S K N K

AAGAACATTATAAAAGCAGTGTCTGGTAAGCAAAACAGACAGTATTTATATAGGAAATA
61 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
TTCTTGTAATATTTTCGTCACTAGACCATTGCTTTGTCTGTCATAAATATATCCTTTAT

E H Y K S S D L V S K T D S I Y I G N S

GCAGTTTTTCAGACATATCATGGAGAGCCCCTGCCCGTAAATTAGAAGGTGTACATGGAA
121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
CGTCAAAAGTCTGTATAGTACCTCTCGGGGACGGGCCATTTAATCTCCACATGTACCTT

S F Q T Y H G E P L P G K L E G V H G I

TTATTCTGAGATCAAGTACACCTCTGGGATTTGATGAAGTACTCTCTATGATTCAGGATT
181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
AATAAGACTCTAGTTCATGTGGAGACCCTAAACTACTTCATGAGAGATACTAAGTCTTAA

I L R S S T P L G F D E V L S M I Q D S
                                           -> TnphoA
CCAGTGGTATTCCCATTGTA AAAACATAACCACGAAAGATGTGATCAGTGGAGGGGCTGACT
241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
GGTCACCATAAGGGTAACATTTTGATGGTGCTTTCTACACTAGTCACCTCCCGGACTGA

S G I P I V K H T T K D V I S G G A D

```

that the promoter regions for each mutant gene were encoded within these fragments; this was demonstrated by the observation that the *TnphoA*-containing clones p13a, p55a, p137c, 138b and 181c were all isolated as *Bam*HI fragments and are all *phoA* positive irrespective of their orientation with respect to the promoter of *lacZ* in pUC18 (see **Figure 3.12**).

3.2.7.1 Design of locus-specific DNA probes. Plasmids p55a and p181c were used to isolate DNA fragments which could be used as probes against the chromosomal loci defined by the two groups of CVD206::*TnphoA* mutants #55/#138 and #13/#137/#181 respectively. Double digestion of either plasmid with *Bam*HI and *Hind*III results in the formation of four DNA fragments: these include two common fragments (pUC18 vector, 2.6 kb, and an internal *TnphoA Bam*HI-*Hind*III fragment, 1.86 kb) and two plasmid specific fragments (a chimeric *Hind*III fragment which is part *TnphoA*/part upstream sequence and a *Bam*HI-*Hind*III fragment which is composed solely of the sequence upstream of *TnphoA* and which would therefore be locus-specific, p55a - 2.3 kb, p181c - 1.4 kb) (**Figures 3.11B and 3.12**). The locus-specific *Bam*HI-*Hind*III fragments from each plasmid were isolated from agarose gels using Polymer wool, as described in Materials and Methods, and labelled to form probes 55a and 181c using either ³²P-dCTP (library screening) or the Fluorescein Gene ImagesTM system (Southern blot analysis).

3.2.7.2 Southern blot analysis of E2348-69 using probes 55a and 181c. Chromosomal DNA was prepared from E2348-69 and *E. coli* K-12 strain DH5 α using the CsCl method. 0.5 μ g of each chromosomal DNA was digested with *Bam*HI and the fragments separated on 1% agarose gel electrophoresis prior to transfer to nylon membranes by Southern blotting. Blots were then hybridised separately under high stringency conditions with either the 55a or 181c probes (see above). Probe 181c identified a single band in both E2348-69 and DH5 α (approximately 6 kb) indicating that this locus is also present in *E. coli* K-12; it appeared however that the *Bam*HI fragment identified in DH5 α was slightly smaller in size than the corresponding band

in EPEC (**Figure 3.15B**, for confirmation of band sizes see also **Figure 3.18B**). In contrast probe 55a identified a band of approximately 5 kb in E2348-69 which was not present in DH5 α , indicating that the locus associated with CVD206::Tn*phoA* mutants #55 and #138 is not in *E. coli* K-12.

3.2.7.3 Construction of an E2348-69 *Bam*HI library in pBluescript. 2 μ g of CsCl purified E2348-69 chromosomal DNA, digested to completion with *Bam*HI, and 0.2 μ g of CsCl purified pBluescript DNA, which had been digested with *Bam*HI, were ligated overnight at 16°C. 1 μ l of the 20 μ l ligation was transformed by electroporation into DH5 α and the resultant transformants selected on LUA plates containing ampicillin and Xgal. The transformation produced approximately 800 colonies of which 80% were white indicating that they carried *Bam*HI insertions; it was therefore estimated that the remainder of the library would realise approximately 15,000 transformants of which 12,000 would contain *Bam*HI inserts. The remainder of the ligation was then transformed into DH5 α and spread onto a single ampicillin LUA plate. The transformant colonies were washed from the plate into 5 ml of LUB; 0.5 ml of culture was retained to be replated immediately for library screening (see below), and the remainder used to produce plasmid DNA by the miniprep method which was stored at -20°C for future use.

3.2.7.4 Screening of the E2348-69 *Bam*HI library using probes 55a and 181c. Stock library culture was diluted and plated onto LUA plates containing ampicillin and Xgal to give approximately 1000 colonies per plate. Four plates (approximately 3200 white colonies) were incubated overnight and then colony blotted to Hybond-NTM membranes as described in Materials and Methods. The colony blots were hybridised at high stringency with an equal mix of 55a and 181c ³²P-dCTP labelled probes and potential positives were identified by autoradiography. Each plate provided several potential positives (screening with ³²P-dCTP labelled probes also has a tendency to produce false positives) and all white colonies in the vicinity of those positives were patchplated to fresh LUA plates and onto duplicate Hybond-NTM membranes. The

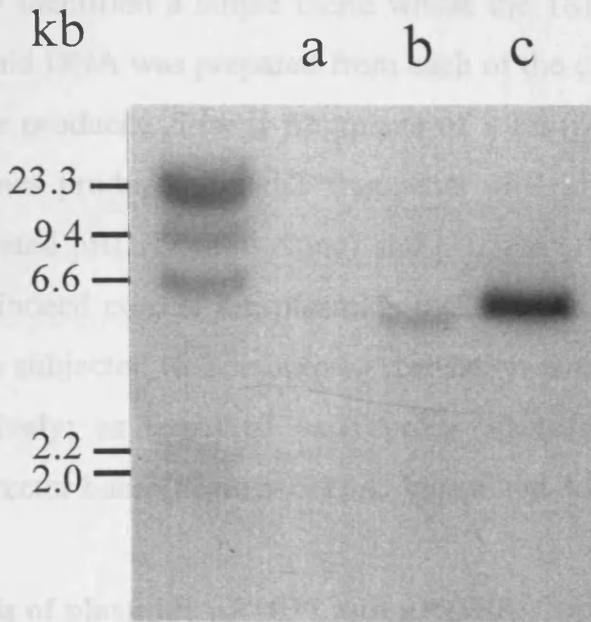
3.15 Southern blot analysis of E2348-69 and DH5 α chromosomal DNA using probes specific to the CVD206::*TnphoA* mutant loci. CsCl prepared chromosomal DNA from E2348-69 and *E. coli* K-12 DH5 α , and plasmid DNA from potential pBluescript/E2348-69 *Bam*HI chromosomal library clones pRDH7 and pRDH8 were digested with *Bam*HI and the fragments separated by 1% agarose gel electrophoresis prior to Southern transfer to Hybond-NTM membranes. The blots were hybridised with the CVD206::*TnphoA* mutant locus-specific probes 55a and 181c (section 3.2.7.1; **Figure 3.12**) as indicated below.

(A) DH5 α (lane a) E2348-69 (lane b) pRDH7 (lane c) - probe 55a.

(B) DH5 α (lane a) E2348-69 (lane b) pRDH8 (lane c) - probe 181c.

A

membranes were processed as described in section 3.2.1.1 and the colony blots probed were hybridised separately with either the 55a or 181c ³²P-dCTP labelled probe. The 55a probe identified a single clone whilst the 181c probe identified two separate clones. Plasmid DNA was prepared from each of the clones and digested with BamHI: the 55a clone produced a single band of 5.4 kb (pRDH7) and 5.4 kb, whilst both 181c clones produced bands of 5.4 kb and 6.3 kb. The two plasmids were designated pRDH7 and pRDH8 (the 181c clone). To confirm that the clones were indeed clones of the same gene, pRDH7 and pRDH8 were digested with BamHI and then subjected to Southern blot analysis using the 55a and 181c probes respectively. The results are shown in Figure 3.2.7.2. The closed BamHI fragment between the two probes is 2.2 kb (pRDH7, lane a).

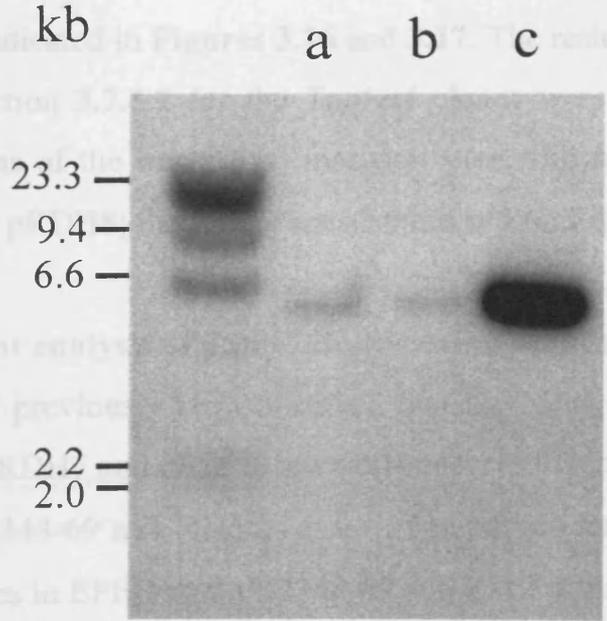


3.2.8 Further analysis of plasmids

3.2.8.1 Restriction enzyme-mapping of pRDH7 and pRDH8. Both pRDH7 and pRDH8 plasmids were prepared using EcoRI gradients and were mapped for useful restriction enzyme sites (i.e. those containing sites found within the phagescript MCS: SphI and BstXI). Multiple restriction enzyme sites were mapped using double digests with enzymes which had single sites; the approximate positions of all the restriction sites identified are indicated in Figures 3.2.7.3 and 3.2.7.4.

B

3.2.8.2 Southern blot analysis of pRDH7 and pRDH8. The plasmids were digested with BamHI and the resulting fragments were separated on an agarose gel. The gel was stained with ethidium bromide and the bands were visualised under short wave UV light. The results are shown in Figure 3.2.7.5. The closed BamHI fragment between the two probes is 2.2 kb (pRDH7, lane a).



3.2.8.3 Southern blot analysis of pRDH7 and pRDH8

and pRDH8. It had previously been shown that the plasmids pRDH7 and pRDH8 were derived from pRDH7 and pRDH8. To confirm this, Southern hybridisation analysis was performed using the 55a and 181c probes. The results are shown in Figure 3.2.7.6. The closed BamHI bands of pRDH7 and pRDH8 were 5.4 kb and 6.3 kb respectively. The 55a probe did not hybridise to pRDH8 DNA at all (section 3.2.7.2). To confirm and clarify these

membranes were processed as described in section 2.11.1 and the colony blots produced were hybridised separately with either the 55a or 181c ³²P-dCTP labelled probe. The 55a probe identified a single clone whilst the 181c probe identified two separate clones. Plasmid DNA was prepared from each of the clones and digested with *Bam*HI; the 55a clone produced *Bam*HI fragments of 3 kb (pBluescript) and 5.4 kb, whilst both 181c clones produced *Bam*HI fragments of 3 kb and 6.3 kb. The two plasmids were designated pRDH7 (55a clone) and pRDH8 (181c clone). To confirm that the clones were indeed correct the plasmids pRDH7 and pRDH8 were digested with *Bam*HI and then subjected to Southern hybridisation analysis using the 55a and 181c probes respectively; as expected each probe identified the cloned *Bam*HI fragment but not the vector band (Figures 3.15A, lane c and 3.15B, lane c).

3.2.8 Further analysis of plasmids pRDH7 and pRDH8.

3.2.8.1 Restriction enzyme mapping of pRDH7 and pRDH8. Both pRDH7 and pRDH8 plasmids were prepared using CsCl gradients and were mapped for useful restriction enzyme sites (i.e. those hexameric sites found within the pBluescript MCS, *Sph*I and *Bgl*II). Multiple restriction enzyme sites were mapped using double digests with enzymes which had single sites; the approximate positions of all the restriction sites identified are indicated in Figures 3.16 and 3.17. The restriction enzyme mapping data obtained in section 3.2.5.2 for the *TnphoA* clones were used to determine the approximate locations of the transposon insertion sites with respect to the restriction maps of pRDH7 and pRDH8; these sites are also marked on Figures 3.16 and 3.17.

3.2.8.2 Southern blot analysis of EPEC strains using probes derived from pRDH7 and pRDH8. It had previously been observed that the locus specific probes used to clone the plasmids pRDH7 and pRDH8 gave different results in Southern hybridisation analysis between E2348-69 and DH5 α i.e. the 181c probe identified *Bam*HI bands of slightly different sizes in EPEC strain E2348-69 and K-12 strain DH5 α , and probe 55a did not hybridise to DH5 α DNA at all (section 3.2.7.2). To confirm and clarify these

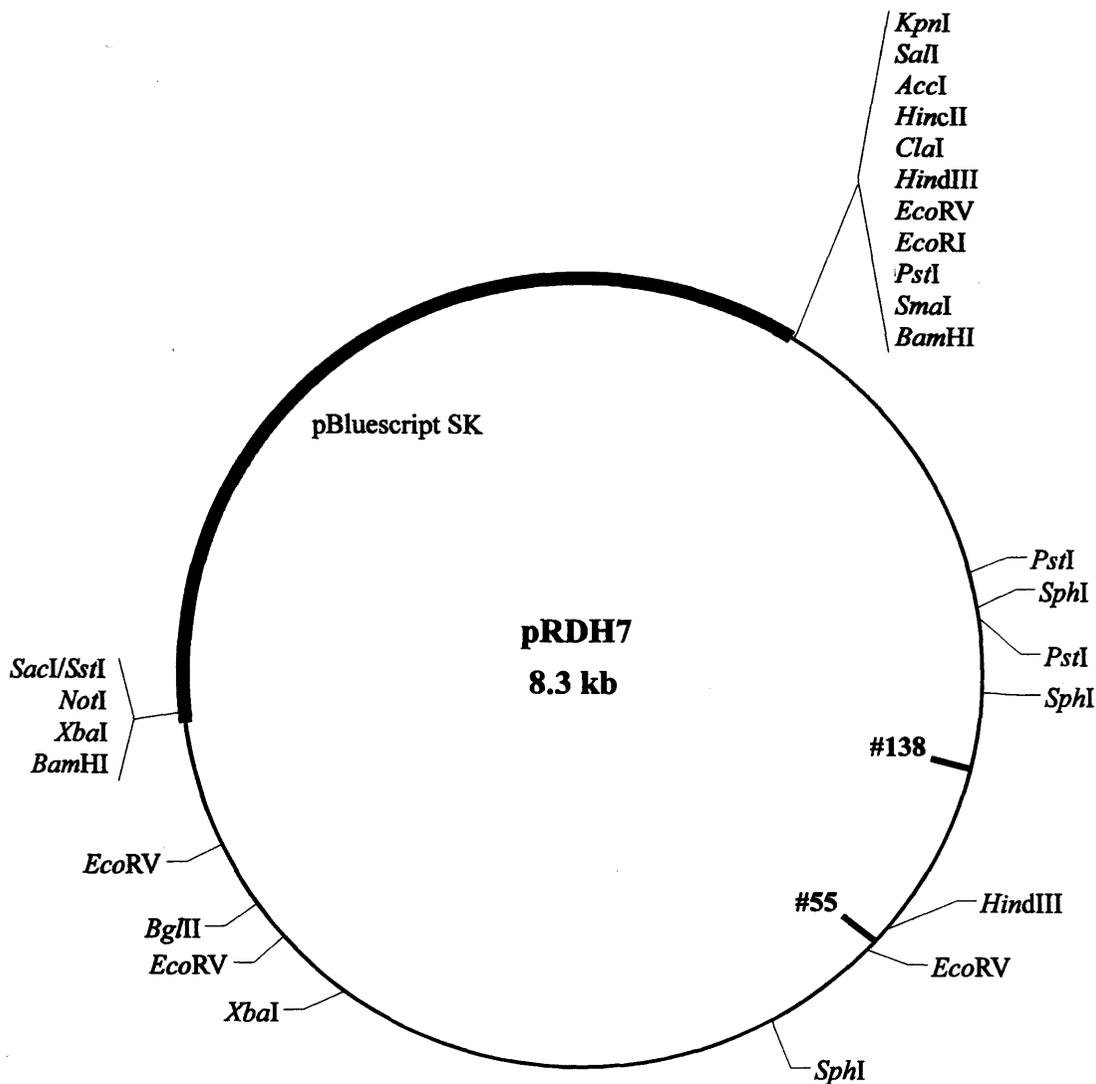


Figure 3.16 Map of plasmid pRDH7. The pBluescript/E2348-69 *Bam*HI library clone pRDH7 was mapped using the commonly used hexameric restriction enzymes indicated above. The *Bam*HI insert was estimated to be approximately 5.4 kb. The equivalent sites at which *TnphoA* is inserted in CVD206::*TnphoA* mutants #55 and #138 are indicated.

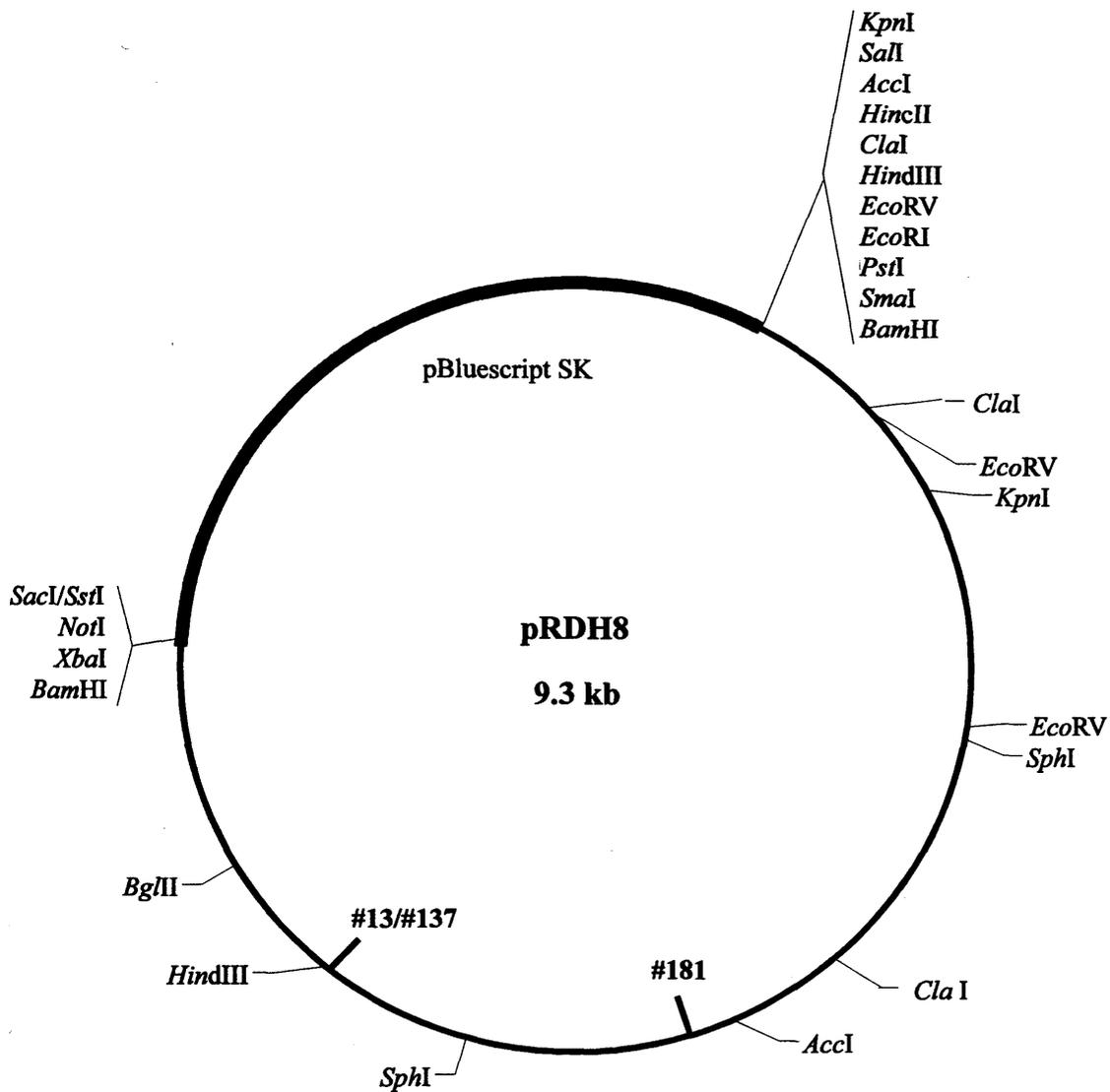


Figure 3.17 Map of plasmid pRDH8. The pBluescript/E2348-69 *Bam*HI library clone pRDH8 was mapped using the commonly used hexameric restriction enzymes indicated above. The *Bam*HI insert was estimated to be approximately 6.3 kb. The equivalent sites at which *TnpA* is inserted in CVD206::*TnpA* mutants #13/137 and #181 are indicated.

results it was proposed to prepare DNA probes which included the entire locus, which had been cloned in pRDH7 and pRDH8, for use in Southern blot analysis.

Plasmids pRDH7 and pRDH8 were digested with *Bam*HI and the resultant fragments separated by 1% agarose gel electrophoresis; the 5.4 kb and 6.3 kb *Bam*HI inserts, respectively, were cut from the gel, purified using Polymer wool and labelled to form probes using the Fluorescein Gene Images™ kit.

Chromosomal DNA from each of the CVD206::*TnphoA* mutants, E2348-69, DH5 α and EAF⁻ strain MAR001 were digested with *Bam*HI, separated on 0.8% agarose gel electrophoresis and transferred to Hybond-N™ membranes by Southern blotting. The Southern blots were then hybridised with the pRDH7 and pRDH8 probes described above. The pRDH8 probe hybridised with a single *Bam*HI band of approximately 6.6 kb in E2348-69, a smaller band of approximately 6 kb in DH5 α , and in each of the CVD206::*TnphoA* mutants #13, #137 and #181 the probe identified a doublet which represented the *Bam*HI fragments from the 5' and 3' regions of the transposon insertion (Figure 3.18B). The pRDH7 probe also identified *TnphoA* derived *Bam*HI fragment doublets in the the CVD206::*TnphoA* mutants #55 and #138 and a band of approximately 6 kb in E2348-69; it did not, however, identify hybridising fragments in DH5 α or the EPEC strain MAR001 (Figure 3.18A). The absence of hybridisation with MAR001 strongly suggested that the locus cloned in pRDH7 was from the EPEC EAF plasmid pMAR2; however this required further confirmation as it has been observed that MAR001 has also lost the 5.5 kb "cryptic" plasmid found in E2348-69 (R. Haigh, unpublished observation) and which is just large enough to encode the pRDH7 *Bam*HI insert.

3.2.9 Attempt to complement the mutations in the CVD206::*TnphoA* mutants by introduction of plasmids pRDH7 or pRDH8.

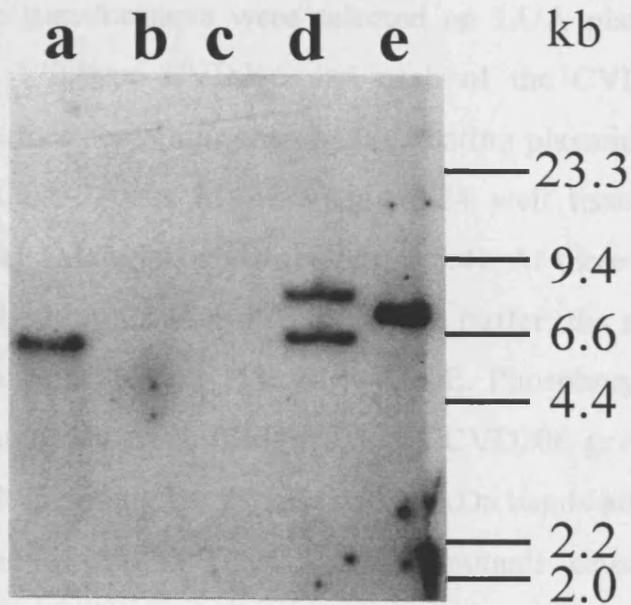
In theory the wildtype genes cloned in pRDH7 and pRDH8 should be able to complement the mutations in their respective CVD206::*TnphoA* mutants. CsCl plasmid

Figure 3.18 Southern blot analysis of EPEC strains using pRDH7 and pRDH8 specific probes. CsCl prepared EPEC chromosomal DNA was digested with *Bam*HI, separated by 0.8% agarose gel electrophoresis and transferred to Hybond-N™ membranes by Southern blotting. Individual blots were hybridised with pRDH7 (5.4 kb *Bam*HI insert) or pRDH8 (6.3 kb *Bam*HI insert) specific probes as indicated below.

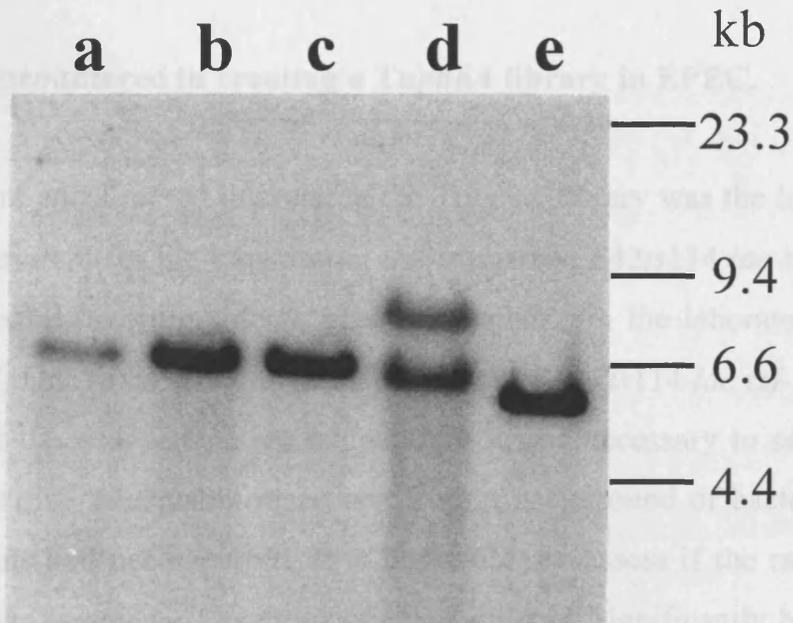
(A) E2348-69 (lane a) MAR001 (EAF⁻) (lane b) DH5α (lane c) CVD206::*TnphoA* #55 (lane d) CVD206::*TnphoA* #138 (lane e) - probe pRDH7.

(B) E2348-69 (lane a) CVD206::*TnphoA* #13 (lane b) CVD206::*TnphoA* #137 (lane c) CVD206::*TnphoA* #181 (lane d) DH5α (lane e) - probe pRDH8.

A



B



preparations of pRDH7 and pRDH8 were used in electroporations to transform CVD206::*TnphoA* mutants #55 and #138, or CVD206::*TnphoA* mutants #13, #137 and #181, respectively; positive transformants were selected on LUA plates containing kanamycin and ampicillin. Wildtype CVD206 and each of the CVD206::*TnphoA* mutants together with its partner containing the complementing plasmid were used to infect fully differentiated Caco-2 cells in the standard 24 well tissue culture cell protein-phosphorylation assay (Materials and Methods, 2.5.4). At three and six hours infected Caco-2 cells were lysed in 2x SDS-PAGE sample buffer, the samples boiled for 10 min and the proteins separated on 15% SDS-PAGE. Phosphorylated proteins were then identified using autoradiography (Figure 3.19). CVD206 gave the standard EPEC phosphorylation result including the 29 kDa and 21 kDa bands after three hours infection; as expected none of the CVD206::*TnphoA* mutants caused significant phosphorylation, even after six hours infection. Significantly none of the mutants carrying pRDH7 or pRDH8 induced any protein phosphorylation indicating that the cloned loci were insufficient in themselves to complement the transposon mutations.

3.3 Discussion

3.3.1 Problems encountered in creating a *TnphoA* library in EPEC.

The major problem encountered in creating the *TnphoA* library was the lack of a good plasmid delivery system for the transposon; unfortunately, F42*ts114 lac zzf-1::TnphoA* was the only *TnphoA*-carrying suicide plasmid available in the laboratory when this work commenced. Due to the continued maintenance of F42*ts114 lac zzf-1::TnphoA* in strain CVD206 at the non-permissive temperature it was necessary to select colonies carrying *phoA*-positive transposon insertions from a background of bacteria in which transposition events had not occurred. It is impossible to assess if the randomness of the selection of *phoA*-positive transpositions was affected significantly by the system used; however, it seems probable that *TnphoA* insertions into very weakly expressed genes or *phoA* fusions which showed only low alkaline phosphatase activity would

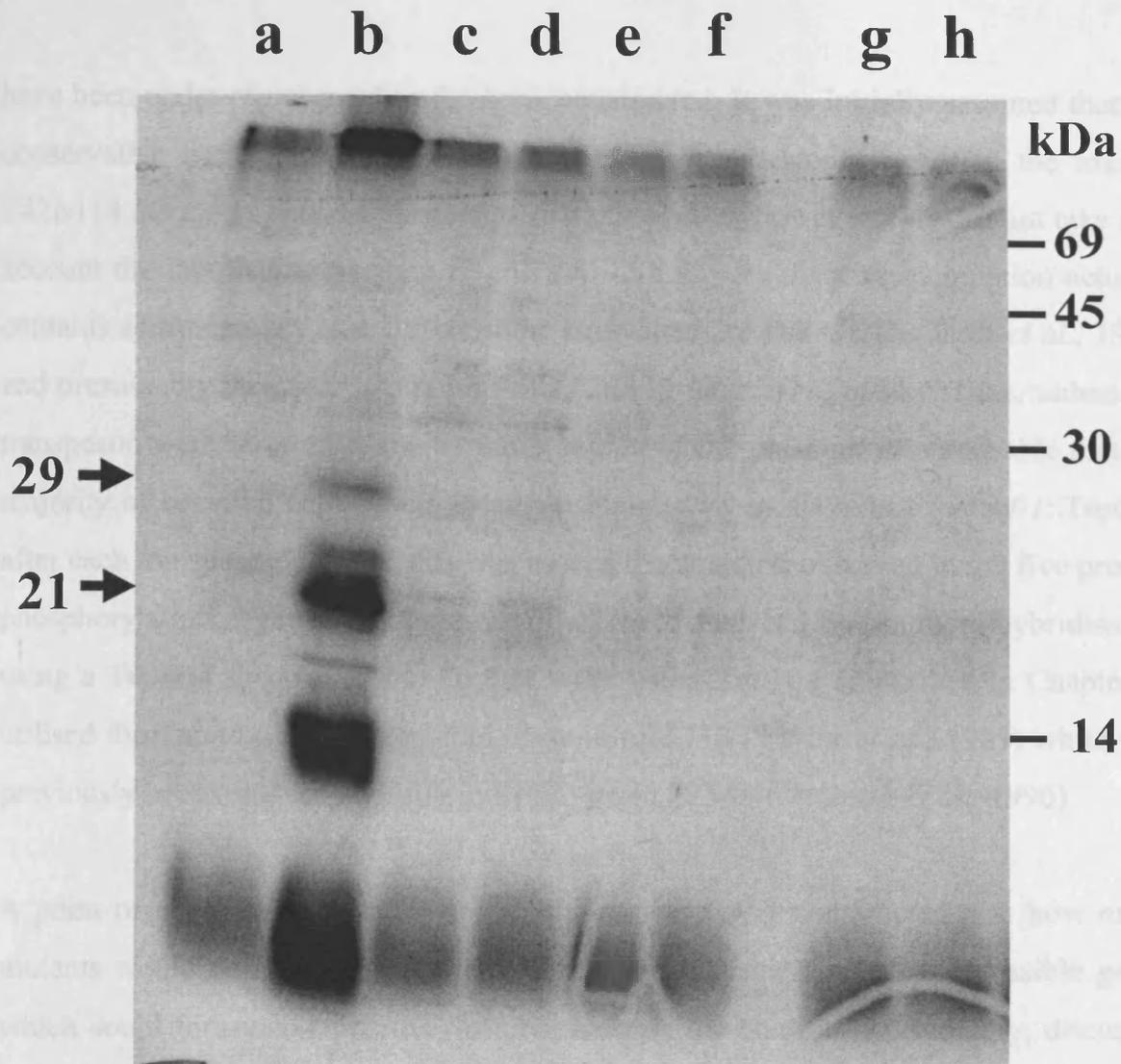


Figure 3.19 Protein phosphorylation profiles of Caco-2 cells infected with CVD206::*TnphoA* mutants complemented with pRDH7 or pRDH8. Confluent and differentiated Caco-2 cells grown in 24 well tissue culture dishes were preincubated in MEM containing $^{32}\text{P}_i$ phosphate (0.1 mCi/ml) for 4 hours prior to infection with approximately 1×10^7 bacteria and incubated for 3 hours at 37°C under 5% CO_2 . Nonadherent bacteria were washed free and the Caco-2 cells lysed in 2x SDS-PAGE buffer. Samples were boiled for 10 min, the proteins separated by 15% SDS-PAGE and phosphorylated proteins determined by autoradiography (see also Materials and Methods, section 2.5.4). Gel loading order: uninfected Caco-2 (lane a), CVD206 (lane b), CVD206::*TnphoA* #181 pRDH8 (lane c), CVD206::*TnphoA* #138 pRDH7 (lane d), CVD206::*TnphoA* #55 pRDH7 (lane e), CVD206::*TnphoA* #181 (lane f), CVD206::*TnphoA* #138 (lane g), CVD206::*TnphoA* #55 (lane h). (Data for CVD206::*TnphoA* #13 pRDH8 and CVD206::*TnphoA* #137 pRDH8 are not shown)

have been under-represented in the bank constructed. It was initially assumed that the conservative nature of the *TnphoA* transposition event would result in the loss of F42ts114 *lac zzf-1::TnphoA* plasmid after transposition; however, this did not take into account the fact that an average *E. coli* cell in a rapidly dividing population actually contains approximately two chromosome equivalents of DNA (Neidhardt *et al.*, 1990) and presumably therefore two copies of F42ts114 *lac zzf-1::TnphoA*. Thus, unless the transposon were inserted in the terminus region of the plasmid, it is probable that the majority of bacterial cells would retain an intact copy of F42ts114 *lac zzf-1::TnphoA* after each transposition event; this was indeed the situation observed in the five protein phosphorylation negative mutants when they were analysed by Southern hybridisation using a *TnphoA* specific probe. Further work using *TnphoA* (described in Chapter 4) utilised the *TnphoA*-carrying suicide plasmid pRT733 (Taylor *et al.*, 1989) which has previously been used successfully in EPEC strain E2348-69 (Jerse *et al.*, 1990)

A point of concern when the *TnphoA* library was being constructed was how many mutants would be required in order to represent accurately all of the possible genes which could form *phoA*-positive fusions. Despite the continued possibility, discussed above, that the *TnphoA* mutant bank might have been under-represented with weakly-expressed genes or alkaline phosphatase fusions, it appears that the figure of 192 mutants used in the primary screen was adequate to cover all of the mutants which could be identified using the F42ts114 *lac zzf-1::TnphoA* based mutation system employed. This conclusion is based upon the observation that though five independent protein phosphorylation negative mutants were isolated they represent only two loci; furthermore the transposon insertions within these loci map to only 600 bp and 1.4 kb apart respectively, indicating repeat insertions either within the same gene or within closely linked genes.

3.3.2 A role for bundle-forming pili (BFP) in the induction of host cell protein phosphorylation ?

Five of the CVD206::*TnphoA* mutants isolated were determined to be unable to induce serine-threonine protein phosphorylation in either HEp-2 or Caco-2 cells, even after extended incubation for six hours. Each of phosphorylation-negative mutants were wildtype for secretion of EPEC virulence proteins and FAS test positive, retaining the ability to cause the formation of attaching and effacing lesions on HEp-2 cells (when complemented with a plasmid encoded *eaeA* gene); however, they were all found to be completely deficient in the production of BFP. The BFP⁻ phenotype was consistent with the observed marked reduction in the level of adherence and invasion with HEp-2 cells and the change in the pattern of adherence from localised to diffuse. Western blot analysis demonstrated that all of the mutants continued to express the BfpA protein (bundlin) indicating that their failure to produce BFP resulted from a defect in pili biogenesis and was not due to altered regulation of the *bfpA* gene.

Analysis of cloned DNA fragments encoding the regions upstream of each of the *phoA*-positive *TnphoA* insertions revealed that the five mutations were in novel genes and that these were localised within two loci. The wildtype loci were cloned from an EPEC/pBluescript chromosomal library as *Bam*HI fragments to form the recombinant plasmids pRDH7 (5.3 kb insert) and pRDH8 (6.4 kb insert). A DNA probe derived from pRDH7, which contains the insertion sites for the CVD206::*TnphoA* mutants #55 and #138, failed to hybridise with the *E. coli* K-12 strain DH5 α , nor with the plasmid-deleted EPEC strain MAR001, indicating that the pRDH7 locus was not only EPEC-specific but also plasmid encoded. As strain E2348-69 carries two plasmids it was initially unclear which encoded the pRDH7 locus; however, comparison of the restriction enzyme map of plasmid pRDH7 with that of the E2348-69 5.5 kb "cryptic" plasmid (S. Clarke, personal communication) revealed incompatible restriction enzyme sites (e.g. *Xba*I) indicating that the *TnphoA* insertions in mutants #55 and #138 were encoded on the EAF plasmid pMAR2. Plasmid pMAR2 had previously been shown to

encode two genes, *bfpA* and *bfpP*, which were separated by approximately 8 kb but which were both required for BFP production and expression of the LA phenotype (Donnenberg *et al.*, 1992; Zhang *et al.*, 1994). Zhang *et al.* (1994) had predicted that, analogous to the *tcp* operon of *Vibrio cholerae* (Oigerman *et al.*, 1993), the region between *bfpA* and *bfpP* would encode genes required for BFP biogenesis. It was therefore proposed, based upon the plasmid location and the BFP⁻ phenotype, that the *TnphoA* insertions in mutants #55 and #138 represented genes in the BFP biogenesis operon.

A DNA probe derived from the plasmid pRDH8, which contains the insertion sites of the CVD206::*TnphoA* mutants #13/#137 and #181, was found to hybridise with chromosomal DNA derived from *E. coli* K-12 strain DH5 α ; this indicated that the pRDH8 locus was not EPEC-specific and was probably chromosomally located. The pRDH8-specific *Bam*HI fragment identified in DH5 α was slightly smaller in size than in EPEC; however, it was unclear if this was simply due to the loss of a *Bam*HI site in the EPEC strain or if there were real chromosomal differences. Donnenberg *et al.* (1990) had previously reported the isolation of E2348-69 mutants with chromosomally encoded *TnphoA* insertions which were both non-invasive and unable to produce BFP. These mutations were subsequently found to be in the gene *dsbA* (Zhang and Donnenberg, 1996), whose protein product is required for the formation of cysteine disulphide bonds in periplasmic proteins (Bardwell *et al.*, 1991), and is essential for the stability of BfpA prior to its assembly into BFP (Zhang and Donnenberg, 1996; Donnenberg *et al.*, 1997). However neither of the open reading frames identified adjacent to the *TnphoA* insertions in mutants #13/#137 and #181 showed any homology with the *dsbA* gene nor with the *E. coli* K-12 87.1 min region where it is located. It has been predicted that the biogenesis of the *Pseudomonas aeruginosa* type IV pilus requires the protein products of between 20 and 40 genes (Hobbs *et al.*, 1993); therefore it is probable that BFP will require a similar quantity, some of which will undoubtedly be genes common to *E. coli* K-12.

Significantly, despite good circumstantial evidence that the the pRDH7 and pRDH8 plasmids encoded complete wildtype genes and their promoter sequences, their introduction into the respective CVD206 mutants failed to complement the *TnphoA* mutations. The most probable explanation for this result is that the mutated genes are organised as part of an operon, and that the transposon insertions result in polar effects upon downstream genes; this would then require the introduction of the complete operon to achieve complementation.

The *TnphoA* mutant screen described in this chapter was intended to identify EPEC genes specifically required for the induction of serine-threonine phosphorylation of host cell proteins such as MLC. Whilst each of the five mutants identified were indeed phosphorylation-negative it seems probable that this phenotype was due to the massive reduction in the adherence of the mutant strains to the HEp-2 cells rather than the loss of a specific phosphorylation induction mechanism itself. The reduction in HEp-2 adherence was entirely consistent with the observed loss of BFP production (Donnenberg *et al.*, 1990; Donnenberg *et al.*, 1992; Sohel *et al.*, 1993); however in the absence of any published work on bacteria-to-host signalling via pili or fimbriae it seems unlikely that the binding of BFP is directly responsible for the induction of host cell protein phosphorylation. Furthermore data have been obtained previously which demonstrated that a plasmid free (and therefore BFP-negative) EPEC strain, MAR001, was still able to induce phosphorylation of 21 kDa and 29 kDa proteins after extended incubation to six hours (Haigh *et al.*, 1995; **Figure 3.1**). The expression of intimin has been shown not to be specifically required for host cell phosphorylation (Haigh *et al.*, 1995) in a wildtype EPEC background; however it seems probable that an additional decrease in HEp-2 cell adherence due to the *eaeA* mutation in CVD206 may explain the difference between the phosphorylation phenotypes of MAR001 and the BFP⁻ CVD206::*TnphoA* mutants.

3.3.3 Future work.

From the results obtained in this chapter it appears that the *TnphoA* mutagenesis strategy adopted has failed to identify mutants which are specifically deficient in host cell phosphorylation. There may be one of two possible explanations for this; (i) the proteins required for phosphorylation induction are weakly expressed and therefore their genes were not represented in the original bank of *TnphoA* mutants, or (ii) the induction of host cell phosphorylation does not require proteins which are exported via the Sec-dependent pathway and therefore it is impossible to obtain *phoA*-positive *TnphoA* mutants in this process. As discussed above the lack of representation of genes within the *TnphoA* mutant library can be easily resolved by the use of a more EPEC-friendly suicide vector delivery system; however, if *phoA*-positive phosphorylation deficient mutants were impossible to isolate this would mean the necessity for a much larger library of mutants which, using the current radioactive methods, would be impractical to screen. However further screening of transposon mutants may not be required due to the recent completion of the DNA sequence of the LEE pathogenicity island from EPEC strain E2348-69 (Elliot *et al.*, 1998a). Despite the demonstration that the expression of the *eaeA* product intimin and the formation of AE lesions is not a prerequisite for host cell phosphorylation (Haigh *et al.*, 1995) it has been shown previously that the phosphorylation phenotype was restricted to AE lesion-forming *E. coli* (Baldwin *et al.*, 1990). As all AE lesion-forming *E. coli* carry the LEE pathogenicity island (McDaniel *et al.*, 1997) it is tempting to speculate that the genes required for induction of host cell phosphorylation might also be encoded within the LEE. Furthermore in their recent report Elliot *et al.* (1998b) describe the identification, by systematic mutagenesis, of genes within the LEE which are not essential for AE lesion formation; they then speculate that these genes may encode other EPEC virulence functions. Screening of mutants in these genes for the loss of the ability to induce host cell phosphorylation may assist in the assignment of a definite role for the phosphorylation of MLC in EPEC pathogenesis.

Chapter 4. DNA sequencing and characterisation of the genes present in plasmids pRDH7 and pRDH8.

4.1. Introduction.

The previous chapter described the isolation and characterisation of five CVD206::Tn*phoA* mutants which were unable to elicit a protein phosphorylation response in cultured human epithelial cells. DNA probes derived from these mutants allowed the identification and cloning of two loci, one chromosomal and the second apparently encoded on the EPEC EAF plasmid pMAR2. Preliminary DNA sequence from the two cloned *Bam*HI fragments in plasmids pRDH7 (5.4 kb) and pRDH8 (6.3 kb) suggested that they largely contained novel sequence. This chapter describes the strategy adopted for the rapid DNA sequencing of the complete *Bam*HI inserts of both pRDH7 and pRDH8, and the analysis and partial characterisation of the genes determined.

A number of strategies have been developed in order to sequence cloned DNA fragments in the size range 1 kb to 10 kb; the choice of method is largely dependent upon the time, money and the quality of DNA sequencing available. The simplest strategy involves subcloning smaller fragments using suitable existing restriction enzyme sites and primer walking across these subclones. Primer walking involves sequencing from a known primer site (usually commencing with vector-derived primers) followed by design of an oligonucleotide homologous to the end of the sequence obtained to serve as primer for the next sequencing step; the process is repeated until one strand of the fragment is determined. This method is both time-consuming and expensive because primer design is dependent upon data yet to be obtained and it requires the production of numerous untested oligonucleotide primers, some of which may prove unreliable. A more directed approach involves the construction of nested deletions (Henikoff, 1987) using exonuclease III. The

exonuclease III enzyme catalyses the removal of 5' mononucleotides from recessed or blunt 3'-hydroxyl termini of double stranded DNA; however protruding 3' termini are completely resistant (Weiss, 1976). Therefore by careful choice of restriction enzymes it is possible to linearise a plasmid such that the linear fragment will only be sensitive to exonuclease III digestion at one end; after exonuclease III digestion the resultant single stranded DNA is removed using S1 nuclease and the plasmid religated, now containing a deletion. By varying the length of the exonuclease III step it is possible to create a series of deletions of increasing length across a cloned fragment which can then be used with a single well-characterised vector-derived oligonucleotide primer to sequence the fragment. As a simpler alternative to making nested deletions, a number of researchers have used transposon sequencing as a fast and effective way to sequence cloned DNA fragments (Liu *et al.*, 1984; Sedgewick and Morgan, 1994). This method involves the creation of a bank of random transposon insertions within the DNA fragment to be sequenced; these transposon insertions can then be used as priming sites, using an oligonucleotide primer which binds within the terminus of the transposon, to obtain DNA sequence from multiple sites within the fragment (it also allows sequencing of both strands at once depending upon the orientation of the transposon). As with the Henikoff method, only characterised oligonucleotide primers are used for sequencing reactions and many such reactions can be run concurrently, therefore making the method both inexpensive and rapid.

The strategy adopted for sequencing the cloned inserts in pRDH7 and pRDH8 was an adaptation of the transposon sequencing method described above. However the choice of *TnphoA* as the transposon used also allowed the concurrent identification of genes encoding exported proteins. *TnphoA* was introduced into plasmids pRDH7 and pRDH8 using a suicide vector system and transconjugant plasmids which had *TnphoA* integrated at regular intervals (400-500 bp) within the *Bam*HI inserts were identified for DNA sequencing. Primers which anneal to DNA sequences present within each of the termini of *TnphoA* were then used to sequence from the transposon into the DNA adjacent to either side of the insertion site. As each transposon insertion gave

approximately 600 bp DNA sequence on one strand in one direction and the same on the other strand in the opposite direction it was possible to rapidly form a series of overlapping sequences on both strands.

4.2 Results.

4.2.1 DNA sequencing of pRDH7 and pRDH8.

4.2.1.1 Construction of *TnphoA* insertions in plasmids pRDH7 and pRDH8.

TnphoA was introduced into CC118 pRDH7 and CC118 pRDH8 by conjugation with strain SM10 λ *pir* pRT733 as described in Materials and Methods. pRT733 is a *TnphoA*-containing derivative of the narrow host range vector pJM703.1 (Miller and Mekalanos, 1988; Taylor *et al.*, 1989); it is dependent upon the presence of the π protein (provided *in trans* by the presence of λ *pir* transducing phage in its host strain; Kolter *et al.*, 1978) for replication from its *oriR6K* origin of replication. Thus after transfer to CC118, which does not contain λ *pir*, the vector is unable to replicate and selection for the kanamycin resistance gene of *TnphoA* directly selects for transposition events. Conjugation mixtures of SM10 λ *pir* pRT733 with CC118 pRDH7 or CC118 pRDH8 were spread onto LUA plates containing rifampicin, kanamycin and ampicillin to select for CC118 containing both the plasmid and *TnphoA*. Identification of transconjugants where *TnphoA* had integrated into the plasmid (as opposed to those containing plasmid but with *TnphoA* integrated into the CC118 chromosome) required transfer of the plasmid to a fresh background strain. Colonies were washed from the plate into LUB and DNA was prepared using the plasmid miniprep method. Plasmid DNA was then transformed into a fresh CC118 strain using the low efficiency CaCl₂ method (this would achieve strains carrying single plasmid variants) and colonies were selected on ampicillin, kanamycin LUA plates. 100 colonies were picked at random from both the pRDH7 and the pRDH8 plates and restreaked onto ampicillin, kanamycin, XP, LUA plates; these were designated pRDH7-1 to pRDH7-100 and

pRDH8-1 to pRDH8-100. Blue colonies, indicating transposition events into genes encoding exported proteins (14/100 for pRDH7, 4/100 for pRDH8; see **Figure 4.1**), were selected and their plasmids purified by the CsCl method.

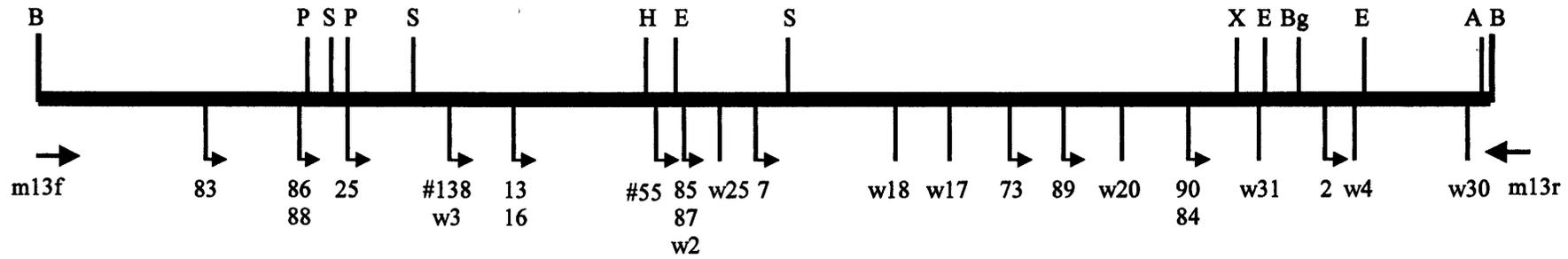
4.2.1.2 DNA sequencing of *phoA* positive *TnphoA* insertions in pRDH7 and pRDH8. The pRDH7 and pRDH8 derived plasmids identified to contain *phoA* positive *TnphoA* insertions were sequenced using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit as described in Materials and Methods. The primers used were PA1 as previously described (section 3.2.6; section 2.12) and PB1 which was designed to anneal approximately 60 bp within the 3' end of *TnphoA*, thus allowing determination of DNA sequences downstream of the transposon (section 2.12; **Figure 3.2**). Sequencing reactions were performed in duplicate and each sequence pair compared using the BESTFIT program (section 2.15) to determine the extent of the reliable sequence; typically this was 500-600 bp for each sequence pair.

4.2.1.3 Identification of *bfpA* in pRDH7. The 14 pairs of DNA sequences determined from the *phoA* positive *TnphoA* insertions in pRDH7 were analysed using the programs BESTFIT and GAP (section 2.15). From the data obtained it was immediately possible to align the sequences to form two regions of contiguous sequence (contigs), one of which overlapped with the sequence from the m13f primer and the other with the m13r primer. The *TnphoA* insertions in the mutants #138 and #55 were determined to be between insertions pRDH7-25 and pRDH7-16, and pRDH7-13 and pRDH7-87 respectively (**Figure 4.1**). The DNA sequences were individually compared against the GenBank database using the BLAST program (section 2.15). The majority gave no significant homology; however 3 gave 99% identity with sequences for the gene *bfpA*, which encodes the major structural subunit of the bundle-forming pilus (BFP) in EPEC (Donnenberg *et al.*, 1992; Sohel *et al.*, 1993). The *TnphoA* insertions in pRDH7-86 and pRDH7-88 were inserted approximately 40 bp downstream of the published *bfpA* sequence (GenBank Accession number Z12295; Donnenberg *et al.*, 1992) whilst pRDH7-83 was inserted within the

Figure 4.1 Diagrammatic representation of the *TnphoA* insertions used for transposon sequencing of plasmids pRDH7 and pRDH8. *TnphoA* insertions are indicated by vertical lines below the the plasmid map; for *phoA*-positive insertions the orientation of the *phoA* gene is also indicated by an arrow (exact locations of all *TnphoA* insertions are shown in Appendix 1 and Appendix 2). The sites of annealing and orientation of primers m13f, m13r and the primers used to complete the pRDH8 DNA sequence are indicated by horizontal arrows.

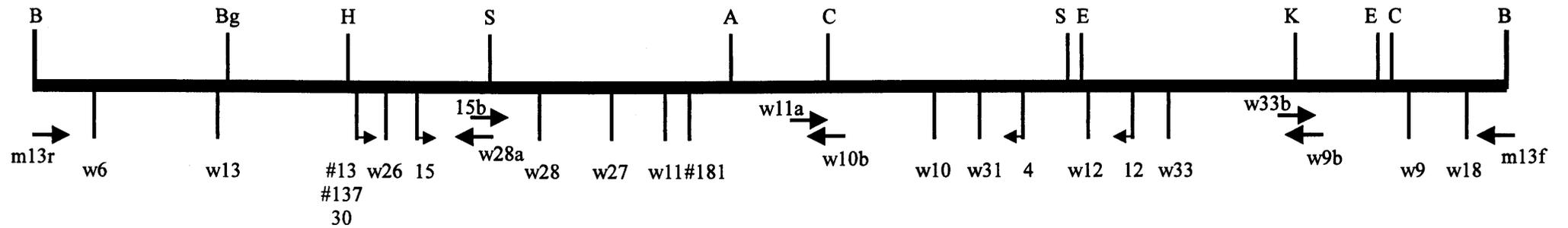
(Restriction enzyme sites: A-*AccI*, B-*BamHI*, Bg-*BglII*, C-*Clal*, E-*EcoRV*, H-*HindIII*, K-*KpnI*, P-*PstI*, S-*SphI*, X-*XbaI*)

pRDH7



1 kb

pRDH8



bfpA coding region and was predicted to form a PhoA fusion to the first 116 amino acids of the mature BfpA protein.

4.2.1.4 DNA sequence analysis of *phoA*-positive *TnphoA* insertions in pRDH8.

Analysis of the DNA sequence data using BESTFIT and GAP allowed the co-localisation of the 4 *phoA* positive *TnphoA* insertions in two pairs; pRDH8-30 was found to be 250 bp upstream of pRDH8-15 (with respect to *phoA* orientation), and at a separate locus pRDH8-4 was 460 bp downstream (with respect to *phoA* orientation) of pRDH8-12 (**Figure 4.1**). The transposon insertion site of pRDH8-30 was found to be identical to those of the original mutants #13 and #137. A BLAST comparison search of the GenBank database using the DNA sequence determined between pRDH8-30 and pRDH8-15 (i.e. the region downstream of the *TnphoA* insertion in mutants #13/#137) showed a high degree of homology (BLAST algorithm E value $< 10^{-132}$) with the *E. coli* K-12 sequence *umudca* (GenBank accession number M29542). This 363 bp fragment, located approximately 10 kb downstream of the *E. coli* genes *umuD* and *umuC* (mapped to 25.5 min; Bachmann, 1983) had been sequenced from laboratory *E. coli* strain AB1157 (van de Putte *et al.*, 1984; Sedgewick *et al.*, 1988). Sedgewick *et al.* (1988) had identified a 12.5 kb region which displayed high restriction site conservation between *E. coli* and strains of *Escheirchia aurescens*, *Escherichia dispar* and *Escherichia alkalescens*. Sequencing the boundaries of this region (i.e. upstream of *umuDC* and within the *umudca* sequence), Sedgewick *et al.* (1988) identified potential Tn3-like transposon insertion sequences (**Figure 4.2**) which they proposed might form the remains of an ancestral sequence insertion event into the *Escherichia* genome. The high degree of homology between the *umudca* sequence and those from pRDH8-30 PB1 and pRDH8-15 PA1 (92% identity; **Figure 4.2**) suggested that the pRDH8 *Bam*HI fragment was located at the equivalent site in the EPEC chromosome. However there were a number of insertional and deletion differences between the two DNA sequences such that the open reading frame identified in pRDH8 into which the *TnphoA* transposons in EPEC mutants #13 and #137 were inserted was completely disrupted in the sequence from *E. coli* K-12. It was impossible to determine if this was

due to sequencing errors (there was evidence of at least one 4 bp duplication in the umudca sequence; **Figure 4.2**) or due to an accumulation of mutations in the *E. coli* K-12 strain. BLAST searches of the GenBank database with the pRDH8-4 and pRDH8-12 DNA sequences failed to produce significant matches against *E. coli* chromosomal sequences.

4.2.1.5 Mapping of *TnphoA* insertions in pRDH7 and pRDH8. In order to complete the DNA sequencing of pRDH7 and pRDH8 it was important to determine the exact location of the *TnphoA* insertions already sequenced and identify further clones with transposon insertions which could be used to determine the DNA sequence separating the contigs already formed. The *phoA*-negative colonies identified in section 4.2.1.1 were chosen as the source of potential new transposon insertions for DNA sequencing (72 clones were selected and renamed; pRDH7-w1 through -w36, and pRDH8-w1 through -w36). The sites of *TnphoA* insertion in both these and the pRDH7 and pRDH8 plasmids carrying *phoA*-positive *TnphoA* insertions was achieved by a combination of restriction enzyme mapping and PCR (summarised in **Figure 4.3**). Plasmid DNA was prepared from each clone using the miniprep method, digested with *Bam*HI and the resultant products sized by agarose gel electrophoresis. Plasmids which failed to give a DNA fragment of 2.9 kb (the size of linearised pBluescript) when digested with *Bam*HI were immediately discarded; it was assumed that these clones carried *TnphoA* inserted in the pBluescript vector rather than the insert. The remaining plasmids were then digested with *Eco*RI and *Eco*RI + *Sst*I and the products analysed side by side on agarose gels (**Figure 4.3A**). *TnphoA* contains two *Eco*RI sites which are 0.7 kb and 1.1 kb respectively into the *phoA* terminus (**Figure 3.2**) and both pRDH7 and pRDH8 contain unique *Eco*RI and *Sst*I sites which are at opposite ends of the cloned *Bam*HI fragments within the divided pBluescript multiple cloning site (MCS) (**Figures 3.16 and 3.17**). By identifying and sizing the smallest DNA fragment released by each pair of digestions (ignoring the 2.9 kb pBluescript fragment released in all *Eco*RI/*Sst*I digestions) it was possible to deduce both the orientation of the transposon and the approximate distance of the insertion site relative to either the

EcoRI site or the *SstI* site in the MCS. These provisional enzyme-mapped locations were confirmed by PCR using combinations of PA1 or PB1 and m13f or m13r primers to amplify the smallest fragments predicted by the presumed orientation and location of the transposon (**Figure 4.3B**). A general PCR reaction was used successfully to amplify fragments up to 3 kb in size using a 1/1000 dilution of the miniprep DNA as template. The PCR program consisted of: one cycle of 95°C for 5 min, 55°C for 30 sec, and 72°C for 5 min; and 36 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 5 min. PCR products were analysed on 1% agarose gels and sizes estimated against DNA molecular weight standards. The PCR confirmed locations of the *phoA*-negative transposon insertions were compared with those of the *phoA*-positive *TnphoA* insertions already sequenced and plasmids were chosen for sequencing which would fill the gaps between current contigs (9 for pRDH7 and 13 for pRDH8); the positions of these transposon insertions are indicated in **Figure 4.1**.

4.2.1.6 The complete DNA sequence of pRDH7. Duplicate DNA sequences were determined for the 9 *phoA*-negative *TnphoA* insertion-carrying pRDH7 plasmids identified in section 4.2.1.3 using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit with primers PA1 and PB1. The sequences obtained were compared, with the contigs previously determined using the BESTFIT program and a consensus sequence was assembled for pRDH7 using the SEQED program (section 2.15). The complete *BamHI* fragment of pRDH7 was determined to be 5347 bp in length and the annotated sequence is shown in **Appendix 1**.

4.2.1.7 Completion of the DNA sequencing of pRDH8 using primer walking. Duplicate DNA sequences were determined using primers PA1 and PB1 for the 14 *phoA*-negative *TnphoA* insertion-carrying pRDH8 plasmids. These sequences were aligned and compiled using the BESTFIT and SEQED programs into 4 contigs covering in total approximately 5 kb. It was estimated, based upon the contigs obtained and the restriction enzyme map of pRDH8, that the 3 remaining gaps were each less than 500 bp in size; therefore it was determined to close them by primer walking. Six

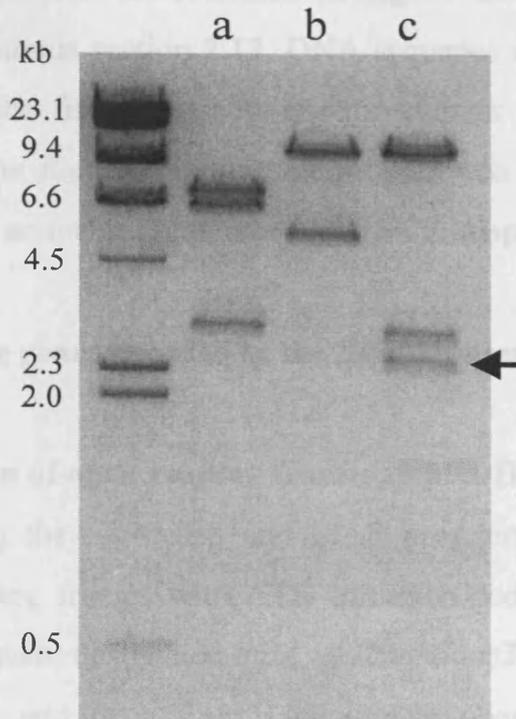
Figure 4.3 Summary of the strategy for mapping the sites of *TnphoA* insertions in plasmids pRDH7 and pRDH8.

(A) Restriction enzyme mapping of plasmid pRDH8-15. Plasmid pRDH8-15 DNA was digested with the following restriction enzymes a) *Bam*HI, b) *Eco*RI, c) *Eco*RI and *Sst*I, and the products separated by electrophoresis on a 1% agarose gel. The *Bam*HI digest released an intact 2.9 kb pBluescript fragment indicating that the *TnphoA* insertion is located within the cloned insert and not the vector. The smallest DNA product released in either the *Eco*RI or *Eco*RI and *Sst*I digestions was the 2.1 kb *Eco*RI-*Sst*I fragment (indicated by an arrow). This fragment represents the distance between the first *Eco*RI site in the 3' end of *TnphoA* and the *Sst*I site in pRDH8; by subtracting the 700 bp of *TnphoA* sequence included in this fragment it is possible therefore to calculate that the insertion site is approximately 1.4 kb within the *Bam*HI insert in pRDH8 and that the 3' (*phoA*) end of *TnphoA* is orientated towards the *Sst*I site (as shown in Figure 4.1B).

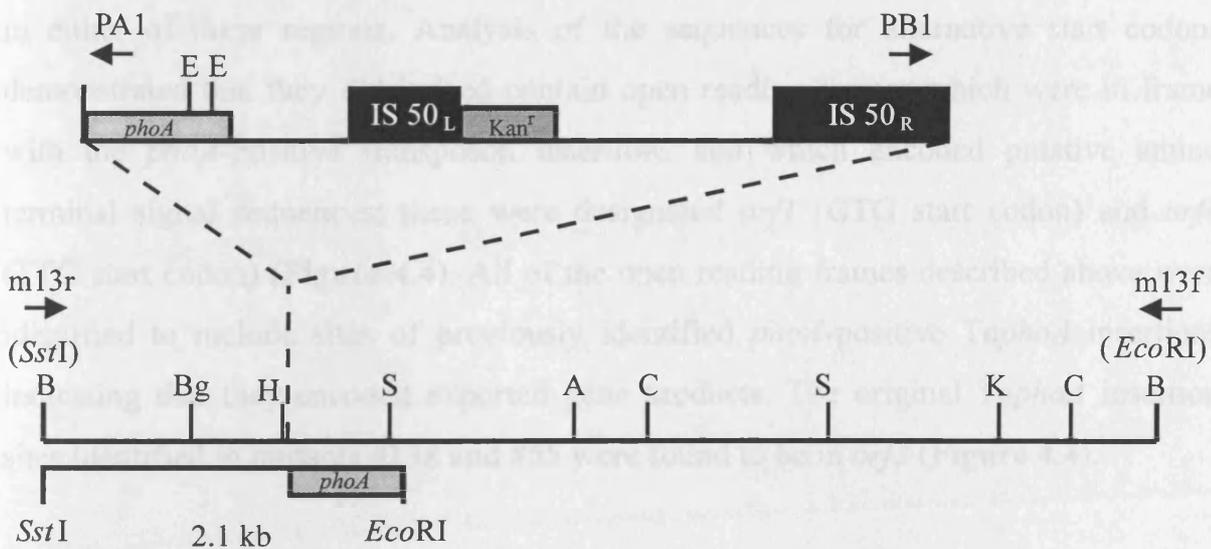
(B) Diagrammatic representation of the *TnphoA* insertion in plasmid pRDH8-15. The site of the *TnphoA* insertion was located within pRDH8-15 by restriction enzyme mapping as described above (the 2.1 kb *Sst*I-*Eco*RI fragment is indicated below the map). The position of the transposon insertion was confirmed by PCR using primers PA1 and m13r which produced a 1.4 kb fragment.

(A - *Acc*I, B - *Bam*HI, Bg - *Bg*III, C - *Cl*aI, E - *Eco*RI, H - *Hind*III, K - *Kpn*I, S - *Sph*I)

A



B



pRDH8-15

oligonucleotide primers were designed which were homologous to DNA sequences approximately 50 bp within the ends of the completed contigs; the positions and orientation of these primers are indicated on **Figure 4.1** and their sequences included in Materials and Methods section 2.12. DNA sequence data from these primers were sufficient to allow the final gaps between the contigs to be closed. The completed DNA sequence of the *Bam*HI fragment of pRDH8 was determined to be 6334 bp in length; the complete annotated sequence is shown in **Appendix 2**.

4.2.2 Analysis of the genes encoded in the *Bam*HI insert of pRDH7.

4.2.2.1 Identification of open reading frames in pRDH7. Initial analysis of the DNA sequence data using the FRAMES and MAP programs (section 2.15) identified 3 complete open reading frames with ATG initiation codons and consensus ribosome binding sites; these were designated *bfpA*, *orf2*, and *orf3*. FRAMES also indicated the existence of the 5' end of a fourth incomplete open reading frame which was designated *orf5'* (**Figure 4.4**). From the locations of the *phoA*-positive insertions determined in the *TnphoA* mapping of pRDH7 (section 4.2.1.5) it was apparent that there would also be open reading frames located between *bfpA* and *orf2*, and between *orf3* and *orf5'*; however there was no evidence of ATG start codons associated with good ribosome binding sites or obvious signal sequences for the open reading frames in either of these regions. Analysis of the sequences for alternative start codons demonstrated that they did indeed contain open reading frames which were in frame with the *phoA*-positive transposon insertions and which encoded putative amino terminal signal sequences; these were designated *orf1* (GTG start codon) and *orf4* (TTG start codon) (**Figure 4.4**). All of the open reading frames described above were identified to include sites of previously identified *phoA*-positive *TnphoA* insertions indicating that they encoded exported gene products. The original *TnphoA* insertion sites identified in mutants #138 and #55 were found to be in *orf2* (**Figure 4.4**).

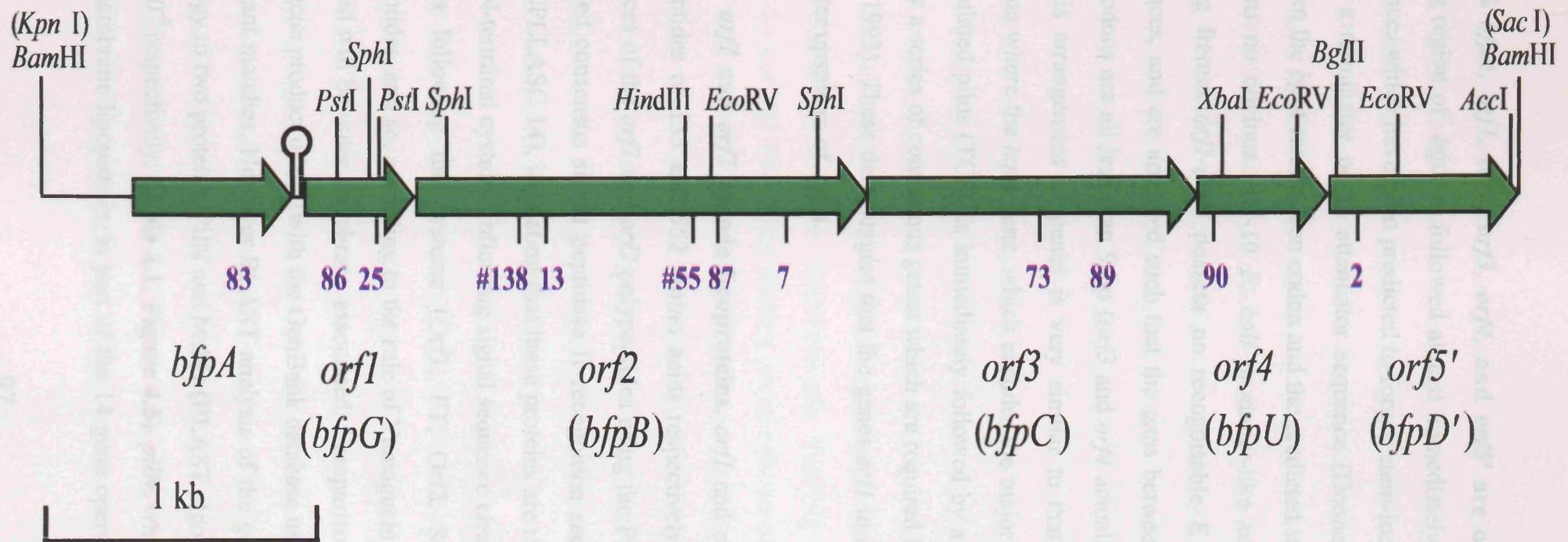


Figure 4.4 Diagrammatic representation of the open reading frames identified in plasmid pRDH7. The 5.3 kb *Bam*HI fragment of pRDH7 is depicted in the same orientation in which the DNA sequence is presented in Appendix 1 (the *Kpn*I and *Sac*I sites indicate the orientation with respect to the pBluescript KS sequence). The predicted open reading frames are represented as green arrows and the sites of *phoA*-positive *TnphoA* insertions are indicated in blue. The stem-loop predicted between *bfpA* and *orf2* (*bfpG*) is indicated.

4.2.2.2 *bfpA*, *orf1*, *orf2*, *orf3*, *orf4*, and *orf5'* are organised as an operon. The coding region of *bfpA* is followed almost immediately by two 11 bp inverted repeat sequences which have been predicted to form a stem-loop structure which might act as either a terminator or an attenuator sequence (Donnenberg *et al.*, 1992). The gap between the *bfpA* termination codon and the predicted start codon of *orf1* is 45 bp and contains no obvious -35/-10 *E. coli* promoter-like sequences. Similarly the open reading frames *orf1-orf5'* possess no recognisable *E. coli* promoter or terminator sequences, and are arranged such that the gaps between termination and subsequent start codons are all less than 9 bp (*orf3* and *orf4* actually overlap by 4 bp; **Appendix 1**). This arrangement of genes is very similar to that of the *tcp* operon of *Vibrio cholerae* where the *tcpA* gene, which encodes the major structural subunit of the toxin co-regulated pilus (TCP), is immediately followed by a predicted attenuator loop and then by a series of contiguous genes which are required for TCP biogenesis (Ogierman *et al.*, 1993). These data suggest that the genes *orf1* to *orf5* are co-transcribed from a promoter upstream of *bfpA*.

4.2.2.3 *orf1* and *orf2* encode lipoproteins. *orf1* and *orf2* were predicted to encode polypeptides of 133 and 552 amino acids respectively. Analysis of the N-terminal sequences of the *orf1* and *orf2* polypeptides using the PSORT program (section 2.15) identified consensus signal peptidase II recognition sequences (VILFFCLLVSC 18; LPLLCPLLASC 14), indicating that these proteins are modified by addition of a lipid to the N-terminal cysteine following signal sequence cleavage (von Heijne, 1986). The residues following the cysteine (*Orf1*, FT; *Orf2*, SG) are uncharged for both polypeptides, and so, according to the rule of Yamaguchi *et al.* (1988), the lipoproteins produced will be outer membrane associated. Comparison of the amino acid sequence of the gene product of *orf1* with the GenBank database using BLAST failed to give any significant matches. However BLAST analysis of the gene product of *orf2* identified homology to two proteins, PilN and MshD (BLAST algorithm E values $< 2 \times 10^{-21}$ and $< 7 \times 10^{-8}$ respectively; **Table 4.1; Figure 4.5**). *pilN*, which encodes a 560 amino acid outer membrane lipoprotein, is part of the 14 gene operon required for the biogenesis

of the thin pilus from the *E. coli* plasmid R64, recently identified as a member of the type IV pilus family (Kim and Komano, 1997). *mshD* also encodes an outer membrane lipoprotein, of 559 amino acids, which is required for assembly of the *Vibrio cholerae* type IV pili-like mannose-sensitive hemagglutinin (MSHA) (Hase *et al.*, 1994). Orf2 contains a region of poly-serine between amino acids 220 and 250 which is partially conserved in both MshD and PilN (Figure 4.5); the function of this region is unknown however secondary structure predictions (Chou and Fasman, 1978) indicate a high β -turn probability which might suggest a role in the physical separation of functionally different N-terminal and C-terminal domains.

4.2.2.4 *orf3* and *orf4*. The open reading frame of *orf3* is predicted to encode a 402 amino acid polypeptide of 45.3 kDa. Analysis of the Orf3 polypeptide with program PSORT identified a non-cleavable N-terminal signal sequence indicating that it was an integral inner membrane protein; this was confirmed by program TMPRED (section 2.15) which predicted two transmembrane spanning segments between amino acids 170-186 and 367-383. Open reading frame *orf4* encodes a polypeptide of 156 amino acids with a consensus signal peptidase I site predicted to cleave between amino acids 18 and 19 (VNA EV) which would produce a mature protein of 17.7 kDa. PSORT predicts Orf4 to be a periplasmic protein possibly associated with the outer membrane. Comparison of the polypeptides encoded by *orf3* and *orf4* with the GenBank database failed to identify any significant matches.

4.2.2.5 *orf5'* encodes a member of the PulE/OutE/ExeE/XpsE/XcpR family of transport proteins. The partial open reading *orf5'* encodes a polypeptide of 231 amino acids which by analysis with program PSORT appears not to possess a consensus *sec*-dependent signal sequence nor any hydrophobic transmembrane regions, and is therefore predicted to be cytoplasmic. However it had already been demonstrated that introduction of plasmid pRDH7-2, which carries a *TnphoA* insertion in-frame with *orf5'*, into DH5 α resulted in colonies which were strongly blue on LUA medium containing XP indicating alkaline phosphatase activity (pRDH7-2 is predicted to form

a chimeric polypeptide consisting of the first 34 amino acids of Orf5' fused to PhoA; **Figure 4.4; Appendix 1**). It remains unclear whether the gene product of *orf5* is exported or if the *phoA*-positive nature of the chimeric polypeptide formed by the insertion in pRDH7-2 is indeed an artefact; PSORT predictions for the chimeric polypeptide suggest that it would be localised to the cytoplasm. Comparison of the Orf5' polypeptide sequence with the GenBank database using BLAST identified a single significant match (BLAST algorithm E value < 0.81) to the N-terminus of the *Vibrio cholerae tcpT* gene product (28.4% identity over 230 amino acids; **Table 4.1; Figure 4.6**). *tcpT* encodes a 503 amino acid ATP-binding protein, predicted to be localised to the cytoplasm, which is required for production of the *Vibrio cholerae* toxin co-regulated pilus (TCP) (Oigerman *et al.*, 1993). TcpT has been identified to belong to a family of ATP-binding proteins which are involved in secretion of proteins via the general secretion pathway in Gram-negative bacterial species (Oigerman *et al.*, 1993; Pugsley, 1993). The family includes PuleE, the pullulanase export protein of *Klebsiella oxytoca*, OutE of *Erwinia chrysanthemi*, ExeE of *Aeromonas hydrophila*, XpsE of *Xanthomonas campestris* and XcpR of *Pseudomonas aeruginosa* (Posset *et al.*, 1992; Lindberg and Collmer, 1992; Jiang and Howard, 1992; Dums *et al.*, 1991; Bally *et al.*, 1992). A consensus nucleotide binding motif (Walker box) has been identified for the family which is present at residues 236-246 of the TcpT sequence (GXTGSGKXXXL; Whitchurch *et al.*, 1991). The amino acid sequence of Orf5' at the equivalent point, as indicated in the alignment in **Figure 4.6**, has only limited homology to this motif. However, as the Orf5' polypeptide sequence is incomplete and terminates immediately after this point, it is possible that the current alignment is incorrect and that an ATP-binding site may exist in the unsequenced region.

4.2.2.6 Comparison of pRDH7 sequences to the *bfp* operon. Shortly after the sequencing and analysis of the pRDH7 *Bam*HI insert was completed the entire sequence of the *bfp* operon was reported separately by two groups for the EPEC strains E2348-69 (O127:H6) (Stone *et al.*, 1996) and B171-8 (O111:NM) (Sohel *et al.*, 1996). The sequence of pRDH7 and the open reading frames identified above agree largely

Gene product	Size (Amino acids)	MW (Da)	Predicted location ^a	pI	Homologue(s) ^b		Identity ^c (Similarity)	GenBank Accession Number
					Protein	Bacterial species		
BfpA (Mature form)	193 (180)	20,269.8 (18,731.9)	OM/P	8.42 (5.30)	TcpA	<i>V. cholerae</i>	26.9 (33.9)	X74730
Orf1 (BfpG) (Mature form)	133 (120)	14,791.8 (13,267.9)	OM (Lipoprotein)	7.79 (7.02)	None			
Orf2 (BfpB) (Mature form)	552 (535)	58,423.3 (56,565.9)	OM (Lipoprotein)	5.59 (5.39)	PilN MshD	<i>E. coli</i> <i>V. cholerae</i>	24.3 (31.3) 21.4 (28.4)	D88588 L19085
Orf3 (BfpC)	402	45,301.3	IM	5.49	None			
Orf4 (BfpU) (Mature form)	156 (138)	17,710.3 (15,684.8)	P/OM	6.60 (5.51)	None			
Orf5' (BfpD)	231	N/A	C	N/A	TcpT	<i>V. cholerae</i>	28.4 (39.6)	X64098

Table 4.1 Physical characteristics of the gene products of pRDH7. ^a subcellular location was predicted using the program PSORT (C, cytoplasm, IM, inner membrane, P periplasm, OM, outer membrane). ^b homologues were identified using the gapped-BLAST program (Atschul *et al.*, 1997). ^c similarity to homologues was calculated using the Wisconsin GCG program BESTFIT using a gap penalty of 2. N/A - data not available.

```

                20                40                60
MshD : -----MRKIVASVVTSVGGSMG-----HRDPVFKQANQAIETNSQIQIIPPSVEADLM : 54
PilN : MKKSHQSRMKLAVPPIAMIALSISCTFSEINKMQKKAQDSAHAREKVSALSAAKQKQATWLNQWLN : 70
Orf2 : ----MKLGRYSPLPLPLASISGNFYKDNLGVINDILHSDTSLKSKEHYKSSLVSKTUSIYIG : 66

                80                100                120                140
MshD : DMDTLTASERK-----TLQRFIIQA-----EDVEKAFFASLVQGTEYSAAHPAVTGRIILN : 109
PilN : VVPVAQVSRFKK--QTAPACYTQARKGEITLQGLGQRTEVCTEVIITPDAANSTLECGARQMIGT : 138
Orf2 : NSSFQTYHGEELPGKLEGVHGILSSSTPLGFDEVLSMQDSSGIVK--HTTKDVTSCVSSKSLAAT : 134

                160                180                200
MshD : TDVTLDEALG---VVRLYGFVVKGKVIQVYPAF----LRTTIPVDYLQFKRTGRSLTIIITGTI : 170
PilN : PDPDEGRIP--LSSLGSTMTTSTQPLTINNLMWGGDINGLLELMRSRSGTARMDCRIVYLETERT : 206
Orf2 : VAEKMSSTGKSTDFHLLLEVSSHQMDNYGALSTFLKVAANYNNTYESGRVATNEETKR : 204

                220                240                260                280
MshD : TNTDTNSNSSSSSSSSSSSSSSSSDGSSSN--SNSNRSDARGTETIETTNSDVPLLEKAVAQLLGSG : 238
PilN : YPLHMN--TKTSSSSSVSSSGSTSTMATSGQDNSASGDATSSQSTTVGQYDLYEDIRKTEAMLTPEK : 274
Orf2 : FSISLPGGKTYTSKNSLSSDSNSSSSSSG---SSGSSPSDEALKFDSDVFAKDIENSKLILGSD- : 269

                300                320                340
MshD : GQTVIVNPQAGVPLAYDEIRQVFLGISQQRHROVISEAKVEVLSGYQVINSKAFS--- : 305
PilN : -RWLASSSTLITVDTAVQEAARVDEQSIINROVALNQVSVSNRNEFLDNLVYSLHS : 343
Orf2 : -SSISSTTSVIVTSSPNMKKININTLAQLERQVTIDAYNVVTSSSDLAMSLEALLK---H : 335

                360                380                400                420
MshD : NGANYKIGSGSITQDSNENPITSVPGDAIENLLGGQSNVVISSGFDAVSFMTQGLNVLSSPRVT : 375
PilN : AGATLNNASGDFTGATSAG--VSIIDTATLAAKF-----SGSSLKAKSAGDVSVVTSQEST : 401
Orf2 : NGGVLGSVSTSNFAATSGT--PSFTRYNGNCD-----ANQVLLNLAKKGVSVVTSASVT : 391

                440                460                480
MshD : ASNNOKAVIKVGTDEYYVTDLSSVVGTDNAQASPDILTLPFSGISLSDVTPQIDQCNVVTHVHPAVI : 445
PilN : VTLTPVPIQADQTVYVAQSAITTTT--DVGATITLPGMITGNNMTLPLIQKTGNLQLQOMNFNSD : 469
Orf2 : TMSGQRPVPLKVGNDRTYVVEIGTVLSQ--SS-TSTASTSTVSGALNLLPVAIDCNILLYGVTLSE : 458

                500                520                540                560
MshD : VEQQTKKILYRSEE---IPLARSISRESDEVIKAKDQVVIIGGLMSVVDQSKVPLGDDVALGH : 512
PilN : PP-TIRSSTSKDGN-SYIEMPYTKLRSLSKVNEKE--GQSVVTCFDQNTT-TSKATTPANPFGG : 534
Orf2 : LVGSNNGDQATVIGTVIQPNVDSTTFVDSMENN--GNTLLAGYEKRNE-SDQVGTSTFKLLG : 525

                580                600
MshD : LFRNTKLTQKEEVILLKPIVVGVTWQKELERSRSLLEQWFPDSQ : 559
PilN : SQ---KNESTVILIPFPSSGGNN----- : 560
Orf2 : AL---NSASAVTICETRIIDLKASSE----- : 552

```

Figure 4.5 Amino acid alignment of pRDH7 Orf2 (BfpB) with MshD of *Vibrio cholerae* and PilN of *E. coli*. The deduced protein product of *orf2* (*bfpB*) of pRDH7 shows 24.3% identity (31.3% similarity) to PilN and 21.4% identity (28.4% similarity) to MshD (see also Table 4.1). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues which are identical in two of the polypeptides are boxed in grey, those identical in all three are boxed in black. The region between 220 and 250 residues contains a conserved poly-serine motif which may play a role in the physical separation of N-terminal and C-terminal domains.

```

                                20                                40
TcpT : MSIDIKYLSRIDREEFFFKDSRLCKKFDEEREVLTLLLEFDTKERVNL : 50
Orf5 : --KLTKTEKTSDIMFERFKRVSEIVTGDGGE----LELTVEQRKYFLIF : 44

                                60                                80                                100
TcpT : LKKDKVYKYFLVSDANHKLLIANVTEQQA DLSIEKDIMKASSATAY : 100
Orf5 : KNGD-----FLVSSCHMKHHLVQMLREIATRKG--PNLTIYEVN----- : 82

                                120                                140
TcpT : GASDIHFREDRICKKFRVNGTMIDYRELSSEADSLMFVLYNVMATTK : 150
Orf5 : -LKDIRLLYEASLKTQON--NG-----QDLPVEKRASMLFECAEMR-- : 122

                                160                                180                                200
TcpT : ETTWNRKLPQDANTIIVINFKAYRFR-YAHMPLFGGGKN-YHAVVRIITY : 198
Orf5 : -----VSDLHIKQYDAEADIYIRKDGDMELLROIESNTAHSILASLY : 164

                                220                                240
TcpT : PSNMFVCTNYQDIGYNEADTDAIAR---ILNTSYGFTVSGTTGSGKSTS : 245
Orf5 : NNAQDSDATYKINAYQAARIVASKSRLALPPVIQAVRQFNPLGQGGRYL : 214

                                260
TcpT : EKKYIELLFFNKYKG-- : 260
Orf5 : LARELYTDKSEKQKEMD : 231

```

Figure 4.6 Amino acid alignment of pRDH7 Orf5 (BfpD) with TcpT of *Vibrio cholerae*. The deduced protein product of the incomplete open reading frame *orf5'* of pRDH7 (*bfpD*) showed 28.4% identity (39.6% similarity) to TcpT (see also Table 4.1). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues that are identical in both sequences are boxed in black, residues which share similar chemical characteristics are boxed in grey.

with the sequence reported by Stone *et al.* (1996), and so the nomenclature agreed between the laboratories of Gary K. Schoolnik (Stanford University) and Michael S. Donnenberg (University of Maryland) for the *bfp* operon has been incorporated in this work alongside my own; that is *orf1* \equiv *bfpG*, *orf2* \equiv *bfpB*, *orf3* \equiv *bfpC*, *orf4* \equiv *bfpU*, and *orf5* \equiv *bfpD*.

The only major differences in the pRDH7 DNA sequence were in the *bfpB* gene (*orf2*); there are five single basepair errors in the pRDH7 *orf2* sequence which result in amino acid changes (2 of which are conserved changes) compared to the *bfpB* sequences of Stone *et al.* (1996) or Soheli *et al.* (1996). However the pRDH7 *orf2* sequence also identifies errors in both of the previous sequences; Soheli *et al.* (1996) report two GC pairs apparently inverted resulting in residues 273-276 of their BfpB sequence being different (it is impossible to determine if they are correct due to the strain difference), whilst Stone *et al.* (1996) include an additional 3 bp resulting in a frameshift and the addition of an extra codon such that residues 185-189 of their BfpB sequence are incorrect, they also show a premature termination codon and therefore lose the terminal glutamate residue. In each case mentioned above the pRDH7 *orf2* sequence agrees exactly with nucleotide sequence from the other laboratory.

4.2.3 Analysis of the genes encoded by the *Bam*HI insert of pRDH8.

4.2.3.1 Identification of open reading frames in pRDH8. The DNA sequence of the pRDH8 *Bam*HI insert was analysed using the FRAMES and MAP programs and five complete open reading frames were identified, now designated *orf2*, *orf4*, *emtA*, *tag* and *ompX*; the 3' end of a sixth incomplete open reading frame, now designated *orf1'*, was also identified (**Figure. 4.7**). The open reading frames *emtA* and *tag* contained the *TnphoA* insertion sites identified in the original mutants #13/#137 and #181 respectively; *emtA* also contained the insertion sites of two of the *phoA* positive *TnphoA* insertion mutants identified in section 4.2.1.4 (pRDH8-15 and pRDH-30). After the initial analysis it was noted that there was a 950 bp region separating the

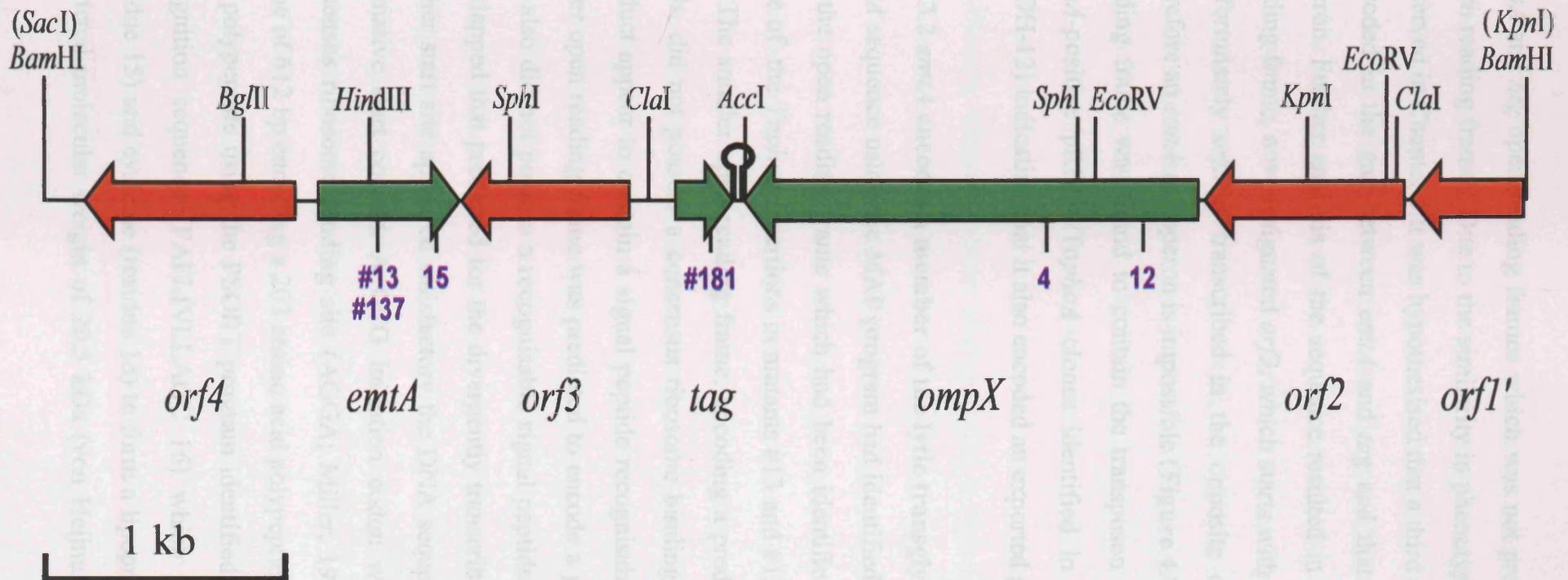


Figure 4.7 Diagrammatic representation of the open reading frames identified in pRDH8. The 6.4 kb *Bam*HI fragment of pRDH8 is depicted in the same orientation in which the sequence is presented in Appendix 2 (the *Sac*I and *Kpn*I sites indicate the orientation with respect to pBluescript KS). The open reading frames which are predicted to encode exported proteins i.e. those with *phoA*-positive *TnphoA* insertions (indicated in blue), are coloured in green whilst those thought to be cytoplasmic are coloured red. The Rho-dependent terminator which is predicted to be between *tag* and *ompX* is indicated by a stem-loop structure.

emtA and *tag* open reading frames which was not predicted to encode any significant open reading frames. Due to the similarity in phenotype of mutants #13/#137 and #181 observed in Chapter 3 it was hypothesised that a third as yet unidentified gene could be encoded in the gap between *emtA* and *tag* and that the three genes might form an operon. Further analysis of the sequence resulted in identification of a 735 bp open reading frame, now designated *orf3*, which starts with the alternative start codon GTG. Unfortunately *orf3* is transcribed in the opposite direction to its neighbours and therefore an *emtA-tag* operon is impossible (Figure 4.7; Appendix 2). The *ompX* open reading frame was found to contain the transposon insertion sites of the two other *phoA*-positive pRDH8/*TnphoA* clones identified in section 4.2.1.4 (pRDH8-4 and pRDH-12) indicating that it also encoded an exported gene product.

4.2.3.2 *emtA* encodes a member of the lytic transglycosylase family. Analysis of the *emtA* sequence using the MAP program had identified two potential ATG start codons for the open reading frame which had been identified to be in-frame with the *phoA* gene of the *TnphoA* insertions in mutants #13 and #137, and pRDH8-15 and pRDH8-30. The smaller open reading frame, encoding a predicted polypeptide of 167 amino acids, did not possess a consensus ribosome binding site nor did its predicted gene product appear to contain a signal peptide recognisable by the PSORT program. The larger open reading frame was predicted to encode a polypeptide of 241 amino acids; this also did not possess a recognisable signal peptide and furthermore its start codon overlapped that predicted for the divergently transcribed open reading frame *orf4*. As neither start site appeared satisfactory the DNA sequence was reanalysed looking for alternative start codons. A GTG initiation codon was identified downstream of a consensus ribosome binding site (AGGA; Miller, 1992) that gave an open reading frame of 612 bp encoding a 203 amino acid polypeptide. Analysis of the N-terminus of this polypeptide using the PSORT program identified a consensus signal peptidase II recognition sequence (FAFLIVLLAGC 16) which would cleave between glycine (residue 15) and cysteine (residue 16) to form a lipoprotein of 188 amino acids with a predicted molecular weight of 20.5 kDa (von Heijne, 1986; Table 4.2; Figure 4.8).

The two amino acids following the cysteine were both the neutral residue serine and so it was predicted that the mature lipoprotein would be localised to the outer membrane (Yamaguchi *et al.*, 1988). When analysed against the GenBank database using BLAST the *emtA* polypeptide showed significant similarity (BLAST algorithm, E value < 10^{-30}) over its entire length with two proteins, MltC from *E. coli* and y761, a protein predicted by the *Haemophilus influenzae* genome project (Dijkstra and Keck, 1996a; Fleischman *et al.*, 1995; **Figure 4.8**). MltC belongs to the *E. coli* lytic transglycosylase family of proteins which are involved in the metabolism of the bacterial peptidoglycan cell wall (Dijkstra and Keck, 1996a). Like lysozyme, the lytic transglycosylases cleave the β -1,4-glycosidic bond between the *N*-acetylmuramic acid and *N*-acetylglucosamine of peptidoglycan; however, due to an intramolecular transglycosylase reaction following cleavage, their final product is 1,6-anhydromuramic acid (Höltje and Tuomanen, 1991; **Figure 4.9**). Four lytic transglycosylases have been characterised in *E. coli*; three are lipoproteins, MltA, MltB, and MltC, and have been shown to be associated with the outer membrane (Ursinus and Höltje, 1994; Ehlert *et al.*, 1995; Dijkstra *et al.*, 1995; Dijkstra and Keck, 1996a), whilst the fourth and best characterised is the soluble lytic transglycosylase, Slt70, which has been shown to be associated with the outer surface of the cell wall (Höltje *et al.*, 1975). The Slt70 protein has been crystallised and its 3-dimensional structure elucidated (Rozenboom *et al.*, 1990; Thunnissen *et al.*, 1994). The crystallography data obtained allowed the identification of 3 amino acid motifs which are predicted to be in the active site of the enzyme and which have been identified in all of the known *E. coli* lytic transglycosylases (Dijkstra and Thunnissen, 1994; Dijkstra and Keck, 1996a). These motifs (the numbers indicate the site in the Slt70 protein sequence); glu478-ser479 (ES), gly493-leu494-met495-gln496 (GLMQ), and ala551-tyr552-asn553-x554-gly555 (AYNXG) were found to be highly conserved and appropriately spaced in the sequence of EmtA when compared to MltC (**Figure 4.8**).

Confirmation of the lytic transglycosylase nature of EmtA was provided by the recent identification and characterisation of the *emtA* gene from *E. coli* K-12; the deduced

protein sequence of this gene is 99% identical with that of EPEC EmtA (Joachim-Volker Höltje, personal communication). Kraft *et al.* (1998) overexpressed and purified the EmtA protein, which had an apparent molecular mass of 22 kDa. The enzyme demonstrated an endo-specific transglycosylase activity on isolated murein glycan chains unlike the exo-specific activities of the previously characterised *E. coli* lytic transglycosylases. Because of this enzymatic nature and its characterisation as a lipoprotein (determined by labelling with ^3H -palmitate; Kraft *et al.*, 1988) it was proposed to designate the protein EmtA, for endo-specific membrane-bound lytic transglycosylase.

4.2.3.3 The open reading frame *tag* encodes an integral inner membrane protein.

The *TnphoA* insertion identified in mutant #181 was located within a 255 bp open reading frame; due to its co-localisation with the *emtA* gene and the similarity of the phenotypes between mutant #181 and the *emtA* mutants #13 and #137 this open reading frame was tentatively designated *tag* for transglycosylase associated gene. *tag* is predicted to encode an 84 amino acid polypeptide with a molecular weight of 8.5 kDa; analysis of the polypeptide using program PSORT predicted an uncleavable *sec*-dependant signal sequence indicating that it would be an integral inner membrane protein (Pugsley, 1993). Separate analysis with program TMPRED indicated three transmembrane spanning regions and **Figure 4.10A** shows the most energetically likely of the two orientations proposed for the Tag polypeptide in the inner membrane. This orientation is further confirmed by the location of the Tag/PhoA fusion site of the *phoA*-positive *TnphoA* insertion in mutant #181 within a predicted periplasmic loop (**Figure 4.10A**). Comparison of the amino acid sequence of the Tag polypeptide with the GenBank database using BLAST identified a single high homology match (BLAST algorithm E value $< 1 \times 10^{-21}$) to the deduced product of a predicted open reading frame *f82*, identified by the *E. coli* genome sequencing project (**Table 4.2**; Blattner *et al.*, 1997). The polypeptide product of *f82* showed 67% identity (89% similarity) with Tag (**Figure 4.10B**); analysis of *f82* with PSORT and TMPRED predicted that it

Gene Product	Size (Amino acids)	MW (Da)	Predicted location ^a	pI	Homologue(s) ^b		Identity ^c (Similarity)	Genbank Accession Number
					Protein	Bacterial species		
Orf1'	159	N/A	N/A	N/A	yc73	<i>H. influenzae</i>	35.8 (61.6)	U32807
Orf2	285	30,867.7	C	8.48	ModD ModD	<i>H. influenzae</i> <i>R. capsulatus</i>	45.8 (62.9) 36.5 (56.1)	U32825 L06254
OmpX (Mature form)	656 (634)	73,585.4 (71,090.4)	OM	5.76 5.47	OprC	<i>P. aeruginosa</i>	27.1 (34.5)	D28119
Tag	84	8,652.5	IM	7.98	f82	<i>E. coli</i>	67.1 (89.0)	AE000274
Orf3	244	27,852.1	C/IM	8.99	None			
EmtA (Mature form)	203 (188)	22,256.5 (20,496.3)	OM (Lipoprotein)	9.16 8.65	MltC y761	<i>E. coli</i> <i>H. influenzae</i>	37.3 (57.2) 35.5 (55.7)	AE000379 U32760
Orf4	304	33,567.4	C	6.10	MccF	<i>E. coli</i>	31.3 (53.3)	X57583

Table 4.2 Physical characteristics of the gene products of pRDH8. ^a subcellular locations were predicted using the program PSORT (C, cytoplasm, IM, inner membrane, OM, outer membrane). ^b homologues were identified using gapped-BLAST at NCBI (Atschul *et al.*, 1997). ^c similarity to homologues was calculated using the Wisconsin GCG program BESTFIT using a gap penalty of 2.

Figure 4.8 Amino acid alignment of EmtA with MltC of *E. coli* and y761 of *Haemophilus influenzae*. The deduced protein sequence of the *emtA* gene shares 37.3% identity (57.2% similarity) with MltC and 35.5% identity (55.7% similarity) with y761 (see also Table 4.2). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues which are identical in two of the polypeptides are boxed in grey, those identical in all three are boxed in black. The location of the three amino acid motifs predicted to be present in the active site of lytic transglycosylases (**ES, GLMQ, AYNXG**) based upon the crystal structure of the *E. coli* soluble lytic transglycosylase Slt70 (Thunnissen *et al.*, 1994; Dijkstra and Thunnissen, 1994) are each marked with an asterisk (*). The cysteine residue which marks the signal peptidase II cleavage and lipid attachment site is marked with a + sign.

```

+                               40                               60
MltC : MMKKYLALALAIAPLLISCSTTK-KEDTYNEAWVKDINGFDILMGQFAHNTENIWGFKEVV : 59
y761 : -MKKYLLALL--FLYACSNSSNQGINYDEAFKDTQGLDILTSGQFSHNIDRIWGVNELL : 58
EmtA : --MKLRWFAFLIVLLAGCSSKH---DYTNPPYNAKVP----- : 32

```

```

80                               100                               120
MltC : IAGPKDYVKYTDQYQFRSHINFDDGTTITETIAGTEPA AHLRRATIKTLLMGDEPSSVDL : 119
y761 : VASRKDYVKYTD SFYFRSHVSVFDEGNIVLETQODLN---RLHNAIVHTLLMGADAKGIDL : 115
EmtA : ----- : -

```

```

140                               160                               180
MltC : YSDVDITIKKFPFLYGVVDNTLQPTRWEGRASNFADYLLKNRDKSAQQRAAYHLQRH- : 178
y761 : FASG-LVPSRSRPFLLGQVVDHQQCHIANQVIASNFATYLIQNKLQTRRLQNGHTVQFVS : 174
EmtA : ----- : -

```

```

200                               220                               240
MltC : -HMVPHLDKRAHKYIGMVFQASRKYGVDESLLILAIMQTESSEFNPYAVSRSDALGLMQVV : 237
y761 : VPMIANHVEVRARKYLELIRKAAQRYGIDESLILGIMQTESSEFNPYAISYANAIGLMQVV : 234
EmtA : --VQR-----AMQWPFISQKAGAAWGVDPQLITAI IAIESGGNPNVAVSKSNAI GLMQIK : 84
                **                               ****

```

```

260                               280                               300
MltC : QHTAGKDVERSQGKSGTFSRFLFD PASNIITGTAYLAMLNNVYLGGIDNPTSRRYAVIT : 297
y761 : PHTAGRDVFAAMKAGGQPS TRYLYDPANNILACVSYLWILQNYLDGITNPTSRRFAMIS : 294
EmtA : ASTSGRDVYFRMGWSGEP TTSSELKNPERNISMGAAYLNILETGPLAGIEDPKVLQYALVV : 144

```

```

320                               340                               360
MltC : AYNGGAGSVLRVFSNDKIQ ANIINTMTTEGDVYQTSITRHPSAESRRYLYKVNTAQKSYR : 357
y761 : AYNSGAGAVLRVFDNDKDTATYKINQMYEQVLRILTIVHPSSQARNYLLKVDKAQKKFR : 354
EmtA : SYANGAGALLRTFSSDRKKAITKINDLDSDEFLEHVARNH PAPQAPRYIYKLEQLDAM- : 203
*****

```

```

MltC : RE- : 359
y761 : VFR : 357
EmtA : --- : -

```

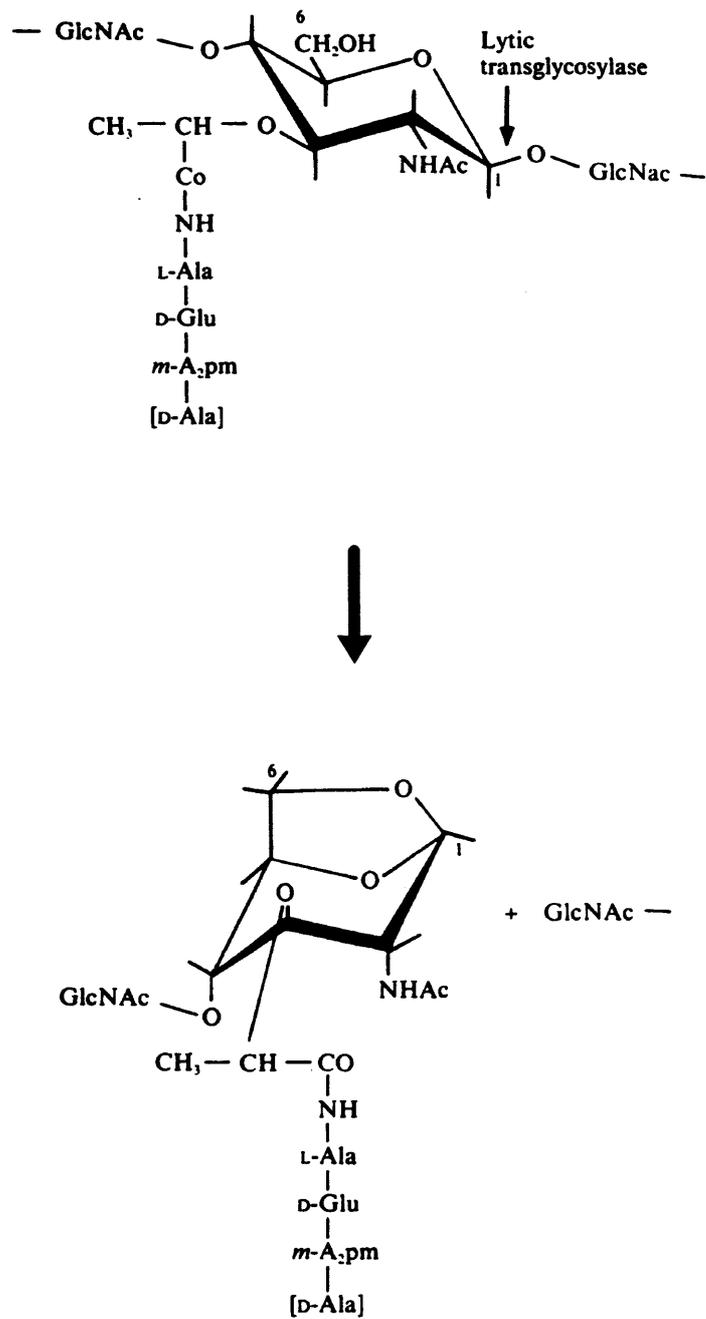
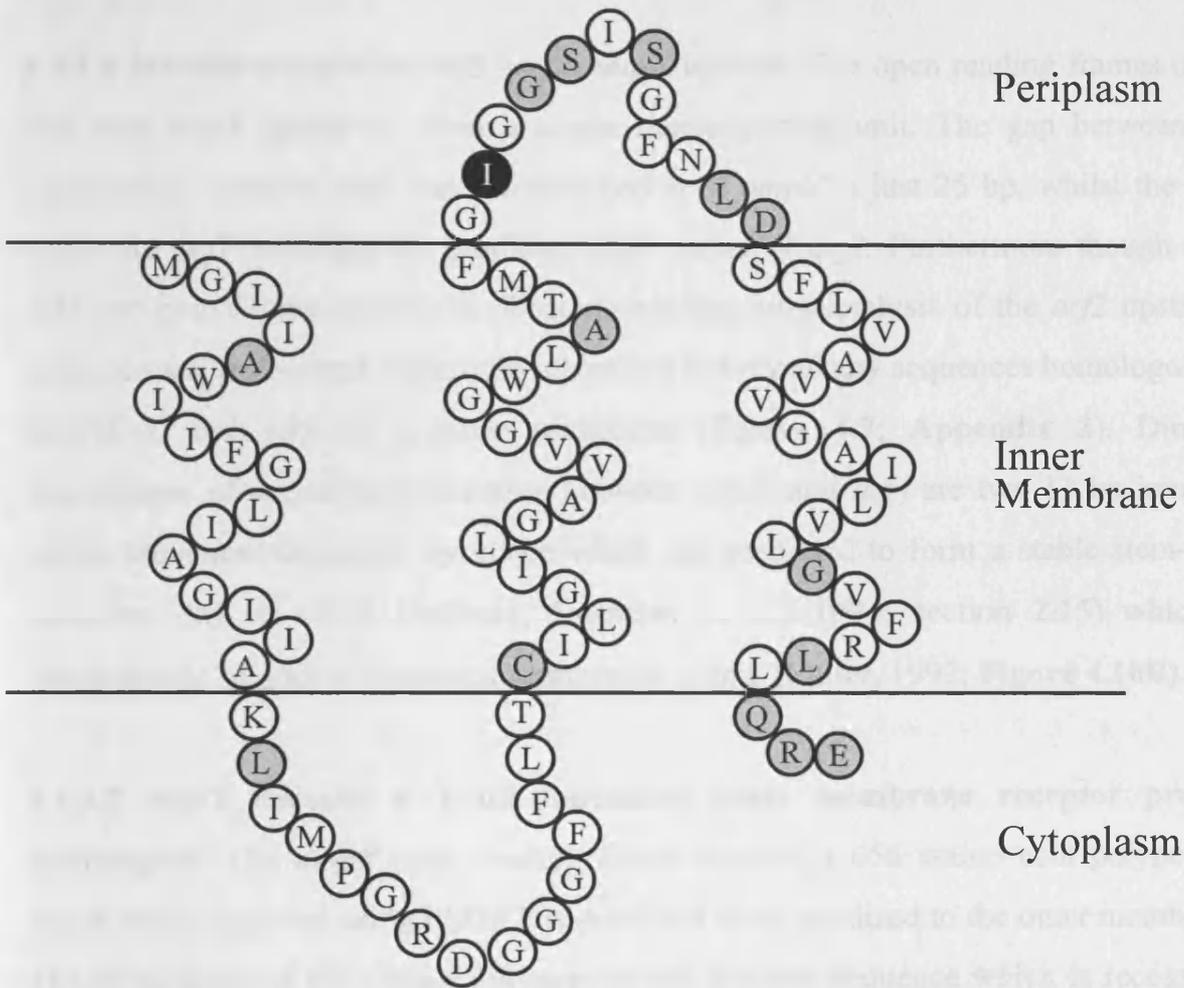


Figure 4.9 The reaction catalysed by the lytic transglycosylases of *E. coli*. (modified from Höltje and Tuomanen, 1991)

Figure 4.10A Predicted transmembrane regions of the Tag protein. Program TMPREDICT (ExPASy tools webpage; expasy.hcuge.ch/www/tools.htm) identified three potential transmembrane spanning regions in the deduced Tag polypeptide and predicted two possible orientations of the protein in the inner membrane; this figure represents the more energetically likely of the two orientations. All amino acid residues are identified by standard single letter code; residues highlighted in grey indicate those which are non-identical between Tag and f82 (see Figure 4.10B). The position of the Tag/PhoA fusion site identified for the *TnphoA* insertion in EPEC mutant #181 is indicated by the isoleucine residue highlighted in black. In the membrane orientation indicated the PhoA domain of the #181 Tag/PhoA fusion would be predicted to be in the periplasm, which would be in accordance with its known *phoA*-positive nature; in contrast the alternative orientation of Tag in the membrane would be predicted to have a cytoplasmically located PhoA domain which would be expected to have a *phoA*-negative phenotype.

Figure 4.10B Amino acid alignment of the Tag polypeptide with f82. The deduced protein product of the *tag* gene from pRDH8 shows 67.1% identity (89.0% similarity) to the deduced protein product of the predicted *E. coli* open reading frame f82 (Blattner *et al.*, 1997) (see also Table 4.2). The alignment was made using the CLUSTALW program of the Wisconsin GCG package; residues that are identical in both sequences are boxed in black whilst residues which share similar chemical characteristics are boxed in grey.

A



B

20 40 60

Tag : MGILAWIIFGLIAGIIAKLIMPGRDGGGGFFLTCLLGIVGAVVGGWLATMFGIGGSISGFN : 60

f82 : MGILSWIIFGLIAGILAKWIMPGRDGGGGFFMTILLGIVGAVVGGWISTLFGFG-KVDGFN : 59

80

Tag : LDSFLVAVVGAIVLVGVERLQRE : 84

f82 : FGSFVVAVIGAVVLFVYRKLK- : 82

encodes an integral inner membrane protein with three transmembrane domains analagous those of the Tag polypeptide.

4.2.3.4 Identification of an *orf1'-orf2-ompX* operon. The open reading frames *orf1'*, *orf2* and *ompX* appear to form a single transcriptional unit. The gap between the termination codon of *orf2* and the start codon of *ompX* is just 25 bp, whilst the stop codon for *orf1'* overlaps the predicted start codon of *orf2*. Furthermore though both *orf2* and *ompX* have consensus ribosome binding sites analysis of the *orf2* upstream sequence and *orf2-ompX* intergenic region did not reveal any sequences homologous to typical *E. coli* -35/-10 promoter sequences (**Figure 4.7; Appendix 2**). Directly downstream of *ompX* (and therefore between *ompX* and *tag*) are two 13 bp inverted repeat sequences separated by 10 bp which are predicted to form a stable stem-loop structure ($\Delta G = -19.29$ kcal/mol; Hofacker *et al.*, 1994; section 2.15) which is characteristic of a Rho-dependent termination signal (Miller, 1992; **Figure 4.16B**).

4.2.3.5 *ompX* encodes a TonB-dependent outer membrane receptor protein homologues. The *ompX* open reading frame encodes a 656 amino acid polypeptide which when analysed using PSORT is predicted to be localised to the outer membrane. The N-terminus of the OmpX polypeptide has a signal sequence which is recognised by signal peptidase I and which is predicted to be cleaved between residues 22 and 23 (AVA AQ); this would produce a mature protein of 634 amino acids, with a molecular weight of 71.1 kDa. Comparison of the OmpX protein sequence with the GenBank database using BLAST identified similarity with a large number of TonB-dependent outer membrane receptor proteins from a number of bacterial species. The highest homology observed was with a copper-regulated protein, OprC, from *Pseudomonas aeruginosa* (27.1% identity; Yoneyama and Nakae, 1996; **Table 4.3**). Comparison of the OmpX polypeptide with other known *E. coli* TonB-dependent outer membrane receptors also revealed regions of high homology (**Figure 4.11**); these included the "TonB box", present in all TonB-dependent receptors, and the Box II and Box III regions which have been demonstrated to show high homology between the TonB-

dependent receptors of several bacterial species (Bäumler and Hantke, 1992). OmpX also possesses the C-terminal consensus region and a terminal phenylalanine which have been proposed to be essential for proper incorporation of Gram-negative bacterial proteins into the outer membrane (Sturyve *et al.*, 1991). *E. coli* K-12 has nine characterised TonB-dependent receptor proteins, the majority of which are involved in the acquisition of iron through the uptake of ferrisiderophore complexes; these include, the ferric-enterobactin receptor (FepA), the ferrichrome receptor (FhuA), the iron dicitrate receptor (FecA), the ferric-coprogen receptor (FhuE), and in some strains the ferric-aerobactin receptor (IutA) (reviewed in Wooldridge and Williams, 1993). There are two further *E. coli* TonB-dependent receptors, Cir and Fiu, for which specific siderophores have not been identified; however they have been shown to be able to transport iron using the monocatecholates dihydroxybenzoic acid (DHBA; an enterobactin precursor) or dihydroxybenzoylserine (DBS; an enterobactin breakdown product) in a manner similar to FepA (Hantke, 1990). A TonB-dependent receptor for haem (ChuA) has also recently been isolated from *E. coli* O157:H7 strain EDL933 (Torres and Payne, 1997). The exception to the iron transport function is the TonB-dependent receptor BtuB which is required for the uptake of vitamin B-12 (Heller *et al.*, 1985). It has been reported that when aligned phylogenetically the TonB-dependent receptors of *E. coli* and *Yersinia* cluster according to their ligand preference (Rakin *et al.*, 1994, Bäumler and Hantke, 1992). Therefore the amino acid sequence of the mature OmpX polypeptide was aligned with the mature polypeptides of the nine known *E. coli* TonB-dependent receptors (plus a tenth, currently uncharacterised, TonB-dependant receptor-like polypeptide, f700, which was identified from the *E. coli* genome sequencing project data by sequence homology; **Figure 4.11**; Blattner *et al.*, 1997) using the CLUSTALW program (Higgins and Sharp, 1989). On the basis of this alignment a phylogenetic tree of the *E. coli* TonB-dependent receptor proteins including OmpX was constructed (**Figure 4.12**). Using the phylogenetic tree obtained it was impossible to assign OmpX clearly to any of the four ligand-based TonB-dependent receptor subfamilies (hydroxamate, citrate, catecholate or porphyrin) proposed by Rakin *et al.* (1994).

Protein	Bacterial species	Ligand	BLAST E value ^a	Identity ^b (similarity)	GenBank Accession number
OprC	<i>Pseudomonas aeruginosa</i>	Unknown	7 x 10 ⁻¹²	27.1 (34.5)	D28119
IrgA	<i>Vibrio cholerae</i>	Vibriobactin	2 x 10 ⁻¹¹	27.5 (36.6)	U72152
CirA	<i>Escherichia coli</i>	Unknown	1 x 10 ⁻⁰⁹	24.0 (33.3)	J04229
OprC	<i>Aquifex aeolicus</i>	Unknown	5 x 10 ⁻⁰⁹	26.3 (37.1)	AE000694
BtuB	<i>Escherichia coli</i>	Vitamin B12	8 x 10 ⁻⁰⁹	29.9 (39.1)	M10112
BtuB	<i>Salmonella typhimurium</i>	Vitamin B12	1 x 10 ⁻⁰⁸	29.1 (38.1)	M89481
PhuR	<i>Pseudomonas aeruginosa</i>	Hemin	2 x 10 ⁻⁰⁸	29.0 (38.8)	AF055999
ChuA	<i>Escherichia coli</i>	Haem	5 x 10 ⁻⁰⁸	27.0 (34.8)	U67920
ShuA	<i>Shigella dysenteriae</i>	Haem	5 x 10 ⁻⁰⁸	26.7 (34.8)	U64516
NosA	<i>Pseudomonas stutzeri</i>	Unknown	9 x 10 ⁻⁰⁸	24.9 (33.3)	M60717

Table 4.3 Outer membrane receptor proteins which show significant homology with OmpX. Proteins with homology to OmpX were identified from the GenBank database using the BLAST program at NCBI. ^aThe BLAST E(xpect) value estimates the statistical significance of the match by specifying the number of matches within a given score that are expected in a search of a database of this size. ^bSimilarity and identity of the protein sequences were calculated using the BESTFIT program of the Wisconsin GCG package using a gap penalty of 2.

TonB-box			Box II		
BtuB	5	PDTLVVTAN	BtuB	565	IANLFDKDY
ChuA	1	TETMTVTAT	ChuA	603	LGNAFDKEY
CirA	5	GETMVVTAS	CirA	608	VLNLGDKDL
F700	3	EQTMIVSAA	F700	636	VDNLFDKEY
Fiu	10	DDTLVVEAS	Fiu	694	VYNLFDTDY
FecA	22	GFTLSVDAS	FecA	705	VKNIFDQDY
FepA	11	DDTIVVTAA	FepA	674	VDNLFDKRL
FhuA	6	EDTITVTAA	FhuA	680	VNNLFDREY
FhuE	5	EETVIVEGS	FhuE	661	VNNLFDKTY
IutA	5	DETFVVSAN	IutA	660	IENLFDRDY
OmpX	4	SDTLTVWSS	OmpX	605	VNNLFDTKY

Box III

BtuB	105	QRVEYIRGPRSAVYGSDAIGGVVNIITTR
ChuA	98	KRVEIVRGPSALLYGSGALGGVISYDTVD
CirA	106	ERIEVVRGPMSSLYGSDALGGVVNIITKK
f700	108	QNVEVLRGPFSAALYGN-ASGGVMNVTTQT
Fiu	117	EQVEVIKGPSGTDYGRSAPTGSINMISKQ
FecA	190	DAIDVVRGGGAVRYGPQSVGGVVNFVTRA
FepA	120	ERIEVLRGPARARYGNGAAGGVVNIITKK
FhuA	127	ERAEMRGPVSVLYGKSSPGLLNMVSKR
FhuE	120	ERVEVVRGATGLMTGTGNPSAAINMVRKH
IutA	102	HHIEVIFG-ATSLYGGGSGGLINIVTKK
OmpX	96	GAVEVSKGYSSLLQGNQMGAINITQK

Figure 4.11 Consensus amino acid sequences found in known and predicted TonB-dependent receptors of *E. coli* and OmpX. The complete mature sequences of the nine known *E. coli* TonB-dependent receptors plus f700 (see below) were aligned with the deduced protein product of *ompX* using the CLUSTALW program of the Wisconsin GCG package. Shown above are the three regions of high amino acid conservation, or "TonB boxes", common to all TonB-dependent outer membrane receptors (Bäumler and Hantke, 1992); the numbers indicate the position of the first residue in each alignment in the mature protein, bold letters indicate residues identical in at least 50% of the proteins. The gene for Fiu has been described and mapped on the *E. coli* chromosome (Hantke, 1990), however its DNA sequence has not been determined. BLAST homology searches of the recently completed *E. coli* K-12 genome sequence (Blattner *et al.*, 1997) identified two potential open reading frames for *fiu*; f700 and f760 (GenBank accession numbers 1787723 and 1787024). f760 was subsequently identified as *fiu* as it was found to be located within the 18.1 min region of the *E. coli* K-12 chromosome which coincides exactly with the location previously mapped by Hantke (1990). The function of f700 is as yet undetermined, but its significant homology with the other *E. coli* TonB-dependent outer membrane receptors indicates that it, like OmpX, may represent a previously undescribed receptor.

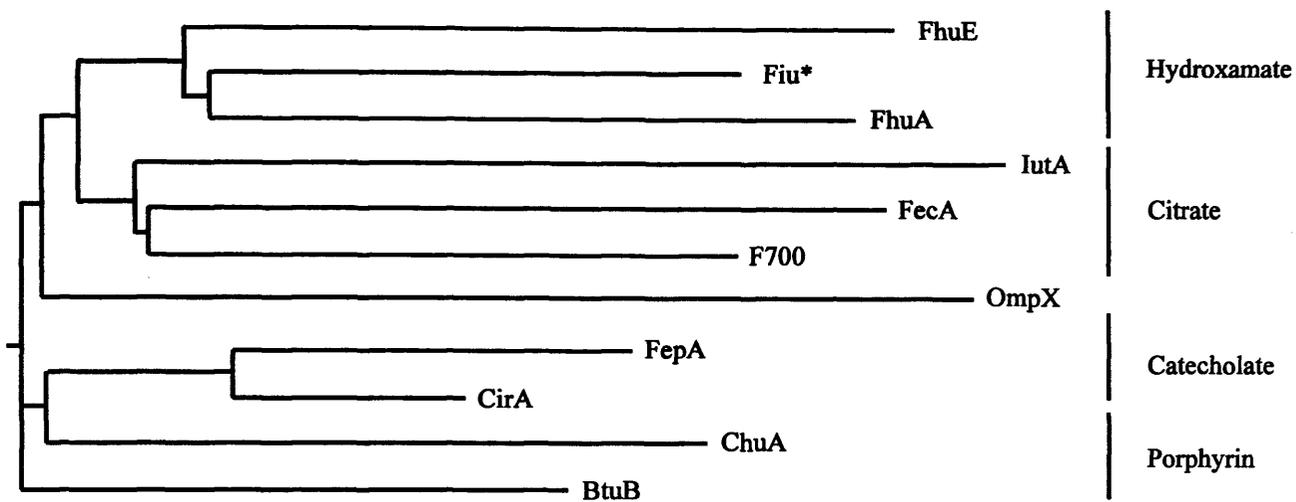


Figure 4.12 Dendrogram representing the phylogenetic analysis of the nine known *E. coli* TonB-dependent outer membrane receptors and OmpX and f700. The mature protein sequences of the *E. coli* TonB-dependent outer membrane receptors were aligned using the program CLUSTALW (available as part of the Wisconsin GCG package, Materials and Methods section 2.15; Higgins and Sharp, 1989) and the output saved in the PHYLIP format. Evolutionary analysis was made using the programs of the PHYLIP V3.7 package (compiled by Felsenstein, 1989): evolutionary distances were calculated using the PROTDIST program (with the Dayhoff PAM matrix option active); and then these distances were used to construct a phylogenetic tree with the FITCH program, using a randomised input order and the global rearrangement option (Fitch and Margoliash, 1967). The final tree was drawn as a dendrogram using the DRAWTREE program.

In the dendrogram depicted above the length of the horizontal lines can be equated to evolutionary distance. Also indicated by vertical lines are the positions of the four ligand-based receptor subfamilies proposed by Rakin *et al.* (1994) from their analysis of the *E. coli* and *Yersinia* TonB-dependent receptors; it is unclear why Fiu (*) which has been shown to be able to take up catecholate siderophores (Hantke, 1990) is localised with the hydroxamate subfamily in this analysis.

4.2.3.6 *orf1'* and *orf2* have homologues in *Haemophilus influenzae* which are located within operons of iron uptake gene homologues. The partial open reading frame *orf1'* encodes a C-terminal polypeptide of 159 amino acids; comparison of this amino acid sequence with the GenBank database using BLAST identified a single significant match (BLAST algorithm E value $< 5 \times 10^{-24}$) to the product of open reading frame *yc73* predicted by the *Haemophilus influenzae* genome sequence (Fleischmann *et al.*, 1995; **Table 4.2**). The *yc73* open reading frame encodes a 268 amino acid polypeptide, predicted by PSORT to be cytoplasmic, which shares 35.8% identity (60% similarity) with *Orf1'* (**Figure 4.13A**) but has no other known homologues. *yc73* is predicted to be the third gene in an operon of three (GenBank accession number U32807; Fleischmann *et al.*, 1995) in which the other two genes show significant homology to the *E. coli* genes *fecD* and *fepC* that encode the integral membrane protein and ATP-binding protein for the ferri-dictrate and ferrienterobactin uptake systems respectively (Pressler *et al.*, 1988; Pierce and Earhart, 1986).

orf2 encodes a predicted polypeptide product of 285 aa with a molecular weight of 30.9 kDa; PSORT analysis failed to identify a consensus N-terminal signal sequence indicating that *Orf2* is also cytoplasmic. A GenBank database search using BLAST identified two high homology matches; *modD* from *Rhodobacter capsulatus* and a predicted open reading frame identified by the *H. influenzae* genome sequence (now also named *modD*) (Blast algorithm E values $< 4 \times 10^{-33}$ and $< 2 \times 10^{-56}$ respectively) whose polypeptides share 36.5% and 45.8% identity, respectively, with the predicted gene product of *orf2* (Wang *et al.*, 1993; Fleischmann *et al.*, 1995; **Figure 4.13B**). *R. capsulatus modD* is the fourth gene of the *mod* operon, which encodes the high affinity molybdenum uptake system; this operon also contains *modA*, which encodes a periplasmic binding protein, and *modB* and *modC* which encode the integral membrane protein and ATP-binding protein respectively of a typical periplasmic transport system (Wang *et al.*, 1993; Ames, 1986). Wang *et al.* (1993) showed that *R. capsulatus modD* mutants required four-fold higher molybdenum concentrations (*modA*, *modB* or *modC* mutants require 500-fold higher), but the function of the cytoplasmically located

ModD was not definitively determined. *E. coli* K-12 contains a high affinity molybdate uptake system homologous to that in *R. capsulatus* and which is encoded by the *modABC* genes (Rech *et al.*, 1994; *modBC* were previously known as *chlJD*). A fourth open reading frame identified downstream of *modC* was predicted to be the *E. coli modD* homologue however it has since been shown that this gene has no homology to *R. capsulatus modD* and is not required for molybdate uptake (Maupin-Furlow *et al.*, 1995). The *modD* gene of *H. influenzae* was named due to its high homology to *modD* of *R. capsulatus*; however it is not part of the *H. influenzae* high affinity molybdate uptake operon and is instead predicted to be the second gene in an operon of five (Genbank accession number U32825; Fleischmann *et al.*, 1995) wherein the others all show homology to genes involved in periplasmic iron uptake in bacteria: i.e. *sfuC* of *Serratia marcescens*, *fepC* of *E. coli*, and *feuA* and *feuB* of *Bacillus subtilis* (Angerer *et al.*, 1990; Pierce and Earhart, 1986; Quirk *et al.*, 1994; Liu *et al.*, 1997).

4.2.3.7 *orf3* and *orf4*. Open reading frame *orf3* is predicted to encode a polypeptide of 244 amino acids with a molecular weight of 27.9 kDa. Analysis of the N-terminus of the Orf3 polypeptide using program PSORT predicts an uncleavable signal sequence but the polypeptide does not contain any membrane spanning regions, and so the protein is predicted to form a peripheral association with the inner membrane. A BLAST search of the GenBank database failed to identify any significant matches with the Orf3 polypeptide.

orf4 is predicted to encode a 304 amino acid polypeptide with a molecular weight of 33.6 kDa. Analysis with the PSORT program failed to identify a signal sequence or membrane spanning regions, and so Orf4 is predicted to be cytoplasmic. A search of the GenBank database using BLAST identified a single high homology match (BLAST algorithm E value $< 2 \times 10^{-18}$) to the product of the *mccF* gene from the *E. coli* plasmid pMccC7 which shares 31.3% identity with the Orf4 polypeptide (Figure 4.14). *mccF* is part of an operon of 6 genes which encode the proteins required for production of the translation-inhibiting peptide antibiotic microcin C7 (MccC7)

Figure 4.13A Amino acid alignment of pRDH8 Orf1' and yc73 of *Haemophilus influenzae*. The deduced polypeptide product of the incomplete open reading frame *orf1'* of pRDH8 showed 35.8% identity (60% similarity) with the C-terminus (residues 113 to 268) of the protein *yc73* identified by the *H. influenzae* genome sequencing project (Fleischmann *et al.*,1995; see also Table 4.2). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues that are identical in both sequences are boxed in black, residues which share similar characteristics are boxed in grey.

Figure 4.13B Amino acid alignment of pRDH8 Orf2 with ModD of *Rhodobacter capsulatus* and ModD of *Haemophilus influenzae*. The deduced protein product of open reading frame *orf2* of pRDH7 showed 36.5% identity (56.1% similarity) to *R. capsulatus* ModD and 45.8% identity (62.9% similarity) to *H. influenzae* ModD (see also Table 4.2). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues which are identical in two of the polypeptides are boxed in grey , those identical in all three are boxed in black.

A

Orf1' : -----IQAWEEEDWSDLPRCDIAVASRSTLVADNRQAMSKLNNO : 39
 yc73 : KQKFGHLHLLTTFHLSWADKWDVDPQADVVLASRSTLVDDDDMDEKLCAK : 150

Orf1' : ARLRVMTTHLVSTSFVSPAQRAGREVIELPNYIFALNVLYQMGIYAHV : 89
 yc73 : AKKRVELTSTQQRHFLDEGVFEAIGREDIGFPTYIYLLNRLYQKGIQANL : 200

Orf1' : DFIRGON-CQQDNSWFERFEQNVSWSLGAINDDERERLYRYQ--QQDAR : 136
 yc73 : NFIETESGCFQ-GESYEDLLASVEEESLGEI SEKEKQGLKAYDRKQANNE : 249

Orf1' : ALATASDVALIWWDSVPQETLR : 159
 yc73 : PLSHGQKVALIWWNVVDRI---- : 268

B

Rhoca : MFLIDETAIMALIREDVPAGDLTTESLGLAAEAG--ELTMARGAMTVACSEEAVRIIQL : 58
 Haein : MYFSDKELLDLLEDIYRGDLTTHALSLENIPAKLEKRNACVFTVAGVSVAEKLLKK : 60
 Orf2 : MFLSQAQIDALLLEDIQGGDLTTRALNIGHQHYYEFLQGG--CVSGISVACKMITT : 58

Rhoca : LGAEAWITASGROVEGGEMLLFARGQVEALLAEWKVAQNLI EWAS--GASATAIVVA : 116
 Haein : IDIQPHLYKEGEWVESGALLSAEGMSEQLHQAWKVVQLVLEWSCFTVAQYTAEMIAN : 120
 Orf2 : LILTIDDSGGSANAGQRLIRAQGNAAALHOGWKAVQNVLEWSLRFIIILKCWRYF : 118

Rhoca : ARAVNEAVTVACTRKSVPGTRATSLRAVTVGGATVHRTGLSDSVLLFA--EHRAEFGG-AD : 173
 Haein : AKSVNPTVAVACTRKSIPNTRKLATNAVLAAGCHHRQGVSETLLVFFTTNHRNLLSIPN : 180
 Orf2 : VN-VTLMVIFACTRKAIPGTRLPSQAIIAAGELHHRAGCAETILLFA--NHRHELHINQ : 175

Rhoca : ALAAQIARLRASCPEKVVVEVADVREALAAQAGAEVLEKFPFQ-VAVVVGIGD : 232
 Haein : TKIVDLRQEAPENKITTLEENYAQFEQMYMAEEDIIQLDKWLLDFTVKKALDLQS : 240
 Orf2 : WSGAINQLRRHAPKIKVVEADAPKEAIALRAQDVLQDKESQ--QATEIQIAPS : 233

Rhoca : WRG--HVAVGGTAAAAPMRRRGRRFW----- : 259
 Haein : KQKDILSVAGGVKNVSDAKLGRLEFITSAPYVPPEDIKVVIEKI--- : 289
 Orf2 : LAPHCTALTGGNLTTLKNLDCGRLEFITSAPYAAAPADIKVSLQPAASI : 285

```

                20                40                60
MccF : MMIQSHPLLAAPLAVGDTIGFFSSSAPATVTAKN-RFFRCYEFQRFTKGFKIVSGKVTG : 59
Orf4 : -----LSLFLHLIAPSGYCIQHAALRGIQRLET--DAGHQYNNVEVIA : 40

                80                100               120
MccF : TDFYRSCTIKERAEFNEIVYN-PDITICMSTIGGDNNSLIPFEDFTDAIIAN--PK : 116
Orf4 : RCERFAGTETERL DLNSLARLTTPTNIIIVAVRGGYASRLLADIDWQALVARQQHDEL : 100

                140               160               180
MccF : IICYSDTTALAGYAKTGITFYGRALPSFGEHPPLVDITYESFIKILFTTRKQSGI : 176
Orf4 : IICGHSDFTAQQGLAHGNVITFSGEMLVANFCAD-ELNAFTEHHFWLALRNE----- : 153

                200               220               240
MccF : YTYTLPEKWSDESINWENKILREKKLYKNNCAFYGGKVEGRVIGGNLNTLFTITGIVGS : 236
Orf4 : ---TFTIEWQGE-----GP-----ICRAEGLWGGNLAMLLIS--IIGI : 186

                260               280               300
MccF : EWMPEILNGDILFIEDSRKSIATIERLFSMFKLNRFDKVSAIILGKHELFDCTAGSKR : 296
Orf4 : PWMPKIENG-IIVLEDINEHPFRVERMLLQIYHAGILPRQKAIILGSFSGSTPNDYDAGY : 245

                320               340               360
MccF : RPYEVLTEILDGKQIPVLDGFDSCSHTHPMLTLPLGVKLAIFDNKNISITEQYLFTSTEK : 356
Orf4 : NLESVYAFVRSRLSIPLITGLDFGHEQRTVTLPLGAHAIDNTREGTQVTISGHVPLKM- : 304

```

Figure 4.14 Amino acid alignment of pRDH8 Orf4 with MccF of *E. coli*. The deduced protein product of *orf4* of pRDH7 shares 31.3% identity (53.3% similarity) with MccF (see also Table 4.2). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues which are identical in both sequences are boxed in black, residues which share similar characteristics are boxed in grey.

(González-Pastor *et al.*, 1995). MccF, which is predicted to be associated with the inner membrane, is not involved in MccC7 synthesis but can confer specific immunity to otherwise microcin-susceptible *E. coli* cells, though the basis for its protective ability is as yet unknown.

4.2.3.8 Identification of the OmpX and EmtA proteins in outer membrane preparations. Of the genes present on pRDH8 only the polypeptides encoded by *ompX* and *emtA* were predicted to contain signal sequences which would result in their localisation in the outer membrane. In order to ascertain if these predictions were correct, outer membrane preparations were made from *E. coli* strain DH5 α and the same strain transformed with plasmid pRDH8. The pRDH8-containing strain showed expression of two additional protein bands, of approximately 20 kDa and 70 kDa, which were absent in the plasmid free control (**Figure 4.15**) These masses were consistent with the predicted sizes for the mature *emtA* and *ompX* gene products, 20.4 kDa and 71.1 kDa respectively. Despite the apparent overexpression of the lytic transglycosylase protein EmtA the growth of the DH5 α strain containing pRDH8 was unimpaired and there was no obvious lysis of the cells; this is in agreement with the data of Kraft *et al.* (1998) who showed that overexpression of EmtA from a *tac* promoter also had no obvious effects upon growth rate. Growth of DH5 α in iron restricted medium resulted in the greatly increased expression of five outer membrane proteins (**Figure 4.15**); these are predicted to be Cir (74 kDa), FhuE (76 kDa) FhuA (78 kDa), FepA (81 kDa) and Fiu (83 kDa) (Hider, 1984; Hantke, 1983). In contrast there was no indication that expression of the OmpX protein (which ran level with the 74 kDa Cir protein) increased in an equivalent manner when DH5 α pRDH8 was grown under the same iron limiting conditions (**Figure 4.15**).

4.2.4 The distribution of the *orf1'-orf2-ompX* operon in *E. coli*.

4.2.4.1 Comparison of pRDH8 sequence with the *E. coli* K-12 genome sequence. Southern blot analysis of *Bam*HI digested chromosomal DNA from DH5 α and E2348-

69 (Figure 3.18) and comparison of the pRDH8 DNA sequence to that of the umudca DNA sequence (section 4.2.1.4; Figure 4.2) had indicated that there were differences between the *E. coli* K-12 and EPEC chromosomal sequences at the *emtA* locus. Subsequent comparison of the DNA sequence of pRDH8 with the recently released sequence of the *E. coli* K-12 genome (Blattner *et al.*, 1997) confirmed that this region is indeed significantly different in E2348-69. The *E. coli* K-12 DNA sequence for the 26.6 min region (GenBank accession numbers AE000218 and AE000219) was found to be almost identical to that of pRDH8 for the region containing *orf4*, *orf3*, *emtA* and *tag*. However immediately after the inverted repeat sequence, which separates the *tag* and *ompX* open reading frames in pRDH8, the identity between the two DNA sequences ceased (Figure 4.16). In the *E. coli* K-12 DNA sequence the *tag* gene is followed by an open reading frame of 441 bp designated *o146*; it does not contain the sequences for *ompX*, *orf2* or *orf1'* although the region between *tag* and *o146* does retain 24 bp which in pRDH8 are predicted to encode the C-terminal 7 amino acids of the OmpX polypeptide (Figure 4.16B). Due to the existence of partial *ompX* sequence in the *E. coli* K-12 genome it was proposed that the arrangement of genes in *E. coli* K-12 was due to a deletion of DNA sequences between *tag* and *o146* in a K-12 ancestral strain.

4.2.4.2 Southern hybridisation analysis of the distribution of *orf1'-orf2-ompX* sequence in *E. coli* isolates. In order to confirm if the deletion of the *orf1'-orf2-ompX* sequence observed was restricted to the *E. coli* laboratory strain K-12 a bank of *E. coli* strains was analysed. These strains included representatives of most of the *E. coli* virotypes including both diarrhoeagenic pathogens (EPEC, EIEC, EHEC, ETEC, DAEC and EAaggEC) and uropathogenic strains, as well as two laboratory *E. coli* strains (Table 4.4). Chromosomal DNA was isolated from each of the strains using the small scale chromosomal DNA purification method. Each chromosomal DNA preparation was digested with *Bam*HI and separated by 1% agarose gel electrophoresis prior to transfer to HybondTM-N membrane by Southern blotting. Two DNA fragments from pRDH8 were used to make probes; these were the complete 6.3 kb *Bam*HI

fragment which contains sequences known to be present in both *E. coli* K-12 and E2348-69, and the 3.3 kb *AccI* fragment which contains just the E2348-69 specific *orf1'-orf2-ompX* operon (Figure 4.16). 50 ng of each fragment were labelled using the Fluorescein Gene ImagesTM system and the probes formed were used in sequential hybridisations with the same Southern blot. All of the strains examined had chromosomal sequences which hybridised with the 6.3 kb *Bam*HI probe though the blot showed some variation in the *Bam*HI fragment sizes between strains (Figure 4.17A). In contrast only 12 of the 23 strains analysed were found to hybridise to the 3.3 kb *AccI* probe; in 11 of these strains the probe hybridised to a *Bam*HI fragment of approximately 6.0 kb whilst in the twelfth strain 86-24 (EHEC, O157:H7) it identified a larger *Bam*HI fragment apparently due to a restriction fragment length polymorphism (Figure 4.17B). The data obtained for the diarrhoeagenic *E. coli* indicated that the *orf1'-orf2-ompX* sequence was present in the DAEC strain, the O157:H7 EHEC strain and in 7/11 EPEC strains; however, it was not present in either the ETEC, EIEC, EAaggEC or O26:H11 EHEC strains tested. In the uropathogenic *E. coli* strains *orf1'-orf2-ompX* was present in 3/4 strains tested, but it was only found in 1/2 strains with the cystitis virotype. As expected the *orf1'-orf2-ompX* sequence was not found in the *E. coli* K-12 strain DH5 α or in the K-12/B hybrid strain HB101.

4.2.5 Cloning and analysis of the region upstream of *ompX* which is deleted in *E. coli* K-12.

Cloning of DNA sequences from a library of chromosomal DNA by colony blotting is a time consuming process and does not guarantee to give the complete region of interest on a single cloned fragment. As an easier alternative it was determined to clone DNA fragments containing the complete *ompX* operon directly using the kanamycin resistance gene present in the *TnphoA* insertion in the *tag* gene in mutant #181 as a selective marker. The choice of useful restriction enzymes for cloning was limited as only a relatively small number of enzymes either do not cut in *TnphoA* or cut only upstream of the kanamycin resistance gene (Figure 3.2); the choices were further hampered by the presence of cleavage sites for a number of these restriction enzymes

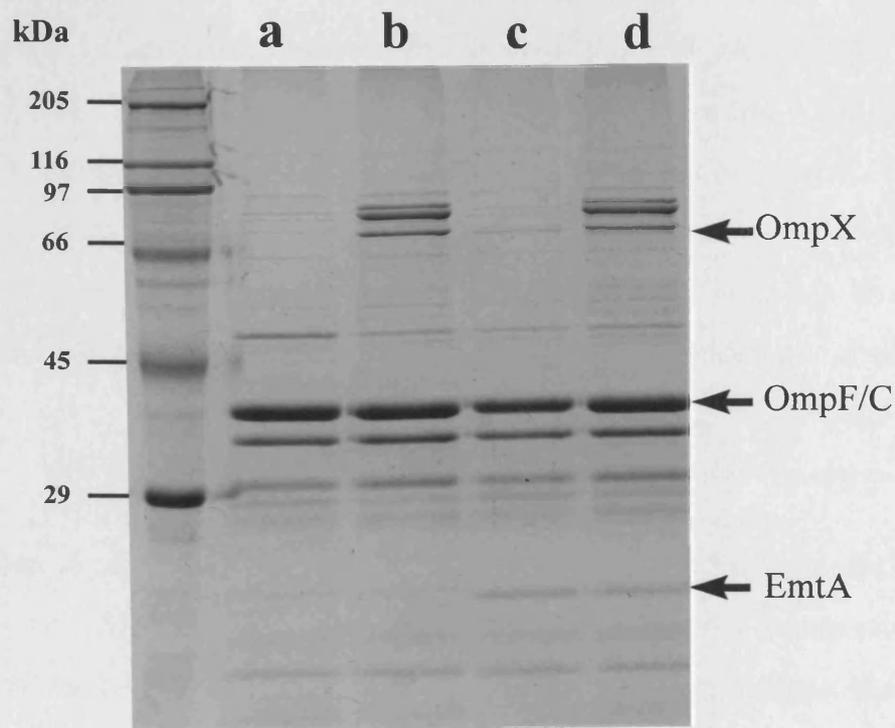


Figure 4.15 Identification of EmtA and OmpX in *E. coli* outer membranes. Outer membranes were prepared (as described in Materials and Methods, section 2.16.3) from overnight cultures of *E. coli* strains DH5 α (lanes a and b) and DH5 α pRDH8 (lanes c and d), which had been grown in LUB (lanes a and c) or in iron restricted medium, LUB containing 300 μ M 2,2'-dipyridyl (lanes b and d). Protein samples were separated on 13% SDS-PAGE gels and visualised by Coomassie staining; equal sample loadings were achieved by normalising the levels of the OmpF/C bands.

Two bands of approximately 20 kDa and 70 kDa (indicated by arrows), which are predicted to be EmtA and OmpX respectively, are evident in DH5 α pRDH8 outer membrane preparations (lane c) but absent in the DH5 α control (lane a). Growth in iron restricted media resulted in the greatly increased expression of five *E. coli* outer membrane proteins in the 70-80 kDa region (lanes b and d). There was no evidence for an equivalent increase in the expression of the 70 kDa OmpX protein or the 20 kDa EmtA protein under iron limiting conditions (lane d).

Figure 4.16A Diagrammatic representation of the *emtA* loci in EPEC strain E2348-69 and *E. coli* K-12. Figure (i) shows the open reading frames identified between the *dadX* and *treA* genes in the 26.7-26.9 min region of *E. coli* K-12 (Blattner *et al.*, 1997). Figure (ii) depicts the restriction enzyme map and open reading frames of the equivalent chromosomal region from EPEC strain E2348-69 identified in plasmid pRDH8; *orf3* and *orf4* are identical to *f224* and *f304* respectively. The 3.4 kb *AccI* fragment used as a probe for *ompX* is indicated as a solid bar. (A-*AccI*, B-*BamHI*, G-*BglII*, C-*ClaI*, E-*EcoRV*, H-*HindIII*, K-*KpnI*, S-*SphI*)

Figure 4.16B Alignment of the pRDH8 DNA sequence with the *E. coli* K-12 26.82 min chromosomal DNA sequence. The sequences of pRDH8 and *E. coli* K-12 (GenBank Accession numbers AE000218 and AE000219) share 95% identity to nucleotide 3120 of pRDH8 and 25% identity beyond this point (identical nucleotides are represented by connecting vertical lines). The seven C-terminal residues of the OmpX polypeptide which appear to be conserved between E2348-69 and *E. coli* K-12, and the two 13 bp inverted repeat sequences which are predicted to form a Rho-dependent terminator loop ($\Delta G = -19.29$ kcal/mol) are shown bold and underlined .

Figure 4.17 Southern blot analysis of multiple *E. coli* virotypes using an *ompX* specific probe. (opposite page). Chromosomal DNA prepared from 24 *E. coli* strains detailed below (Table 4.4) was digested with *Bam*HI, separated on 1% agarose gel electrophoresis and transferred by Southern blotting to nylon membrane. The Southern blot was probed with the complete 6.3 kb *Bam*HI fragment of pRDH8 which contains both *E. coli* K-12 and E2348-69 sequences (A); it was then stripped and reprobed with the 3.3 kb *Acc*I fragment of pRDH8, which contains just the E2348-69 specific *ompX*, *orf2* and *orf1* ' open reading frames (B).

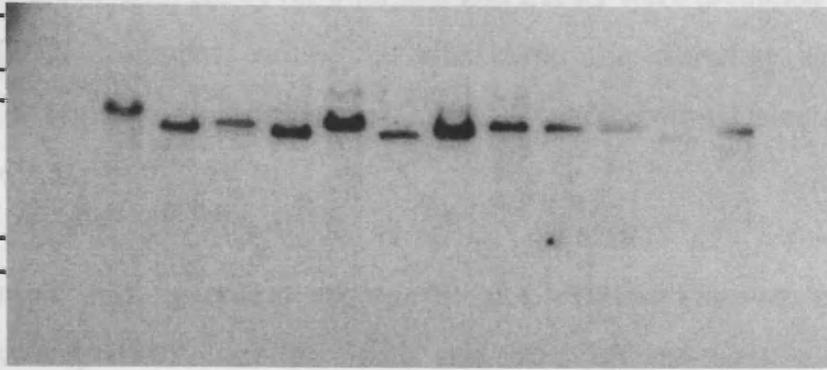
Strain	Serotype	Virotype	Lane	<i>ompX</i>
E2348-69	O127:H6	EPEC	a	+
E11881B	O25:H42	ETEC	l	-
17-2	O3:H2	EAggEC	m	-
1457-75	O124:H3	EIEC	n	-
C1845	O75:NM	DAEC	o	+
86-24	O157:H7	EHEC	p	+
3605-73	O26:H11	EHEC	q	-
CH7/6 (Charles)	O26:H1	EPEC	r	-
261/88	O86:H34	EPEC	s	+
19	O111:H2	EPEC	t	+
KH1/8 (Khoyer)	O114:H2	EPEC	u	-
57	O119:H6	EPEC	v	+
292	O125:	EPEC	w	+
E611	O126:	EPEC	b	-
IR3/8	O127:H-	EPEC	c	+
44/86	O128:H2	EPEC	d	-
53/85	O142:H34	EPEC	e	+
1187	O75:K5:H-	Cystitis	g	-
3071	O6:K2:H1	Septicaemic Pneumonia	h	+
1192	O1:K1:H7	Pyelonephritis	i	+
1152	O6:K13:H1	Cystitis	j	+
HB101	K-12/B	N/A	k	-
DH5 α	K-12	N/A	f	-

Table 4.4 *E. coli* strains used to screen for *ompX*. The diarrheagenic *E. coli* strains b-e and l-w were a generous gift from Dr Stuart Knutton, Institute of Child Health, Birmingham. The uropathogenic strains g-j were selected from a collection of *E. coli* strains supplied by Dr Timo Korhonen, University of Helsinki.

A**kb**

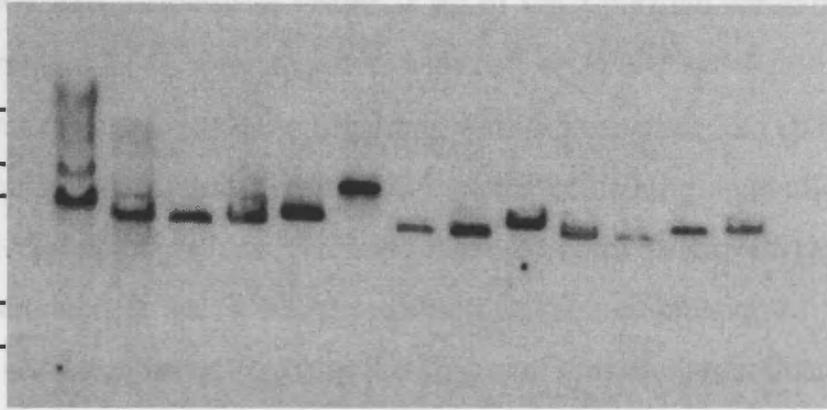
23.1
9.4
6.6
2.3
2.0

a b c d e f g h i j k a



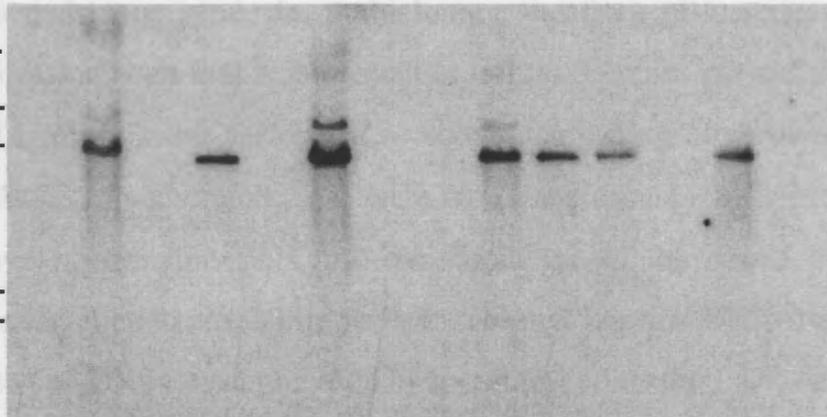
a l m n o p q r s t u v w

23.1
9.4
6.6
2.3
2.0

**B**

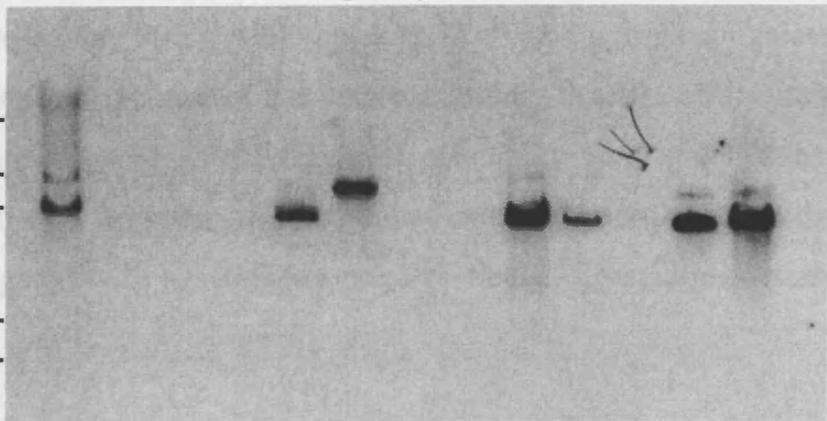
a b c d e f g h i j k a

23.1
9.4
6.6
2.3
2.0



a l m n o p q r s t u v w

23.1
9.4
6.6
2.3
2.0



(*Cla*I, *Eco*RV and *Kpn*I) in the sequence immediately downstream of the *TnphoA* insertion site in #181 (Figure 4.7). However both *Sst*I and *Eco*RI were predicted to cut in #181 to release DNA fragments which included both the complete *TnphoA* kanamycin gene and a region of chromosomal DNA of undetermined size which extends upstream of *orf1*.

4.2.5.1 Cloning of *ompX* and upstream regions from CVD206::*TnphoA* mutant #181. It was initially attempted to clone the *Eco*RI and *Sst*I fragments containing the kanamycin gene directly into pBluescript vector. 2 µg of CsCl purified chromosomal DNA prepared from mutant #181 were digested with *Sst*I or *Eco*RI and ligated with 200 ng of similarly cut pBluescript DNA. Ligations were transformed into DH5α by electroporation and transformants selected on LUA plates containing ampicillin and kanamycin; despite several attempts it proved impossible to obtain kanamycin resistant transformants in this manner. It was thought possible that the combination of a high copy number plasmid and large insert size (the *Sst*I fragment must be larger than 13 kb and the *Eco*RI fragment must be at least 10 kb) could be responsible for the lack of transformants, and so an alternative cloning vector was sought. Plasmid pWKS30 contains an ampicillin resistance gene, the multicloning site from pBluescript and a pSC101 origin of replication such that it is present at only 6-8 copies per cell (Wang and Kushner, 1991). Ligations were performed as above using 2 µg of chromosomal DNA and 200 ng of pWKS30 vector and these were then transformed into DH5α. The *Sst*I ligation again gave no transformants, but transformation of the *Eco*RI ligation resulted in a single kanamycin and ampicillin resistant colony. Plasmid DNA from this transformant, designated pRDH44, was prepared by the miniprep method. Digestion of pRDH44 with *Eco*RI identified 5 DNA fragments in addition to the expected 5 kb pWKS30 fragment; these were 200 bp, 700 bp, 3 kb, 7 kb, and a fragment greater than 20 kb. It was apparent from the size of the inserts that only the 20+ kb fragment was large enough to be the expected *Eco*RI fragment containing the *TnphoA* kanamycin gene. Therefore the 20+ kb fragment was subcloned into pBluescript; *Eco*RI digested pRDH44 DNA was ligated with an 10-fold excess of *Eco*RI digested pBluescript, the

ligation transformed into DH5 α and transformants selected on kanamycin and ampicillin LUA plates. Plasmid DNA was prepared from the resultant colonies and a transformant was identified by restriction digestion and agarose gel electrophoresis that contained only the 20+ kb *EcoRI* fragment in pBluescript; this plasmid was named pRDH45. The *EcoRI* insert of pRDH45 was sized by digestion of plasmid DNA with *NotI* (*TnphoA* contains 2 sites, pBluescript contains 1 site) and separation by agarose gel electrophoresis against standard DNA markers. The digestion gave 5 fragments of approximately 1.4, 3.0, 4.6, 9.0 and 16 kb respectively indicating a total insert size of approximately 31 kb (Figure 4.18A).

4.2.5.2 Partial DNA sequencing and mapping of pRDH45. The DNA sequence of the ends of the pRDH45 *EcoRI* insert was determined with the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit using CsCl purified pRDH45 DNA and the primers m13f and m13r (section 2.12). As expected the DNA sequence of one end of the *EcoRI* fragment (m13r) was identical to the 3' region of *TnphoA*. The m13f derived DNA sequence was identified using the BLAST program to be 99% identical with a sequence from the *E. coli* K-12 genome sequencing project (accession number AE000219; Blattner *et al.*, 1997). This DNA sequence encodes a predicted open reading frame *yehM* which is located approximately 14 kb downstream of the open reading frame *o84 (tag)* in *E. coli* K-12. By comparison of the positions of the *NotI* cleavage sites present in the sequence of pBluescript, *TnphoA* and the *tag-yehM* region of *E. coli* K-12 with the *NotI* fragments identified in pRDH45 it was possible to construct a tentative map for pRDH45 (Figure 4.18B). If it is assumed that the deletion which has occurred in *E. coli* K-12 was not accompanied by further alterations in the chromosomal sequence within the *tag* region it can be calculated that the deleted region containing *ompX* is approximately 10 kb in length.

To determine the DNA sequence of the distal end of the E2348-69 *ompX* upstream region, a sequencing primer rdh45a (Materials and Methods, 2.12) was designed which would anneal to a sequence 150 bp downstream of the proposed deletion point in the

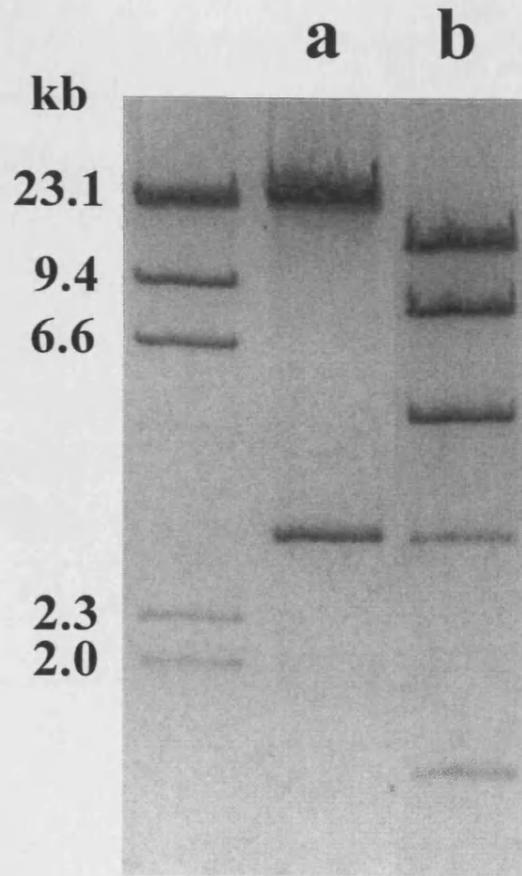
E. coli K-12 sequence and located within the *ol46* open reading frame (**Figure 4.16; Figure 4.18B**). Despite repeated attempts using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing system and proven CsCl prepared pRDH45 plasmid it was impossible to obtain DNA sequence data using the *rdh45a* primer. It is not clear if this is due to primer design or to further alterations in the sequence of E2348-69 compared to *E. coli* K-12 within this region.

4.2.5.3 Identification of the N-terminus of the Orf1 polypeptide and a further upstream gene in the operon. Primer *orf1r* was designed to anneal to a sequence approximately 100 bp from the *Bam*HI site of pRDH8 within the *orf1* open reading frame (**Figure 4.18B; section 2.12**). Duplicate DNA sequences were determined using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit with CsCl prepared pRDH45 DNA and primer *orf1r*; comparison of the duplicate sequences using program BESTFIT showed identity over approximately 600 bp. Analysis of the *orf1r*-derived DNA sequence using the MAP and FRAMES programs identified two potential open reading frames. Comparison of overlapping sequences with pRDH8 identified the first open reading frame as the 5' end of *orf1*. The second 180 bp partial open reading frame, designated *orf5'*, terminated just 9 bp upstream of the predicted *orf1* start codon (**Figure 4.19**). A BLAST search of the GenBank database using the 270 amino acid full length Orf1 polypeptide (molecular weight 31.0 kDa) again identified a single high homology match with the *Haemophilus influenzae* open reading frame *yc73*. Analysis using the GAP program found that the homology between the full length Orf1 and *yc73* polypeptides (36.1% identity) was very similar to that previously observed for the N-terminal sequences (35.8% identity) (**Figure 4.20A**). Analysis of the full length Orf1 polypeptide with program PSORT failed to detect an N-terminal signal sequence and predicted that the protein would be cytoplasmically located. A BLAST search against the GenBank database using the 60 amino acid polypeptide encoded by *orf5'* identified a single significant match (BLAST algorithm E value < 0.001) with the deduced gene product of the *H. influenzae* open reading frame *yc72* (**Figure 4.20B**). The *yc72* open reading frame is located directly upstream of the *yc73* open reading

Figure 4.18A Restriction enzyme mapping of plasmid pRDH45. Restriction enzyme digestions (*EcoRI* - lane a, *NotI* - lane b) of pRDH45 were separated by 1% agarose gel electrophoresis. *EcoRI* fragments were estimated at 3 kb (pBluescript) and 20 kb; *NotI* fragments were estimated at 1.4 kb, 3.0 kb, 5.6 kb, 9.0 kb and 16 kb.

Figure 4.18B Plasmid map of pRDH45. A map of pRDH45 was constructed based upon restriction enzyme digestion analysis (see above) and DNA sequencing data (section 4.2.5.2). Indicated are the regions of known DNA sequence; pBluescript (pBS), *TnphoA*, *E. coli* K-12 (*ychM-treA*) and pRDH8 (*orf1-ompX*). The region containing *ompX* which is deleted from *E. coli* K-12 was estimated to be 10 kb (7 kb of DNA sequence upstream of *orf5* have still to be determined). Also indicated are the sites of the sequencing primers m13f, m13r, *orf1r* and *rdh45a*.

A



B

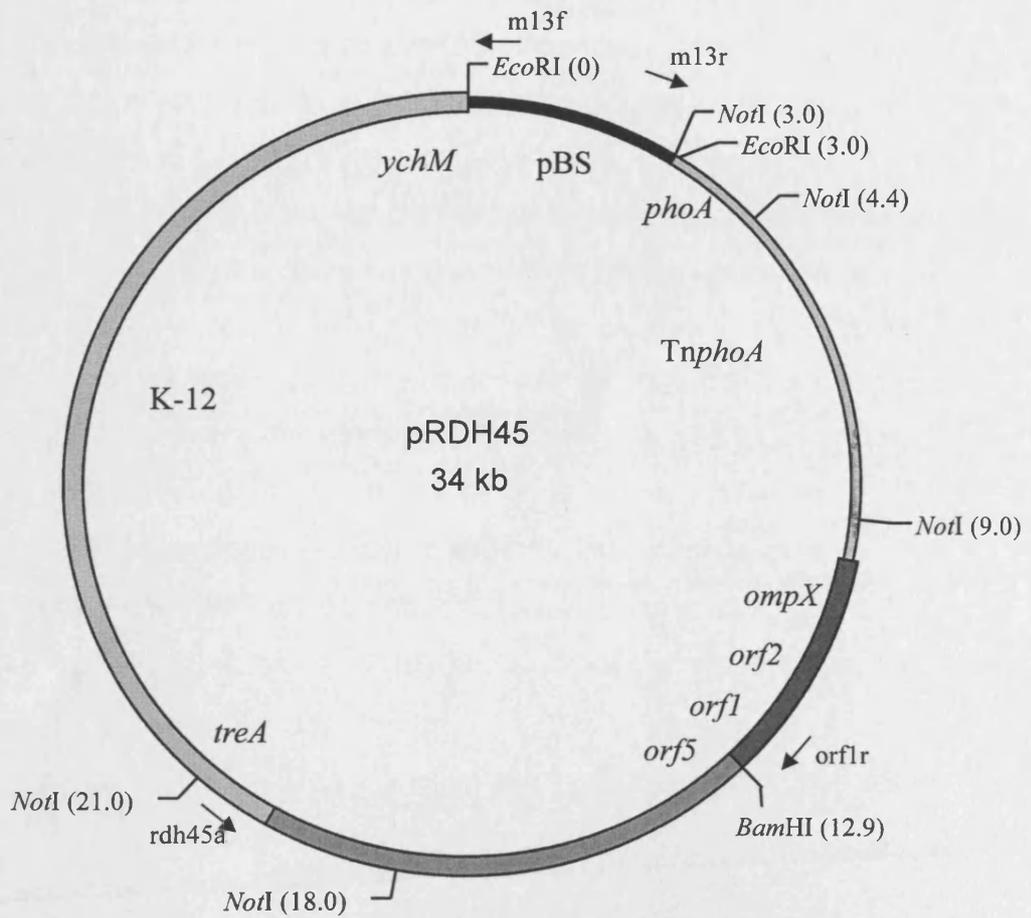


Figure 4.19 Identification of open reading frames in pRDH45. The 600 bp of DNA sequence determined from the plasmid pRDH45 using primer orf1r (see also Figure 4.18) was analysed using the MAP program of the Wisconsin GCG package and found to contain two partial open reading frames encoded on the complementary strand; these were designated *orf1'* and *orf5'*. The stop codon for *orf5'* and start codon for *orf1'* are shown in bold and underlined, and the predicted polypeptide products Orf1' and Orf5' are indicated. The initial 70 bp of the orf1r derived pRDH45 DNA sequence is identical to the last 70 bp of the pRDH8 DNA sequence (Appendix 2); the C-terminal 24 residues of pRDH45 Orf1' (underlined) are identical to the deduced N-terminal 24 amino acids of the pRDH8 *orf1'* gene product.

Figure 4.20A Amino acid alignment of pRDH45/pRDH8 Orf1 with yc73 of *Haemophilus influenzae*. The deduced polypeptide sequence of the *orf1* open reading frame (compiled from pRDH8 and pRDH45 DNA sequences) shows 38.0% identity (45.6% similarity) with the yc73 gene product of *H. influenzae* (see also Table 4.2). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues that are identical in both sequences are boxed in black, residues which share similar chemical characteristics are boxed in grey.

Figure 4.20B Amino acid alignment of pRDH45 Orf5' with yc72 of *Haemophilus influenzae*. The deduced polypeptide sequence of the 3' partial open reading frame *orf5'* of pRDH45 shows 33.9% identity (47.5% similarity) to the C-terminus of the *H. influenzae* gene product yc72 (GenBank Accession number U32807; Fleischer *et al.*, 1995). Sequences were aligned using the CLUSTALW program of the Wisconsin GCG package; residues that are identical in both sequences are boxed in black, residues which share similar chemical characteristics are boxed in grey.

A

Orf1 : MLIDDIIDFADLYLQQLRLAHRTEKTFDHWDRRAEKMAENCASPTDSYLQQLTSKIDLOGA : 60
yc73 : MTTIYDINFALYQQLIACNHYNLPPKIKWDKRAVKMAENLVGKPSAYNQQLLQAMNVQTD : 60

Orf1 : QTLFDGCGPGTVSIALADKLTITTYGVDYSQGLNVAARRAAALKADNHWIQFAWEEDW : 120
yc73 : EAVLDGCGPGTFAVPLAQQGSTIYADYSNGMLDCLAQFKQKFGFLHHTTFHNSWADNW : 120

Orf1 : SDIPRCDAVASRSTLVADMRQAISKLNQAALRVYTHVSTSFYSPAQRAAGREVIE : 180
yc73 : DDVPOADAVLASRSTLVDDDDMTEKLCAKAKRVVLTSTQRHFIDEQFEAIGREDIG : 180

Orf1 : LPNYTIALNVLYQMGIYAHVDFIRGN-CQQDNSLWERFEQNVSWSLGALNDDERERLYR : 239
yc73 : FPTYIALLNRLYQKGIQANLNFIETSGCFQ-GESVEDLLASVEFSLGELSEKELQGLKA : 239

Orf1 : YQ--QPARAATASDWALIWWDSVPQETLR : 270
yc73 : YDRKQANNEPISHGQKAWALIWWNVDR----- : 268

B

yc72 : IRVLEKIKQLQKQNIATITSTHNPQQAALGDNVLDQQFGFQQGDKKHLTLENLAKIY : 244
Orf5' : -----MLMSTHHLHANAIALDSTIQEPPDGRVTQGLPTEQLTTKKLAAIY : 45

yc72 : RTSPELHQLNHNHIEKSL : 263
Orf5' : RVSADQHHHLSAIKH--- : 61

frame in *H. influenzae* and its gene product has significant homology to the *E. coli* ATP-binding inner membrane protein FepC, which is required for the uptake of ferrienterobactin (Pierce and Earhart, 1986).

4.3 Discussion.

4.3.1 pRDH7 encodes the proximal region of an operon for BFP biogenesis.

The *TnphoA* mapping and DNA sequencing of plasmid pRDH7 identified a putative operon containing six open reading frames each of which were predicted to encode extracytoplasmic gene products. The DNA sequence of the first open reading frame in the operon was found to be almost identical with that reported by Donnenberg *et al.* (1992) for the *bfpA* gene, and the predicted products of two further open reading frames shared homology with proteins from other type IV pili operons i.e. Orf2 with MshD and PilN (Hase *et al.*, 1994; Kim and Komano, 1997), and Orf5 with TcpT (Oigerman *et al.*, 1993). This discovery was consistent with the Southern hybridisation data localising the pRDH7 locus to the EAF plasmid, and with the observed loss of BFP production in the CVD206::*TnphoA* mutants #55 and #138.

The subsequent reports of the DNA sequence of the complete *bfp* operon from both E2348-69 (Stone *et al.*, 1996) and B171-8 (O111:NM; Sohel *et al.*, 1996) confirmed that the predicted *orf1-orf5* open reading frames form the proximal genes of a 14 gene cluster which is sufficient to encode the biogenesis of BFP. As predicted the gene *orf2* (*bfpB*), which contains the transposon insertion sites in CVD206::*TnphoA* mutants #55 and #138, has been identified as an outer membrane lipoprotein which is essential for BFP biogenesis (Ramer *et al.*, 1996). Furthermore the demonstration of the production of an mRNA transcript which includes (at least) *bfpA*, *bfpG* (*orf1*), *bfpB* (*orf2*) and *bfpC* (*orf3*) (Ramer *et al.*, 1996) is consistent with the *bfp* genes forming an operon and

therefore explains the failure of the incomplete operon on plasmid pRDH8 to complement the *TnphoA* mutants #55 and #138.

4.3.2 Analysis of the genes encoded on plasmid pRDH8.

The *TnphoA* mapping and DNA sequencing of plasmid pRDH8 identified seven open reading frames of which three were predicted to encode extracytoplasmic proteins. The first four open reading frames, *emtA*, *tag*, *orf3* and *orf4*, were identical to sequences determined to be present at the 26.7 min region in the *E. coli* K-12 strains MG1655 and W3110 (Oshima *et al.*, 1996; Blattner *et al.*, 1997). This was consistent with the previously observed homology of pRDH8 sequences with the *umudca* DNA sequence from *E. coli* K-12 strain AB1157 which had been mapped near to 25.5 min (Sedgwick *et al.*, 1988). The remaining three open reading frames, *orf1*, *orf2* and *ompX*, which were predicted to constitute an operon, were not homologous to sequences in *E. coli* K-12 strain MG1655 and were also shown to be absent from a number of pathogenic *E. coli* strains when analysed by Southern hybridisation. The possible functions of the gene products of some of these open reading frames and a rationale for the observed distribution of the *orf1-orf2-ompX* operon amongst *E. coli* strains are discussed further below.

4.3.2.1 A possible role for the lytic transglucosylase EmtA in transenvelope transport.

The transposon insertion sites in the host cell phosphorylation deficient CVD206::*TnphoA* mutants #13/#137 and #181 were determined to be within the open reading frames *emtA* and *tag*, respectively, both of which have corresponding genes in the *E. coli* K-12 chromosome. The secondary phenotypes of the *TnphoA* mutants #13/#137 and #181 were almost identical to those of the *bfpB* mutants #55 and #138, i.e. they failed to produce BFP and showed very low levels of adherence and invasion with HEp-2 cells. In common with *BfpB* mutants the #13/#137 and #181 mutants

showed normal levels of bundlin expression thus indicating that the mutations had a direct effect upon pili biogenesis rather than *bfpA* gene regulation. Despite the similarity in their mutant phenotypes and the close proximity of their open reading frames it seems unlikely that the *tag* and *emtA* genes are co-transcribed. *emtA* is separated from *tag* by a gap of 950 bp which is predicted to encode an open reading frame, *orf3*, transcribed in the opposite direction; furthermore analysis of the DNA sequence immediately downstream of *emtA* identified a consensus Rho-independent terminator sequence. As it seems most improbable that there will be read through from the *emtA* promoter to *tag*, it must be presumed that the 200 bp gap between the divergent *orf3* and *tag* genes encodes promoter sequences for both genes. As the genes are apparently transcribed independently it seems odd that neither the *emtA::TnphoA* mutants #13 and #137 nor the *tag::TnphoA* mutant #181 were complemented by the introduction their wildtype alleles on the pRDH8 plasmid.

Unlike the mutations in *bfpB* it was not immediately obvious why mutations in *emtA* and *tag* should have an effect on pili production. The *tag* gene is predicted to encode a small integral inner membrane protein with three transmembrane spanning domains and short cytoplasmic and periplasmic loops. This polypeptide shares significant sequence and structural homology with the deduced polypeptide product of an open reading frame, f82, predicted from the *E. coli* K-12 genome sequence (Blattner *et al.*, 1997); however, as neither protein has any other homologues it is currently impossible to speculate why a *tag* mutant should fail to produce BFP. In contrast the outer membrane lipoprotein product of the *emtA* gene had significant homology with the C-terminal catalytic domain of a characterised *E. coli* protein, MltC (Dijkstra and Keck, 1996a). MltC, also an outer membrane lipoprotein, is a member of the lytic transglycosylase family of peptidoglycan hydrolysing enzymes (Höltje *et al.*, 1975; Höltje and Tuomanen, 1991) which are involved in the metabolism of the *E. coli* cell wall (Höltje, 1996a; Höltje, 1996b; Höltje, 1998). EmtA also contains the three conserved amino acid motifs, characteristic of *E. coli* lytic transglycosylases, which are predicted from the crystal structure of the soluble lytic transglycosylase (Slf70; Engel

et al., 1991) to be present in the enzyme active site (Rozeboom *et al.*, 1991; Thunnisen *et al.*, 1994; Dijkstra and Thunnisen, 1994). The recent characterisation by Kraft *et al.* (1998) of an identical gene from *E. coli* K-12 has demonstrated that *emtA* does indeed encode a membrane-bound lytic transglycosylase activity; however, unlike the previously characterised membrane-bound lytic transglycosylases, MltA, MltB and MltC (Ursinus and Höltje, 1994; Dijkstra *et al.*, 1995; Dijkstra and Keck, 1996a), the activity of EmtA is endo-specific, i.e. the enzyme cleaves within glycan strands (Kraft *et al.*, 1998). Because of this novel enzymatic activity it has been suggested that EmtA may be responsible for the formatting of new murein strands to size prior to insertion into the growing cell wall.

The cell wall of *E. coli* is a complex macromolecule which is synthesised in the periplasmic space; it is composed of a network of short repeating oligosaccharide strands cross-linked by peptides which form a bag shaped murein molecule, or sacculus, that completely encompasses the cell (Rogers *et al.*, 1980; Neidhardt *et al.*, 1990). In Gram-negative bacteria the murein sacculus has been shown to be present as a single peptidoglycan layer for 75 to 80% of its structure, but also to exist as a triple layer for between 25 and 30% (Labischinski *et al.*, 1979). The cell wall has previously been regarded solely as a rigid exoskeleton whose function was to protect the integrity of the cell against detrimental environmental influences and it had been assumed that it plays only a minor role in cell permeability compared to that of the Gram-negative outer membrane (Nakae and Nikaido, 1975). In this respect the role and even the very presence of the cell wall has been frequently ignored by researchers working in fields involved in the assembly of trans-envelope structures such as flagella (Macnab, 1992) or protein secretion apparatus (Pugsley, 1993; Thanabalu *et al.*, 1998). Recent work concerning the permeability of the peptidoglycan sacculus has caused some researchers to review the role of the murein cell wall in trans-envelope transfer systems. Koch's model of the structure of single-layer peptidoglycan proposes a fundamental structural unit where neighbouring oligosaccharide chains are crosslinked by peptide chains at intervals of 4 disaccharides (Koch and Woeste, 1992). From this model it has been

calculated that the *E. coli* cell wall would contain holes with a maximum diameter of 5 nm; in theory such holes would allow the passage of globular proteins of no greater than 55 kDa (Dijkstra and Keck, 1996b). This figure is consistent with the 50 kDa cut-off point found by Demchick and Koch (1996) using fluorescein-labelled dextrans to assess the permeability of isolated *E. coli* sacculi. The dimensions of the components of many trans-envelope structures, such as flagella (diameter 13 nm; Stallmeyer *et al.*, 1989), pili (type IV pili diameter 6 nm; Folkhard *et al.*, 1984), and protein secretion systems, are all predicted to exceed the maximum size of the holes available in the cell wall. Dijkstra and Keck (1996b) proposed therefore that one of the common aspects of the assembly of trans-envelope transfer complexes must be the localised rearrangement of the peptidoglycan; this assembly process would be predicted to require the action of specific peptidoglycan metabolising enzymes such as peptidoglycan hydrolases.

Despite the heterologous nature of the different trans-envelope transport systems it is becoming increasingly apparent that they have common aspects which are reflected in the similarity of some of their protein components. Thus the operons required for diverse processes such as the assembly of type IV pili, DNA transfer, filamentous phage assembly and protein-secretion systems all contain groups of homologous genes (Hobbs and Mattick, 1993; Russel, 1994). In support of their theory Dijkstra and Keck (1996b) have identified putative lytic transglycosylase homologues that exist in the operons of several trans-envelope transfer systems (Dijkstra and Keck, 1996b); e.g. Protein 19 (P19), which is encoded adjacent to the origin of transfer of plasmid R1 (Graus *et al.*, 1990), is required for efficient conjugative transfer of plasmid DNA (Bayer *et al.*, 1995). P19 and its protein homologues, Orf169 (F plasmid; Loh *et al.*, 1989) and gene x (plasmid R100; Fee and Dempsey, 1988), were all identified to contain the three amino acid motifs indicative of the lytic transglycosylase active site (Dijkstra and Thunnisen, 1994). In addition transglycosylase active site motifs were found in the *Shigella flexerini* IpgF protein, which is encoded by a gene adjacent to the type III secretion operon involved in secretion of the Ipa proteins (Allaoui *et al.*, 1993),

and its homologue IagB in *Salmonella typhi*, which is essential for the protein secretion-dependent host cell invasion process (Miras *et al.*, 1995).

The observation that an *emtA::TnpHoA* mutant can produce mature bundlin subunits but is unable to make pili indicates that the mutation specifically affects BFP assembly. Considering the strong circumstantial evidence for the involvement of peptidoglycan hydrolases in rearrangements of the cell wall during the assembly of trans-envelope structures it seems reasonable to speculate that the EmtA enzyme is required to facilitate these rearrangements for the assembly of BFP. Chapter 5 describes the construction and phenotypic analysis of a defined non-polar mutation in the *emtA* in order to assess its role in BFP biogenesis.

4.3.2.2 A novel siderophore uptake system in *E. coli*.

From analysis of the DNA sequence of plasmid pRDH8 it was predicted that the open reading frames *orf1'*, *orf2* and *ompX* were organised as a operon. As the termination codon of *orf1'* overlaps the initiation codon of *orf2* and the 25 bp gap between *orf2* and *ompX* does not contain consensus *E. coli* promoter sequences it was presumed that all three genes are co-transcribed from a promoter upstream of *orf1'*. *ompX* appears to be the distal gene as it is immediately followed by two inverted 13 bp repeat sequences which are predicted to form a stable stem-loop structure indicative of a transcriptional terminator (Miller, 1992).

The *ompX* gene is predicted to encode a protein which shares a significant degree of similarity with a large number of TonB-dependent outer membrane receptor proteins from a diverse group of Gram-negative bacterial species. This similarity is particularly high in the so-called "TonB box", Box II and Box III regions, which have previously been identified in all of this class of receptors (Bäumler and Hantke, 1992), and in the C-terminal consensus region which is present in various outer membrane proteins (Struyve *et al.*, 1991; Koebnik, 1993). As with the majority of TonB-dependent outer

membrane receptor proteins studied to date (van der Helm, 1998) the receptors in *E. coli* (excluding the vitamin B12 receptor BtuB; Heller *et al.*, 1985) are involved in iron acquisition. *E. coli* has approximately five high affinity iron uptake systems (Earhart, 1996). This level of multiplicity may seem excessive; however, it merely demonstrates the extent to which iron is essential for the growth of *E. coli*. The necessity for iron is due to the role it plays in the synthesis of a small but very important group of iron containing proteins, which are involved in a variety of essential cellular processes; these include: the haem-containing proteins, such as cytochromes, cytochrome oxidases, catalase and peroxidase; iron sulphur proteins, such as glutamate synthase, succinate dehydrogenase and ferridoxin; and other non-haem, non-iron-sulphur proteins, such as ribonucleotide reductase or superoxidase dismutase (Wooldridge and Williams, 1993; Earhart, 1996). Despite its relative abundance in the natural environment the actual availability of iron is frequently low; this is because at neutral pH, under aerobic conditions, iron rapidly converts from the ferrous (Fe^{2+}) to ferric (Fe^{3+}) form which then readily forms insoluble, and therefore inaccessible, oxy-hydroxy compounds. Furthermore, when bacteria are in an animal host the availability of extracellular iron is further reduced by the action of the specific host iron binding proteins transferrin and lactoferrin which are found in the serum and in the mucosal secretions, respectively (Weinberg, 1984; Bezkorovainy, 1987; Brock 1989; Weinberg, 1989). Consequently mechanisms for the acquisition of iron are a vital part of the virulence strategy of all pathogenic bacteria.

The five well characterised high affinity iron uptake systems found in *E. coli* are all involved in the uptake of siderophores; these are small molecules (typically 500-1000 Da) which have a high affinity for Fe^{3+} ($K_{\text{diff}} > 10^{30}$) (Neilands, 1981a; Neilands, 1981b; Matzanke, 1991). All wildtype *E. coli* strains synthesise the catecholate siderophore enterobactin (O'Brien *et al.*, 1970; Rogers, 1973; Rogers *et al.*, 1977) and some also produce the hydroxamate siderophore aerobactin (Gibson and Magrath, 1969) which are taken up by the outer membrane receptors FepA and IutA respectively (Lundrigan and Kadner, 1986; Krone *et al.*, 1985). In addition *E. coli* can use a number

of hydroxamate siderophores produced by fungi such as ferrichrome, coprogen and rhodotorulic acid via the FhuA and FhuE outer membrane receptors, respectively (Coulton *et al.*, 1986; Hantke, 1983). The fifth high affinity uptake system involves the outer membrane receptor FecA which takes up the siderophore molecule ferridictrate (Wagegg and Braun, 1981). However, OmpX shows greatest similarity to a poorly characterised *E. coli* TonB-dependent outer membrane receptor, CirA (Nau and Konisky, 1989); this protein does not have a known siderophore but, in common with the *E. coli* outer membrane receptor Fiu, has been shown to allow the uptake of iron complexes of the enterobactin-related compounds dihydroxybenzoic acid (DHB) or dihydroxy-benzoylserine (DBS) (Hantke, 1990). OmpX shows even greater similarity to outer membrane receptor proteins from *Pseudomonas aeruginosa* and *Vibrio cholerae*; however, a strict homologue has not been identified in any bacterial species. Furthermore, a phylogentic comparison of OmpX with the known *E. coli* TonB-dependent outer membrane receptor proteins proved unable to classify the receptor into any of the four outer membrane receptor subfamilies proposed by Rakin *et al.* (1994) and therefore gave no indication as to the nature of siderophore it might bind.

Each of the *E. coli* high affinity iron uptake systems comprises of a number of envelope proteins in addition to the outer membrane receptor, some of which are common to all of the different systems (Figure 4.21). Ferrisiderophores will bind to their receptors irrespective of the functional state of the uptake system; however, the subsequent transfer of the ferrisiderophore across the outer membrane requires both the TonB protein (Postle and Good, 1983) and the accessory proteins ExbB and ExbD (Eick-Helmerich and Braun, 1989). It is thought that the TonB-ExbB-ExbD complex transfers energy, supplied by the inner membrane proton-motive force, to the outer membrane receptor protein causing a conformational change which opens a channel through which the siderophore enters the periplasm (Braun, 1995; Earhart, 1996; Letain and Postle, 1997; Moeck and Coulton, 1998). Periplasmic ferrisiderophore is bound by a so-called binding protein and is taken up into the cytoplasm (or in the case of ferridictrate just the iron atom is transferred; Braun, 1995) by an inner membrane

transport system similar to those of the ABC (ATP binding cassette) transporter superfamily (Higgins *et al.*, 1990; Higgins, 1992). There is generally a lower degree of specificity in this transfer process compared to uptake of ferrisiderophore across the outer membrane, i.e. *E. coli* has just three inner membrane transporters and their respective binding proteins (Earhart, 1996). All hydroxamate ferrisiderophores are bound by FhuD, ferrienterobactin and other catecholate derivatives probably all bind to FepB, and ferridictrate binds to FecB (Lundigran and Earhart, 1981; Braun and Hantke, 1991; Braun, 1995). The three *E. coli* cytoplasmic membrane permeases each consist of two hydrophobic membrane protein domains (FepDG, FecCD or the double size FhuB), which are presumed to form the transporter channel, and two molecules of an ATPase (FepC, FecE, or FhuC; Staudenmaier *et al.*, 1989; Koster, 1991; Shea and McIntosh, 1991) which are thought to provide the energy required for the transfer (Ames and Joshi, 1990; Earhart, 1996; **Figure 4.21**).

The preliminary DNA sequencing of plasmid pRDH45 identified the 3' end of a gene, designated *orf5'*, which is contiguous with the *orf1-orf2-ompX* operon. The deduced protein product of *orf5'* shows significant similarity to that of the *Haemophilus influenzae* predicted open reading frame *yc72* (Fleischmann *et al.*, 1995). *yc72* encodes an uncharacterised FepC homologue indicating that *orf5* probably also encodes the ATPase component of a cytoplasmic membrane uptake system. In *E. coli* all of the genes encoding the inner membrane permeases and periplasmic binding proteins for the high affinity iron uptake systems are arranged as operons with the genes for their respective outer membrane receptors; in the case of enterobactin the operon also contains the siderophore biosynthesis genes (**Figure 4.22**; Earhart, 1996; Braun *et al.*, 1998). The existence of the novel *fepC* homologue *orf5* strongly suggests that the *ompX* operon may also encode the genes for the periplasmic binding protein and the hydrophobic membrane protein(s) of a novel cytoplasmic membrane permease. The open reading frames *orf1* and *orf2* are predicted to encode cytoplasmic proteins which have no homology with gene products of any of the characterised *E. coli* siderophore uptake systems. However, the genes encoding their closest homologues, the

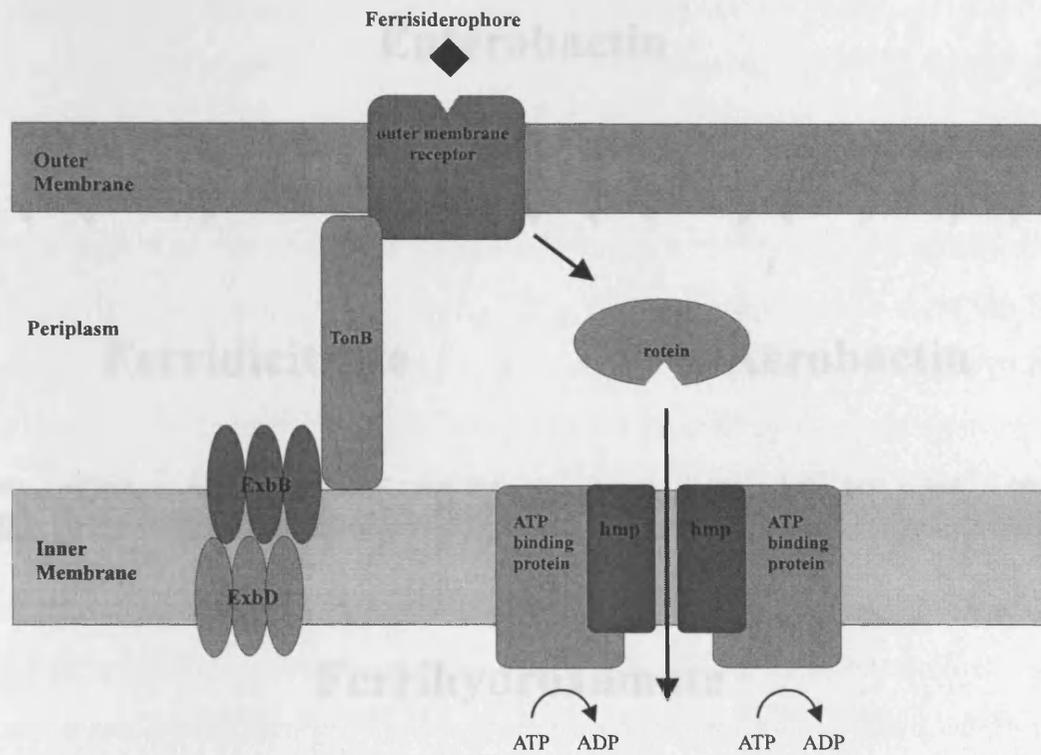


Figure 4.21 General model of TonB-dependent ferrisiderophore uptake in *Escherichia coli*. *E. coli* is able to utilise a wide range of both endogenous and exogenous siderophores via a series of specific high efficiency uptake systems. The ferrisiderophore is bound by a specific outer membrane receptor (e.g. the enterobactin receptor FepA) and is then transported across the outer membrane in an energy dependent manner requiring the the function of TonB and the inner membrane protein complex ExbB/ExbD. A periplasmic binding protein (e.g. FepB) then presents the ferrisiderophore to an inner membrane ABC-like transport system, which usually consists of two copies of an ATP-binding protein (e.g. FepC) and two hydrophobic membrane proteins (hmp; e.g. FepD and FepG), which transports it to the cytoplasm (for a review see Earhart, 1996).

Figure 4.21 The major *E. coli* gene clusters involved in high affinity iron uptake.
(Adapted from Earhart, 1996)

Enterobactin



Ferridictrate



Aerobactin



Ferrihydroxamate

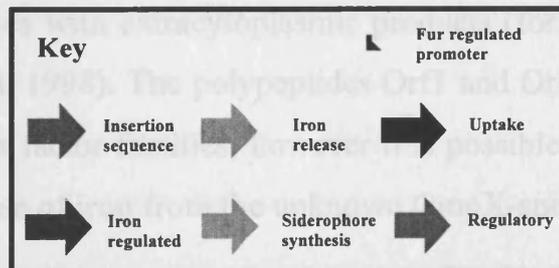


Figure 4.22 The major *E. coli* gene clusters involved in high affinity iron uptake.

(Adapted from Earhart, 1996).

uncharacterised *H. influenzae* proteins *yc73* and ModD (Fleischmann *et al.*, 1995), are both located within operons which consist solely of genes encoding homologues of inner membrane iron uptake proteins. The Orf2 protein also shows homology with ModD, a cytoplasmic protein from *Rhodobacter capsulatus*, which has been shown to induce a four-fold enhancement in the uptake of molybdenum by a specific inner membrane uptake system (Wang *et al.*, 1993). It is currently unclear what the function of the gene products of *orf1* and *orf2* might be but considering their location within the *ompX* operon, and the association of the genes for their *H. influenzae* homologues with inner membrane siderophore permease components, it seems likely that they are also somehow involved in siderophore uptake. There are only two *E. coli* cytoplasmic proteins which have been identified to be involved with iron uptake systems; these are the products of the genes *fes* and *fecI* (Earhart, 1996). The *fes* gene encodes the enzyme enterobactin esterase which is required for release of the iron atom from cytoplasmic ferrienterobactin; Fes hydrolyses the ester bonds of ferrienterobactin yielding three DBS molecules (Greenwood and Luke, 1978; Brickman and McIntosh, 1992). The *fecI* gene product is essential for the positive regulation of the ferridictrate uptake system and is related to a new subfamily of σ^E -like proteins which are involved in expression of genes with extracytoplasmic products (for a review see Crosa, 1997; Missiakas and Raina, 1998). The polypeptides Orf1 and Orf2 show no homology with the well conserved σ factor families; however it is possible that they may be enzymes involved in the release of iron from the unknown OmpX-specific siderophore.

4.3.2.3 The distribution of the *orf1-orf2-ompX* operon amongst pathogenic *E. coli*.

Comparison of the pRDH8 DNA sequence with that determined for the *E. coli* K-12 strain MG1655 (Blattner *et al.*, 1997) demonstrated that the *orf1-orf2-ompX* operon was not present at the 26-27 min region in the K-12 strain, nor was it encoded elsewhere in the chromosome. This was consistent with Southern blot analysis which failed to detect bands hybridising to an *ompX*-specific probe in *E. coli* K-12 strain DH5 α or K-12/B strain HB101. Detailed analysis of the DNA sequence of the *E. coli*

K-12 26.8 min region directly downstream of the open reading frame o84 (*tag*) identified 24 bp of sequence which were 100% identical with the 3' coding sequence of *ompX*. Because of this homology it was proposed that the absence of the *orf1-orf2-ompX* operon in *E. coli* K-12 was due to a chromosomal deletion of undetermined size. Preliminary DNA sequencing of a large *ompX*-containing *EcoRI* chromosomal DNA fragment, cloned from EPEC strain CVD206::Tn*phoA* #181 into pBluescript (pRDH45), allowed the identification of the *E. coli* K-12 open reading frame *ychM* in the region upstream of *ompX* in E2348-69. *ychM* has been mapped to 27.15 min (Rudd, 1998; GenBank Accession number AE00219; Blattner *et al.*, 1997) in the K-12 genome and is approximately 14 kb downstream of the o84 open reading frame (o84 = *tag*). Using the known *E. coli* K-12 genomic DNA sequence and the restriction enzyme map of plasmid pRDH45 it was calculated that the *ompX*-containing locus which is absent from *E. coli* K-12 is approximately 10 kb in size (of which 7 kb is as yet unsequenced).

A recent report (Guyer *et al.*, 1998) has suggested that the *orf1-orf2-ompX* operon is in fact part of a pathogenicity island (Pai) previously identified by the research group of H. Mobley at the University of Maryland; this group reported a novel 50 kb Pai which was inserted in the chromosome of the uropathogenic *E. coli* strain CFT703 (Kao *et al.*, 1997). The Pai was cloned as three overlapping cosmids containing partial *Sau3A*-digested genomic DNA fragments and found to contain the *hly* operon and one of the two copies of the *pap* genes which had previously been reported in CFT703 (Mobley *et al.*, 1993). In contrast with the other Pairs identified in uropathogenic *E. coli* strains (Blum *et al.*, 1994; Swenson *et al.*, 1996), the termini of this island were not found to be inserted within a tRNA gene. Furthermore, the sequence identified at the left hand terminus was the *dadX* gene (26.75 min in *E. coli* K-12) whilst the sequence next to the right hand terminus was identified as a region upstream of the *argA* gene (62.8 min in *E. coli* K-12). It was proposed that the apparent discrepancies in the locations of the Pai termini were due to rearrangements of the CFT703 chromosome (Kao *et al.*, 1997). In a recent abstract by Guyer *et al.* (1998) it was reported that the complete sequence

of the 50 kb Pai had been determined and that homologues to *fepA* (*ompX*), *H. influenzae* *yc73* (*orf1*), *modD* (*orf2*) and *fepC* (*orf5'*) had been identified within the left hand end of the island. Comparison of the terminal DNA sequences of the Pai from CFT703 (Genbank accession numbers AF003741 and AF003742; Kao *et al.*, 1997) with the recently released *E. coli* K-12 genome demonstrated that the proposed lefthand terminal sequence was in fact *E. coli* K-12 chromosomal sequence. The error in the assignment of the left hand terminal junction made by Kao *et al.* (1997) was apparently due to a deletion of approximately 400 bp immediately downstream of the *dadX* gene in the CFT703 genome. Additional data obtained from the University of Maryland group (H. Mobley, personal communication) confirmed that the DNA sequence of the left hand terminus of the CFT703 Pai contained approximately 5 kb of *E. coli* K-12 DNA sequence. In consideration of these new data the University of Maryland group has proposed a new site for the left hand terminal sequence of the CFT703 pathogenicity island (H. Mobley, personal communication); this site is now identical to the basepair with the site at which it has been proposed above that the *orf1-orf2-ompX* operon is deleted from *E. coli* K-12; i.e. 24 bp within the 3' end of *ompX*.

Hacker *et al.* (1997) proposed a set of criteria by which pathogenicity islands could be defined, including: (i) occupation of large chromosomal regions (often greater than 30 kb), (ii) carriage of virulence genes, (iii) a different GC content in comparison to the host DNA, (iv) association with tRNA genes or insertion sequences (IS), and (v) presence in pathogenic strains but absence or sporadic distribution in less pathogenic strains. The Pai of CFT703 fits some of these criteria well; i.e. it is proposed to be 50 kb, it carries the *hly* and *pap* operons, the overall GC content is 42.9% (compared to 51% for *E. coli* K-12), it is more prevalent in pyelonephritic *E. coli* strains than less virulent uropathogenic strains, and the right hand terminus is inserted close to the tRNA gene *metV* (Kao *et al.*, 1997; Guyer *et al.*, 1998). In contrast the identification of the *orf1-orf2-ompX* operon as the left hand terminus of the CFT703 Pai is much less convincing; the 26.8 min region of the *E. coli* chromosome contains no tRNA genes, and the GC content of the 3 kb *orf1-orf2-ompX* DNA sequence from E2348-69 is

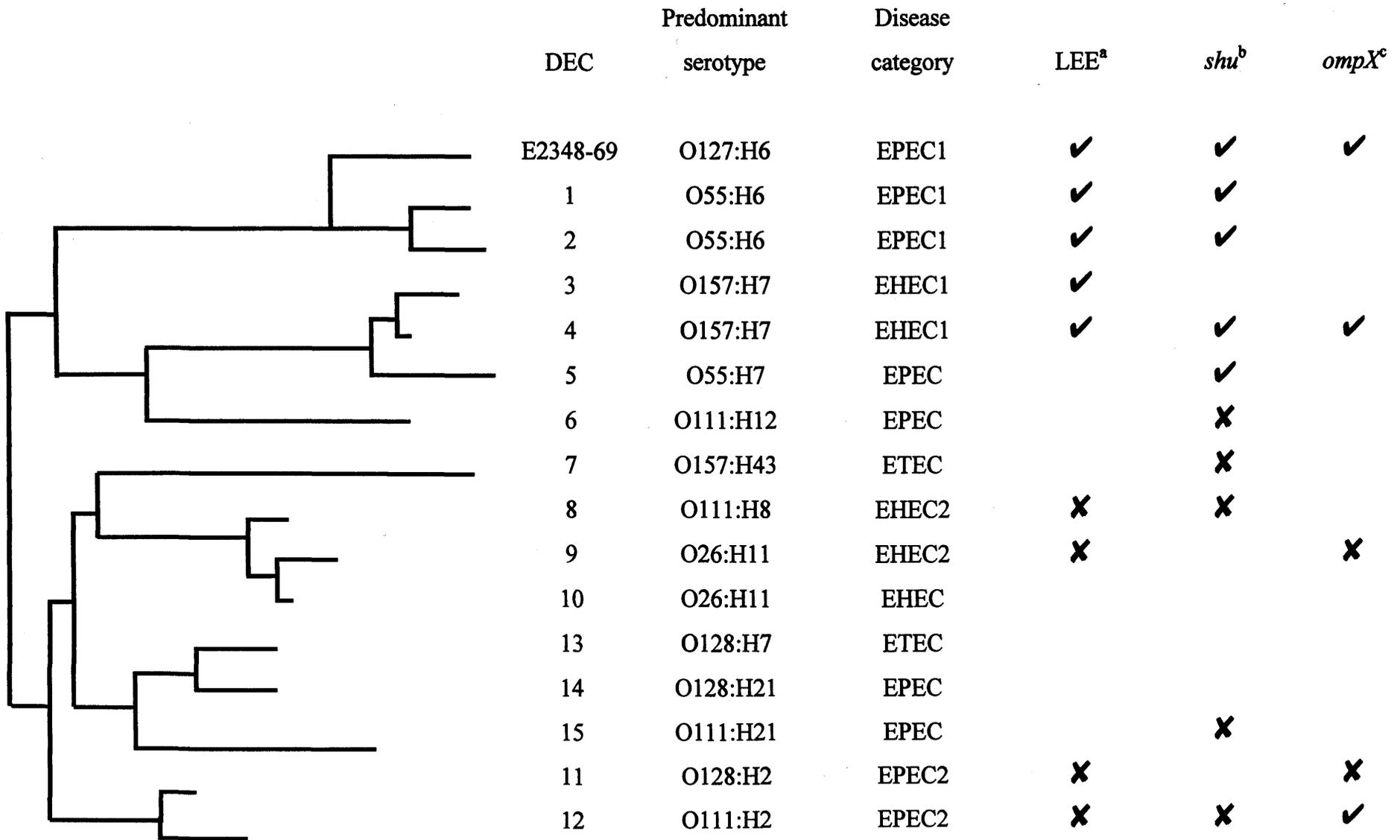
51.1%, almost identical to *E. coli* K-12. Furthermore the DNA sequence obtained from the recombinant plasmid pRDH45 indicates that in strain E2348-69 the *orf1-orf2-ompX* operon is located at 26.8 min but it cannot be part of a similar pathogenicity island. I propose that the inclusion of the *orf1-orf2-ompX* sequences at the left hand terminus of the CFT703 Pai is due to the inadvertent cloning of two independent partially digested *Sau3A* chromosomal fragments into the original cosmid clone (8-3f) identified by Kao *et al.* (1997). This theory is also consistent with the observation that the putative iron transport system (i.e. *orf5-orf1-orf2-ompX*) reported by Guyer *et al.* (1998) at the left hand terminal appears incomplete. Whilst their DNA sequence contains a *fepC* homologue (*orf5*), Guyer *et al.* (1998) do not report further upstream genes that might encode the inner membrane hydrophobic proteins (*fepD* and *fepG* homologues) and the periplasmic binding protein (*fepB* homologue) which would be expected in an *E. coli* high affinity iron transport system (Earhart, 1996; **Figure 4.22**). Confirmation of this hypothesis should be available when the DNA sequence of the region upstream of *orf5'* in pRDH45 is determined or when the complete sequence of the CFT703 Pai is released.

Analysis of the distribution of the E2348-69 *ompX* locus amongst a limited group of representative *E. coli* strains found that whilst it was present in both diarrhoeagenic and uropathogenic strains it was not present in all pathogenic strains. Amongst the diarrhoeagenic biotypes tested *ompX* was found in the DAEC, the EHEC O157:H7 and 64% of EPEC strains, but was absent in EHEC O26:H11, ETEC, EIEC and EAaggEC strains; however, as only one strain of each of the non-EPEC biotypes was tested the wider significance of these results is unknown. Whittam *et al.* (1993) have used electrophoretic typing of polymorphic enzyme loci (Selander *et al.*, 1986; Ørskov *et al.*, 1990) to phylogenetically align 1,300 *E. coli* strains known to cause haemorrhagic colitis or infantile diarrhoea; from this data they were able to identify 15 major diarrhoeagenic *E. coli* clones (DEC) which can be divided into two evolutionary subgroups, DEC1-6 and DEC7-15 (**Figure 4.23**). Furthermore Wieler *et al.* (1997) have identified four disease clusters amongst the DEC strains (Whittam *et al.*, 1993)

based upon the distribution of two different insertion sites for the LEE pathogenicity island; in the EPEC1 and EHEC1 strain clusters the LEE is inserted at the *selC* locus (McDaniel *et al.*, 1995) whilst in EPEC2 and EHEC2 it is inserted elsewhere (**Figure 4.23**). Despite the limited number of strains tested in this study that fall within the classic EPEC and EHEC serogroups represented in the DEC strains it was immediately apparent that the distribution of the *ompX* locus did not seem to follow the accepted clonal phylogeny. The EPEC1 E2348-69 strain and EHEC1 O157:H7 are closely related (DEC1-6 subgroup) and unsurprisingly are both *ompX*⁺; however only one of the group of closely related strains, EHEC2 O26:H11, EPEC2 O128:H2 and EPEC2 O111:H2 (DEC7-15 subgroup) is *ompX*⁺ i.e. EPEC2 O111:H2 (**Figure 4.23**). This distribution pattern cannot be simply explained by descent from a common *ompX*⁺ *E. coli* progenitor strain. Either the acquisition of the *ompX* locus has occurred separately in the major DEC subgroups or the genes were present in an ancestor of all of the strains and have recently been lost from the EHEC2 O26:H11 and EPEC2 O128:H2 strains.

The distribution of the *ompX* locus amongst the DEC strains appears to have some similarities with that of the recently described *Shigella dysenteriae* *shu* locus (Wyckoff *et al.*, 1998); this locus encodes a haem transport operon which includes the outer membrane receptor gene *shuA* (also known as *chuA* in *E. coli* O157:H7; Torres and Payne, 1997) and seven other genes (Mills and Payne, 1995; Mills and Payne, 1997). Wyckoff *et al.* (1998) have shown that the 10 kb *shu* locus is inserted into a common chromosomal site, equivalent to 78.7 min in *E. coli* K-12, in the *S. dysenteriae* serotype 1 strains, *E. coli* O157:H7 strains, EPEC strain E2348-69 and a number of other diarrhoeagenic and uropathogenic *E. coli*. The GC content of the *shu* locus is 50% (*E. coli* or *Shigella* are approximately 51% GC) indicating that either it was acquired a long time ago or it was transferred from an organism with a very similar base content. PCR and Southern hybridisation analysis of the phylogenetically aligned DEC strains (Whittam *et al.*, 1993) have demonstrated that, in common with *ompX*, the apparent inheritance of the *shu* locus is inconsistent with the known DEC clonal

Figure 4.23 The distribution of the LEE insertion site, the *shu* locus and the *ompX* locus amongst a series of phylogenetically aligned diarrhoeagenic *E. coli* clones (DEC). The phylogenetic tree in this figure is based upon the phylogenetic analysis of Whittam *et al.* (1993) and the EPEC and EHEC subgroupings upon the work by Wieler *et al.* (1997). ^a the insertion site for the LEE was found either to be in *selC* (✓) or inserted at another site (✗) (McDaniel *et al.*, 1995; Wieler *et al.*, 1997), ^b PCR identification of the *shu* locus (haem uptake) of *Shigella dysenteriae* (Wyckoff *et al.*, 1998), ^c Southern hybridisation analysis of the *ompX* locus (not using DEC strains). (✓ = present, ✗ = absent)



phylogeny. The distribution indicated either that the *shu* locus had been acquired more than once in the different *E. coli* lineages, or that it had been acquired by an ancestral *E. coli* and then subsequently deleted in some of the lineages e.g. DEC6 (Figure 4.23; Wyckoff *et al.*, 1998). Despite the similarity in the irregular distribution of the *ompX* and *shu* loci, because of the existence of discrete non-coding flanking sequences at the ends of *shu* which are not observed with *ompX*, it is unclear if the mechanisms of deletion/insertion are common.

4.3.2.4 Future work: Analysis of the frequency of deletion of the *ompX* locus.

The DNA sequence data presented within this chapter suggest that the *ompX* operon encodes an *E. coli* siderophore uptake system which currently has no homologues amongst the Gram negative bacteria. Despite reports of an *ompX*-containing pathogenicity island (Guyer *et al.*, 1998) the DNA sequence and phylogenetic data are more consistent with the hypothesis of the introduction of the *ompX* operon into the chromosome of an ancestral *E. coli* and then the subsequent deletion from *E. coli* K-12 and a number of pathogenic strains. The majority of the diarrhoeagenic *E. coli* strains tested in this study were serogroup type strains (supplied by Dr Stuart Knutton, Birmingham) which have been stored and subcultured in the laboratory for a number of years. Therefore it is unclear if the distribution of *ompX* observed in these strains represents a true picture of all pathogenic *E. coli* strains or if it is due to instability induced by growth in the laboratory.

Statistically relevant confirmation of the *ompX* distribution pattern would require a much more extensive analysis of *E. coli* strains preferably using the DEC strains of Wittam *et al.* (1993). It would also be useful to examine a group of fresh clinical isolates of all pathogenic *E. coli* to assess both the relevance of *ompX* to virulence and the hypothesis that deletion of the *ompX* operon may occur during storage or subculturing in the laboratory. The stability of the *ompX* operon in E2348-69 could also be assessed directly using the "island probing" method reported recently for

analysis of the loss of pathogenicity island sequences in *Shigella flexneri* (Rajakumar *et al.*, 1997). This method is based upon the counterselectable properties of the tetracycline resistance protein encoded by *tetAR* (Reyrat *et al.*, 1998). TetAR acts by causing alterations in the outer membrane of Gram negative bacteria such that the cell becomes tetracycline impermeable; however, in *E. coli* these alterations also render the bacteria hypersensitive to the lipophilic chelating agent fusaric acid (Bochner, 1980; Maloy and Nunn, 1981). Thus Rajakumar *et al.* (1997) were able to construct *tetAR* insertions in the *S. flexneri* she Pai using the positive tetracycline resistance phenotype, then over a number of generations select for the appearance of bacteria which had deleted the Pai using the negative fusaric acid sensitivity phenotype. The system has been shown to be sensitive enough to identify a deletion rate of 10^{-5} to 10^{-6} per generation in *Shigella* and should be easily adaptable to EPEC strain E2348-69.

Chapter 5. Construction of non-polar mutants of the *emtA* and *ompX* genes in EPEC strains.

5.1. Introduction.

5.1.1 General introduction.

The results presented in Chapter 3 strongly suggest that the loss of the ability to induce protein phosphorylation in cultured human epithelial cells observed with the CVD206::Tn*phoA* mutants is due to the inability of these strains to adhere efficiently. Whilst there is speculation as to the exact role of the bundle-forming pili (BFP) in adherence of EPEC to intestinal cells *in vivo* (Donnenberg *et al.*, 1997; Hicks *et al.*, 1998) it has been clearly demonstrated that the failure to make BFP results in massively reduced bacterial adherence to HEp-2 cells (Donnenberg *et al.*, 1992). The discovery that the mutations in the CVD206::Tn*phoA* mutants #55 and #138 were in the *bfpB* gene, which has been shown to be required for BFP biogenesis (Ramer *et al.*, 1996), was consistent with the observed decrease in adherence to HEp-2 cells and consequent failure to induce protein phosphorylation seen with these strains. It was not immediately apparent why the mutations in the *emtA* and *tag* genes (CVD206::Tn*phoA* #13/#137 and #181 respectively) should cause a similar loss of BFP production.

The *tag* gene is predicted to encode an integral inner membrane protein with no characterised homologues (section 4.2.3.3) and therefore it is currently impossible to speculate upon its cellular function. The *emtA* gene product, however, has significant homology to the *E. coli* protein MltC, a member of the lytic transglycosylase family of proteins (Höltje and Tuomanen, 1991); furthermore the EmtA protein from *E. coli* K-12 has been demonstrated to possess a novel endo-specific lytic transglycosylase activity (Kraft *et al.*, 1998). Based upon active site motif homology (Dijkstra and Thunnissen, 1994; Dijkstra and Keck, 1996a), lytic transglycosylase

homologues have been identified which are essential for the function of a number of bacterial transmembrane transport systems involved in pili production or protein export; it has been proposed that these proteins could be required for localised peptidoglycan rearrangements to facilitate the assembly of the transfer systems through the cell wall barrier (Dijkstra and Keck, 1996b). Due to the unique endospecific nature of the EmtA lytic transglycosylase activity it has been speculated that EmtA may also be involved in the assembly of transmembrane structures by the formation of discrete tightly controlled holes in the peptidoglycan (Kraft *et al.*, 1998, A. J. Dijkstra, personal communication). Recently it has been shown that two *E. coli* type IV pili genes, *pilT* (plasmid R64 thin pilus operon; Kim and Komano, 1997) and *bfpH* (*bfp* operon; Stone *et al.*, 1996; Soheli *et al.*, 1996), encode polypeptides which contain the lytic transglycosylase active site motifs identified by Dijkstra and Keck (1996a); it has therefore been proposed that these genes may encode lytic transglycosylase activity (Kim and Komano, 1997). The presence of lytic transglycosylase homologues in these type IV pili operons greatly strengthens the theory that specific peptidoglycan rearrangements are required for type IV pili biogenesis. The recent report by Anantha *et al.* (1998) that a *bfpH* mutant was wild-type for BFP production can be interpreted to mean that either the BfpH protein has no function in BFP biogenesis, which seems unlikely, or that in an *in vitro* situation its loss of function can be complemented by other similar chromosomally encoded proteins. It seemed plausible to propose that the observed loss of BFP production in the *emtA* mutants CVD206::*TnphoA* #13 and #137 was a consequence of the involvement of EmtA in the assembly of BFP through the cell wall. However, as the mutations in these strains are due to transposon insertions, it was impossible to discount that the observed phenotype could be due to polar effects upon the expression of other adjacent genes. In order to confirm that the BFP phenotype was due solely to loss of the lytic transglycosylase activity of EmtA it was required to construct a non-polar *emtA* mutation in EPEC.

5.1.2 Construction of non-polar mutations.

By simple definition a non-polar mutation must have no effect upon the expression of the surrounding genes other than by the specific loss of function of its own gene product; however what exactly constitutes a non-polar mutation is more difficult to define. As bacterial genes frequently exist as clusters or operons, the insertion of a transposon or antibiotic resistance cassette into a gene will often affect the transcription of the genes downstream. Link *et al.* (1997) demonstrated that, in *E. coli*, insertional mutations could have radically different phenotypes compared to deletions; e.g. the insertion of a kanamycin resistance cassette into the *yjbJ* gene resulted in improved growth of the mutant strain in LUB compared to parent, whilst in contrast growth of a deletion mutation was unchanged. To address the problem of interrupted transcription of downstream genes several groups have used "non-polar" antibiotic resistance gene cassettes in the disruption of genes within operons; these cassettes contain an antibiotic resistance gene which has no promoter or terminator sequences and which is followed by a consensus ribosome binding site and an initiation codon such that transcription and translation of the 3' end of the disrupted gene and the genes downstream are maintained (Menard *et al.*, 1993; Anantha *et al.*, 1998). As another approach Link *et al.* (1997) advocate the use of in-frame deletions; this involves the deletion of the majority of the gene but retention of a short sequence containing both the initiation and termination codons in-frame, in order to maintain downstream transcription and translation. Whilst data from the Bfp operon (Anantha *et al.*, 1998) have demonstrated that read-through antibiotic cassettes can be non-polar (i.e. the mutants obtained were able to be complemented solely by the mutated gene *in trans*) it is unclear if this will be similar in all cases. Insertion or deletion of large regions of DNA may affect localised DNA supercoiling (Higgins *et al.*, 1988) and therefore the transcription of supercoiling-sensitive genes. Similarly insertions and deletions may affect mRNA secondary structure; this can affect both the rate of initiation of translation and mRNA stability (Draper, 1996). Frameshift or nonsense mutations do not affect the transcription of downstream genes and should not significantly affect

mRNA stability or DNA supercoiling, as such they would seem to be the obvious choice for the construction of non-polar mutations. However, frameshift or nonsense mutations may also have polar effects upon protein expression from genes within operons due to translational coupling (for review see Draper, 1996). Translational coupling has been proposed as a method for the production of stoichiometric levels of proteins from an operon by a requirement for the translation of upstream genes before initiation of translation can occur; e.g. the production of the *E. coli* ribosomal protein L1 is dependent upon translation of the L11 protein, which is encoded upstream, for the denaturation of mRNA secondary structures which otherwise mask the L1 ribosome binding site (Baughman and Nomura, 1983; Draper, 1996).

This chapter describes the construction of mutations in the *emtA* and *ompX* genes by the deletion or insertion of 4 bp within restriction enzyme sites in order to cause frameshifts and premature termination of translation resulting in truncated polypeptides. The rationale for why these mutations should be considered non-polar is discussed further below.

5.1.3 Introduction of non-selectable mutations into *E. coli* using counterselectable markers.

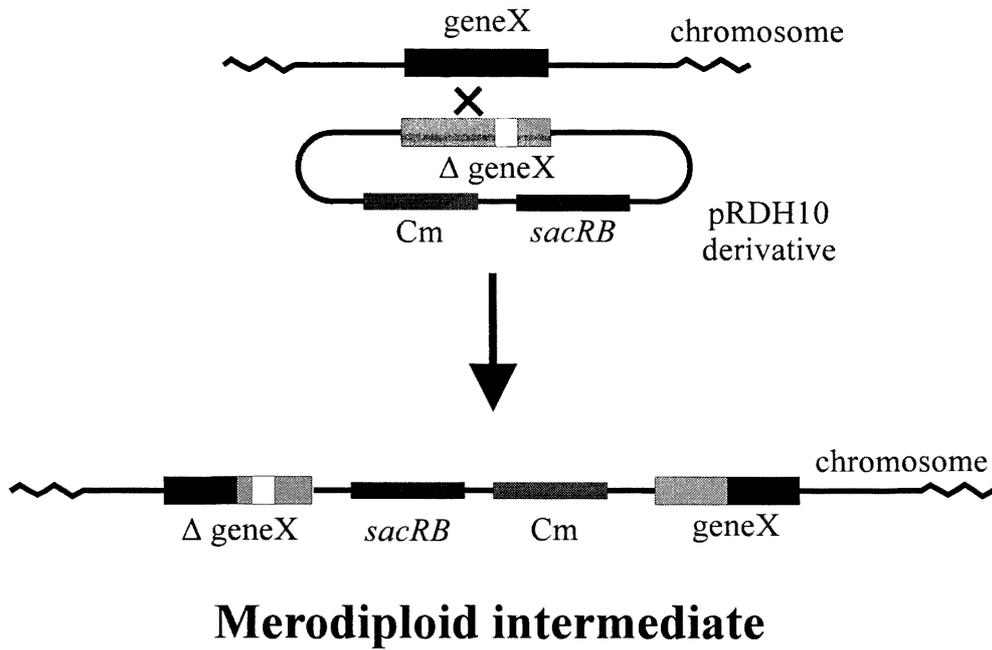
The non-polar, non-selectable mutations constructed in the *emtA* and *ompX* genes were introduced into EPEC using the *sacB* allelic exchange strategy as previously described for EPEC by Donnenberg and Kaper (1991). The procedure, summarised in **Figure 5.1**, is divided into two steps each involving homologous recombination. In the first step derivatives of suicide vector pRDH10 (**Figure 5.2**) carrying the mutant alleles are introduced into EPEC by conjugation. pRDH10 contains an *oriR6K* which is dependent upon the π protein (encoded on the lysogenic phage λ pir) for its replication; thus in EPEC the plasmid fails to replicate and is rapidly lost. Selection for the chloramphenicol resistance gene encoded on pRDH10 identifies exconjugants in which the pRDH10-derived plasmids have integrated into the chromosome at the site of the

wildtype allele, by homologous recombination, to form merodiploids (Figure 5.1A). In step two the chloramphenicol selection is removed allowing secondary recombination events, between the two homologous regions in the merodiploid, which will result in excision of the plasmid integrate; as pRDH10-derivatives are unable to replicate these plasmids will be rapidly lost. Dependent upon the site of the second recombination event the resultant bacteria can retain either the mutant or the wild-type allele in the chromosome (Figure 5.1B). The rate at which secondary recombination events occur in *E. coli* is often low, and can vary greatly depending upon the bacterial strain, the size of the homologous regions and frequently the nature of the mutated gene. It may therefore require the testing of many thousands of colonies for loss of the chloramphenicol resistance marker in order to detect a mutant (Blomfield *et al.*, 1991). A simple method for the solution of this problem has been the inclusion of counterselectable markers within allelic exchange vectors (for a recent review see Reytrat *et al.*, 1998). pRDH10 contains the counterselectable *sacRB* gene (Selbitschka *et al.*, 1992) which, in common with the wildtype *Bacillus subtilis sacB* gene, encodes the enzyme levansucrase (Gay *et al.*, 1983). Expression of levansucrase in *E. coli*, and many other Gram-negative bacteria, results in a lethal toxicity to the cell which is inducible by growth in the presence of sucrose (Gay *et al.*, 1985). Thus growth of pRDH10-containing merodiploids upon sucrose agar (Blomfield *et al.*, 1991) at 30°C can be used as a positive selection for bacteria within which secondary recombination events have occurred which result in plasmid excision. The resultant colonies will be massively enriched for resolved merodiploids which can then be screened for the presence of the non-selectable mutation by PCR or Southern blotting.

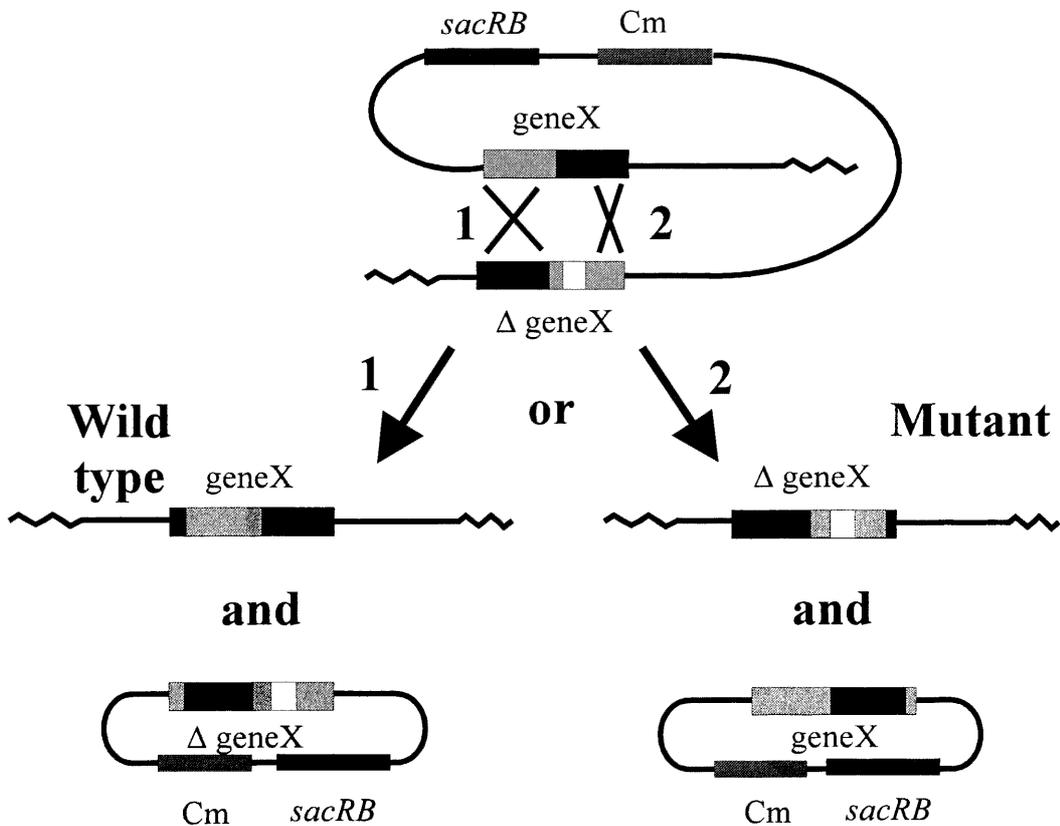
sacB is currently the most widely used counterselectable marker amongst Gram-negative bacteria (Reytrat *et al.*, 1998). In addition to its use in EPEC (Donnenberg and Kaper, 1991) it has also been used to assist in allelic exchange in a wide range of other bacterial species, including *Aeromonas salmonicida*, *Desulfovibrio vulgaris*, *Helicobacter pylori*, *Mycobacterium bovis* BCG, *Pseudomonas aeruginosa*, *Myxococcus xanthus*, *Xanthomonas* spp., *Yersinia enterocolitica*, *Klebsiella*

Figure 5.1 Summary of non-polar mutation construction using a positive-selection allelic exchange vector. Introduction of non-polar unmarked mutations into *E. coli* requires a two step process: (A) The mutated allele Δ geneX is introduced into the wildtype strain on a suicide vector (pRDH10 derivative; see **Figure 5.2**) by conjugation. Selection for the chloramphenicol resistance gene marker (Cm) of the suicide vector identifies bacteria in which a single homologous recombination event has occurred between the plasmid Δ geneX allele and the chromosomal geneX allele resulting in formation of a merodiploid. (B) With the chloramphenicol resistance selection removed random homologous recombination events occur which result in the merodiploid being resolved. The occurrence of secondary recombination events can be very low (10^{-3} - 10^{-6}); however, it is possible to positively select for bacteria in which they have occurred by growth upon 6% sucrose LUA plates. Sucrose is toxic to bacteria which are expressing the *Bacillus subtilis* levan sucrose protein, encoded by the *sacRB* gene in pRDH10, and therefore merodiploids are unable to grow. There are two possible outcomes from a second recombination event depending upon the site at which recombination occurs: (1) the wildtype allele remains in the chromosome and the mutant allele is lost as the reconstituted suicide vector is unable to replicate; (2) the wildtype allele is lost in the suicide plasmid and the mutant allele is left in the chromosome. Where the mutation in geneX is exactly central in the gene fragment cloned into pRDH10 the probability of each recombination event should be 50% (assuming that the mutation does not affect the ability of the bacteria to grow upon the 6% sucrose plates).

A



B



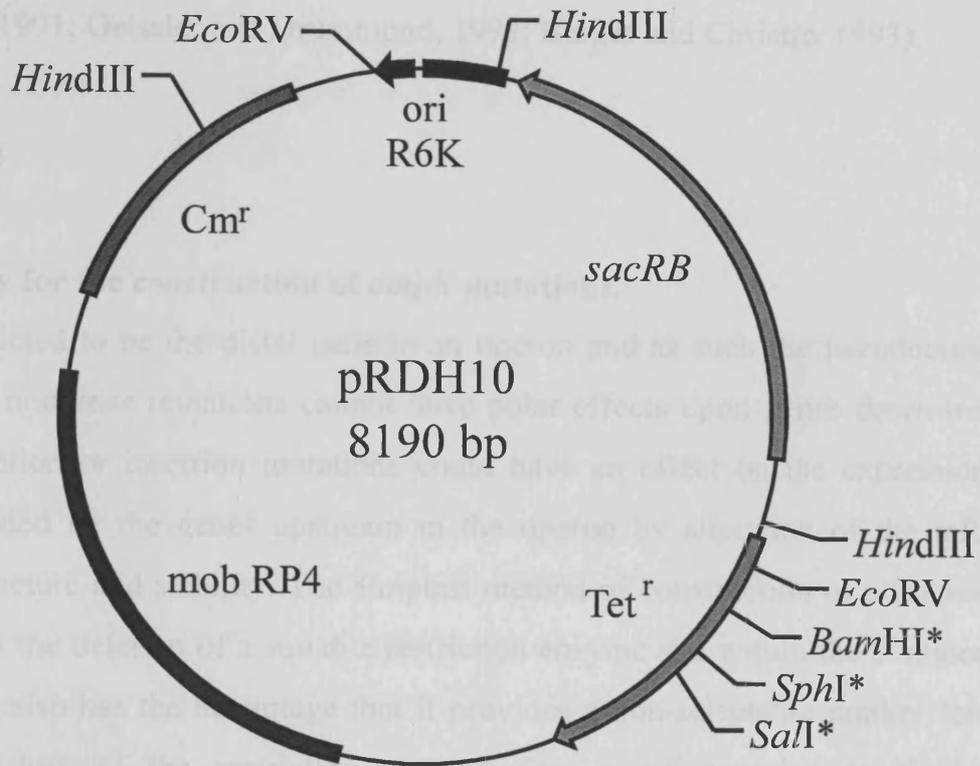


Figure 5.2 Restriction enzyme map of the positive-selection allelic exchange suicide vector pRDH10. Plasmid pRDH10 contains the *oriR6K* origin of replication, the modified *sacRB* gene (Selbichka *et al.*, 1993), a chloramphenicol resistance gene (Cm^r), a tetracycline resistance gene (Tet^r), and the mobilisation region *mobRP4*. Unique restriction enzyme sites which can be used for cloning are indicated by an asterisk (for construction details see **Appendix 3**).

pneumoniae, and *Agrobacterium tumefaciens* (Vipond *et al.*, 1998; Keon *et al.*, 1997; Azad *et al.*, 1997; Copass *et al.*, 1997; Pelicic *et al.*, 1996a; Pelicic *et al.*, 1996b; Wu and Kaiser, 1996; Schweizer and Hoang, 1995; Schweizer, 1992; Kamoun *et al.*, 1992; Kaniga *et al.*, 1991; Geissler and Drummond, 1993; Berger and Christie, 1993).

5.2 Results

5.2.1 Strategy for the construction of *ompX* mutations.

ompX is predicted to be the distal gene in an operon and as such the introduction of frameshift or nonsense mutations cannot have polar effects upon genes downstream; however, deletion or insertion mutations could have an effect on the expression of proteins encoded by the genes upstream in the operon by alteration of the mRNA secondary structure and stability. The simplest method of construction of a frameshift mutation is by the deletion of a suitable restriction enzyme site within the 5' region of the gene; this also has the advantage that it provides a non-selectable marker for the mutation i.e. loss of the restriction enzyme site. *ompX* contains an *SphI* site approximately 500 bp downstream of the initiation codon; a deletion of 4 bp from this site would result in a frameshift in the gene and the introduction of a new termination codon. This new gene sequence would now encode a truncated, and presumably inactive, OmpX' polypeptide of just 133 amino acids (after cleavage of signal sequence, 111 amino acids) (**Figure 5.3**). The cloning strategy which was adopted for the deletion of the *ompX SphI* site and the construction of a pRDH10 derivative carrying the mutant *ompX* allele are summarised in **Figure 5.4** and described below.

It was also proposed to simultaneously construct a chromosomally encoded *ompX/phoA* fusion to investigate the regulation of *ompX* expression. Plasmid pRDH8-12 contains a *phoA*-positive *TnphoA* insertion located approximately 300 bp downstream of the *ompX* start codon which encodes a PhoA fusion to the first 36 amino acids of the mature OmpX protein (Chapter 4 section 4.2.3.1, **Figure 4.7**). This plasmid was used to construct a pRDH10 derivative which contained the

ompX/TnphoA fusion sequence and also adjacent flanking sequences to be used for the directed construction of a *TnphoA* insertion in the chromosomal *ompX* gene by homologous recombination (summarised in **Figure 5.5** and detailed in section 5.2.1.4).

5.2.1.1 Subcloning of a pRDH8 *ClaI* fragment containing the 3' end of *ompX* into pBluescript to generate pRDH35. The *ompX*-carrying plasmid pRDH8 contains two *SphI* sites (**Figure 5.4**) and therefore it was impossible to directly construct the *SphI* deletion in this plasmid. Several points were taken into consideration in the design of a subcloning step; (i) the subcloned fragment could not contain the second *SphI* site from pRDH8, (ii) the subcloned fragment had to be excisable, after deletion of the unique *SphI* site, by digestion with *BamHI* and *SaII*, to allow cloning into pRDH10, and (iii) the efficiency of transfer of the mutation into *E. coli* by allelic exchange would be greatly enhanced if the *SphI* deletion was approximately central within the homologous DNA. The final strategy adopted involved the cloning of a 2.4 kb *ClaI* fragment from pRDH8 into the pBluescript vector (**Figure 5.4**). This fragment contains the majority of the *orf2* gene and the 5' end of *ompX*, and encodes EPEC chromosomal sequences flanking the *SphI* site of 1.0 kb and 1.4 kb respectively. 500 ng of pRDH8 were digested with *ClaI* and the resultant DNA fragments were separated by 1% agarose gel electrophoresis. The 2.4 kb *ClaI* fragment was isolated from the gel and purified using Polymer wool. 50 ng of purified pRDH8 *ClaI* fragment were ligated with 10 ng of *ClaI* digested pBluescript DNA, the ligation reaction was transformed into DH5 α by electroporation and transformants were selected on LUA containing ampicillin and X-gal. White colonies, indicating insertions into the pBluescript *lacZ* gene, were selected and plasmid DNA prepared using the miniprep method. Recombinant plasmids were correctly identified as pRDH35 on production of two DNA fragments, of 2.4 kb and 3.0 kb, when digested with *ClaI* and a single 5.4 kb fragment when digested with *SphI*.

5.2.1.2 Construction of a frameshift mutation in *ompX* by the deletion of the unique *SphI* site from pRDH35. 500 ng of pRDH35 DNA were linearised by

digestion with *SphI* and then incubated with Klenow to remove the 3' overhanging ends. The resultant blunt ended fragments were religated (to form pRDH36) and the ligation mixture transformed into DH5 α by electroporation. Transformants were selected on LUA containing ampicillin. Removal of 3' overhangs using Klenow is not 100% effective and the ligation of blunt ended DNA fragments is considerably less efficient than that of sticky-ended fragments, and so a large proportion of the transformants obtained were likely to carry religated pRDH35 plasmid rather than the deletion plasmid. As an initial screening step the loss of the *SphI* site was used as a positive selection in order to enrich for the small number of transformants carrying the mutation. Transformant colonies were washed from the LUA plates into LUB and these bacterial cells used to prepare plasmid DNA; this DNA was then digested with *SphI* to linearise wild-type plasmid and retransformed into DH5 α using the CaCl₂ method (this ensures a single plasmid variant in each transformant). Transformants were again selected on LUA plates containing ampicillin, and plasmid DNA was prepared from individual transformants using the miniprep method. Plasmids were digested with *SphI* or *ClaI* and the restriction fragments separated by 1% agarose gel electrophoresis. Clones with plasmids which when digested with *ClaI* produced two DNA fragments of 2.4 kb and 3.0 kb, but which would not linearise on digestion with *SphI*, were identified. Plasmid DNA was prepared from a single clone using the CsCl method and the deletion site was sequenced using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit with the sequencing primer ompxr; this primer was designed to anneal approximately 100 bp downstream (with respect to *ompX*) of the *SphI* site (section 2.12). The DNA sequence obtained confirmed that in the clone selected the *ompX* sequence had lost the central 4 base pairs of the *SphI* site as indicated in **Figure 5.3**; the plasmid was designated pRDH36 (**Figure 5.4**).

5.2.1.3 Cloning the *ompX SphI* deletion mutation into delivery vector pRDH10. 500 ng of plasmid pRDH36 DNA were digested with *BamHI* and *SalI* and the resultant fragments separated by 1% agarose gel electrophoresis. The 2.4 kb *ompX* Δ *SphI* fragment was isolated from the gel and purified using Polymer wool; 50 ng of

*Bam*HI/*Sal*I *ompX*Δ*Sph*I fragment were ligated with 10 ng of *Bam*HI/*Sal*I digested pRDH10. The ligation reaction was transformed into *E. coli* strain SY327λpir by electroporation (the pRDH10 *oriR6K* requires the presence of the π protein encoded by the lysogenic phage λpir for replication), and transformants selected on LUA plates containing chloramphenicol. Insertion of fragments into the *Bam*HI and *Sal*I sites of pRDH10 results in the inactivation of the tetracycline resistance gene; therefore transformant colonies were screened for insertions by patch-plating onto duplicate chloramphenicol and chloramphenicol/tetracycline containing LUA plates. Plasmid DNA was prepared from chloramphenicol resistant/tetracycline sensitive colonies and the correct clones were confirmed by digestion with *Bam*HI and *Sal*I to release the 2.4 kb fragment; this recombinant plasmid was designated pRDH38 (Figure 5.4).

5.2.1.4 Subcloning the *ompX/TnphoA* fragment from pRDH8-12 into pRDH10.

Due to the arrangement of restriction sites in plasmid pRDH8-12 it was impossible to clone a single DNA fragment which would contain the *TnphoA* insertion and the flanking *ompX* sequences directly into pRDH10. The final strategy adopted was to subclone the entire pRDH8 *Bam*HI insert plus the *TnphoA* insertion as a partially digested *Bam*HI fragment (Figure 5.5); this would result in *ompX* flanking sequences of 1.6 kb and 4.7 kb respectively. The asymmetry in the size of the flanking regions was not predicted to be a problem in this case as the mutants would be easy to select for due to the kanamycin resistance gene of *TnphoA*. Trial digestions of pRDH8-12 DNA were performed using varying plasmid concentrations and enzyme digestion times to find conditions which favoured production of the desired partially digested *Bam*HI fragment. 1 μg of pRDH8-12 DNA was digested with 1 U of *Bam*HI for 30 min at 30°C and the DNA fragments produced were separated by 0.8% agarose gel electrophoresis. The 14 kb partially digested *Bam*HI fragment, containing the *ompX/TnphoA*, was isolated from the gel and purified using Polymer wool. A ligation was performed with 50 ng of the 14 kb *ompX/TnphoA Bam*HI fragment and 50 ng of *Bam*HI digested pRDH10. The ligation reaction was transformed into SY327λpir using electroporation and transformants were selected on LUA plates containing

A*Sph*I

CTACGATATGCATGCTTCATTTGCCGTCAGCAGCGATCTGGGGTATTTGCAAGTCAGCGGTAGCCAGCTAAAACAGGATT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GATGCTATACGTACGAAGTAAACGGCAGTCGTCGCTAGACCCCATAAACGTTTCAGTCGCCATCGGTCGATTTTGTCCTAA

^aY D M H A S F A V S S D L G Y L Q V S G S Q L K Q D F

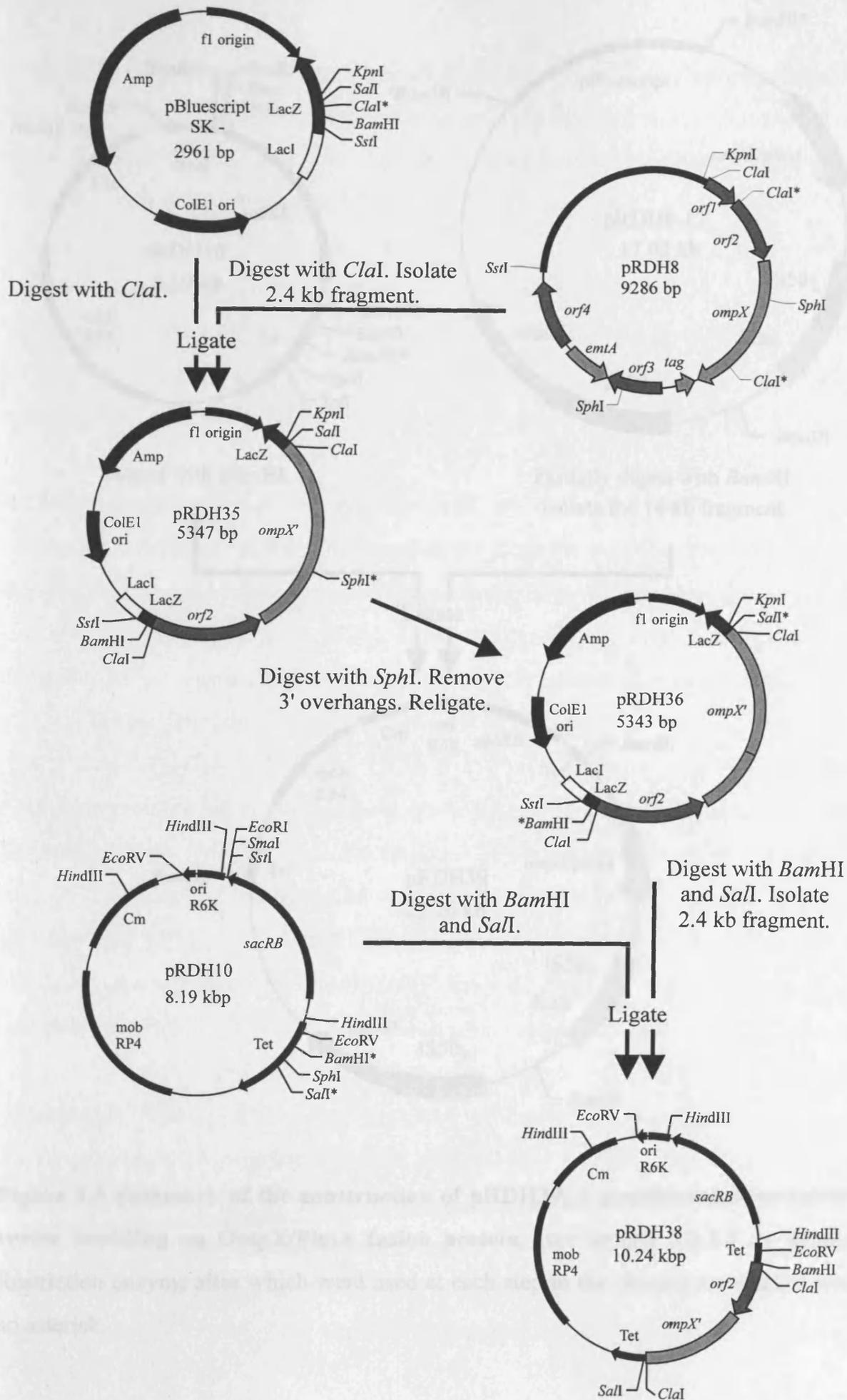
B Δ *Sph*I

CTACGATATGCTTCATTTGCCGTCAGCAGCGATCTGGGGTATTTGCAAGTCAGCGGTAGCCAGCT**TAA**AAACAGGATTTTCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+
 GATGCTATACGAAGTAAACGGCAGTCGTCGCTAGACCCCATAAACGTTTCAGTCGCCATCGGTCGATTTTGTCCTAAAAGA

^bY D M L H L P S A A I W G I C K S A V A S * N R I F S

Figure 5.3 Creation of a frameshift mutant of *ompX* by deletion of an *Sph*I site. (A) represents the partial DNA sequence of pRDH8 (from 4391-4470 bp shown in reverse); ^a indicates the amino acid sequence of the mature OmpX protein (residues 91 to 117). (B) represents the equivalent DNA sequence of pRDH36 where the *Sph*I site has been removed by the exonuclease action of Klenow and religation; the new termination codon produced is indicated in bold and underlined (see also **Figure 5.4**); ^b indicates the amino acid sequence of the truncated protein OmpX' (residues 91 to 111)

Figure 5.4 Summary of the construction of pRDH38, a positive-selection suicide vector containing an *ompX* frameshift mutation (see sections 5.2.1.1-5.2.1.3 for details). Restriction enzyme sites which were used at each step in the cloning are marked with an asterisk.



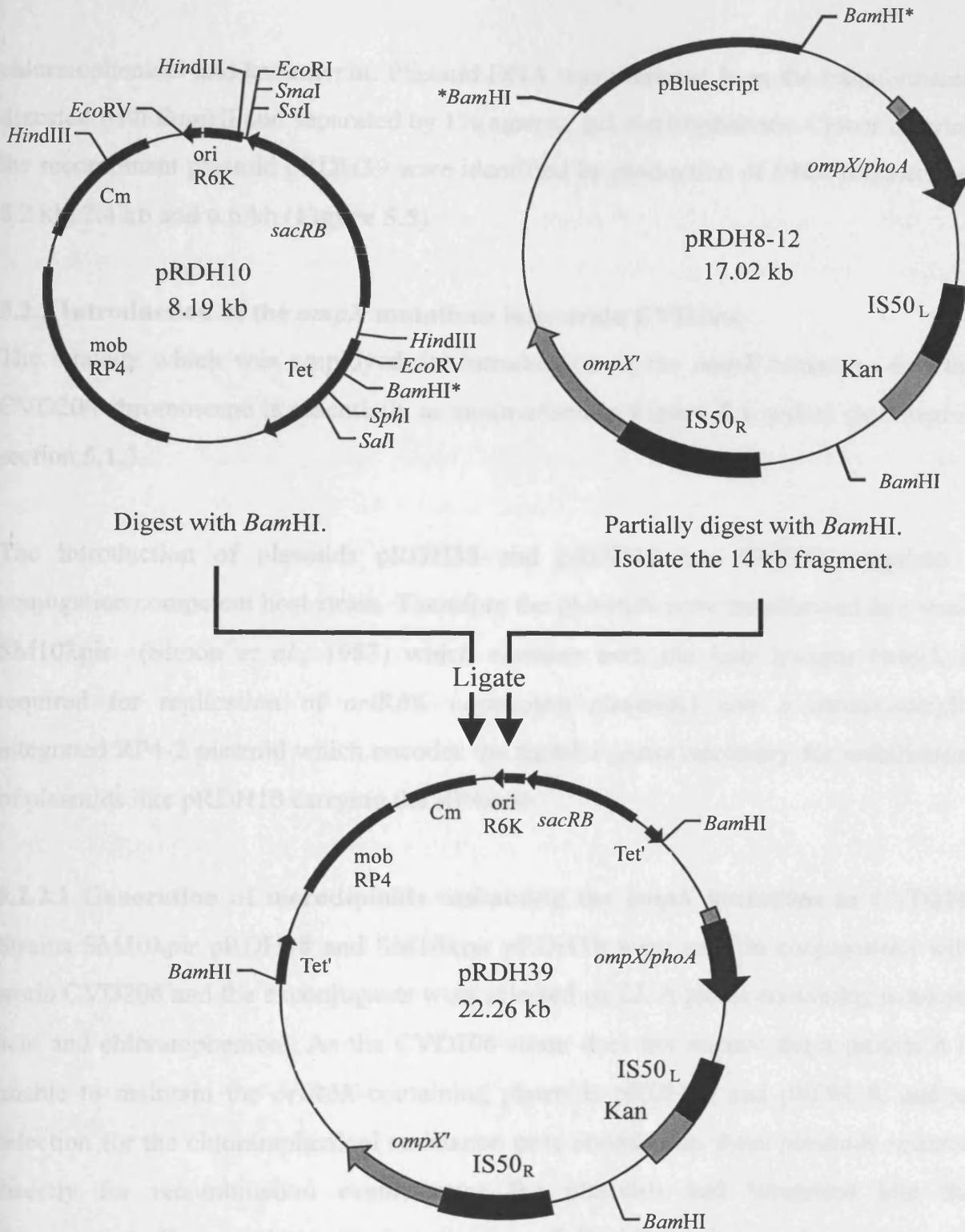


Figure 5.5 Summary of the construction of pRDH39; a positive-selection suicide vector encoding an OmpX/PhoA fusion protein. (see section 5.2.1.4 for details) Restriction enzyme sites which were used at each step in the cloning are marked with an asterisk.

chloramphenicol and kanamycin. Plasmid DNA was prepared from the transformants, digested with *Bam*HI and separated by 1% agarose gel electrophoresis. Clones carrying the recombinant plasmid pRDH39 were identified by production of DNA fragments of 8.2 kb, 7.4 kb and 6.6 kb (Figure 5.5).

5.2.2 Introduction of the *ompX* mutations into strain CVD206.

The strategy which was employed for introduction of the *ompX* mutations into the CVD206 chromosome is essentially as summarised in Figure 5.1 and as described in section 5.1.3.

The introduction of plasmids pRDH38 and pRDH39 into CVD206 required a conjugation competent host strain. Therefore the plasmids were transformed into strain SM10 λ pir (Simon *et al.*, 1983) which contains both the λ pir lysogen (which is required for replication of *oriR6K* containing plasmids) and a chromosomally integrated RP4-2 plasmid which encodes the transfer genes necessary for mobilisation of plasmids like pRDH10 carrying the RP4*mob*.

5.2.2.1 Generation of merodiploids containing the *ompX* mutations in CVD206.

Strains SM10 λ pir pRDH38 and SM10 λ pir pRDH39 were used in conjugations with strain CVD206 and the exconjugants were selected on LUA plates containing nalidixic acid and chloramphenicol. As the CVD206 strain does not encode the π protein it is unable to maintain the *oriR6K*-containing plasmids pRDH38 and pRDH39, and so selection for the chloramphenicol resistance gene encoded on these plasmids selected directly for recombination events where the plasmids had integrated into the chromosome. Due to the extensive regions of flanking sequence homologous to chromosomal sequences present in both pRDH38 and pRDH39 (2.4 kb and 6.3 kb, respectively) it was predicted that there would be a very high plasmid integration rate. Each of the conjugations resulted in the appearance of chloramphenicol and nalidixic acid resistant exconjugants at a rate of approximately 1/100 to 1/1000. With such large regions of flanking DNA there was also concern about the possibility of illegitimate

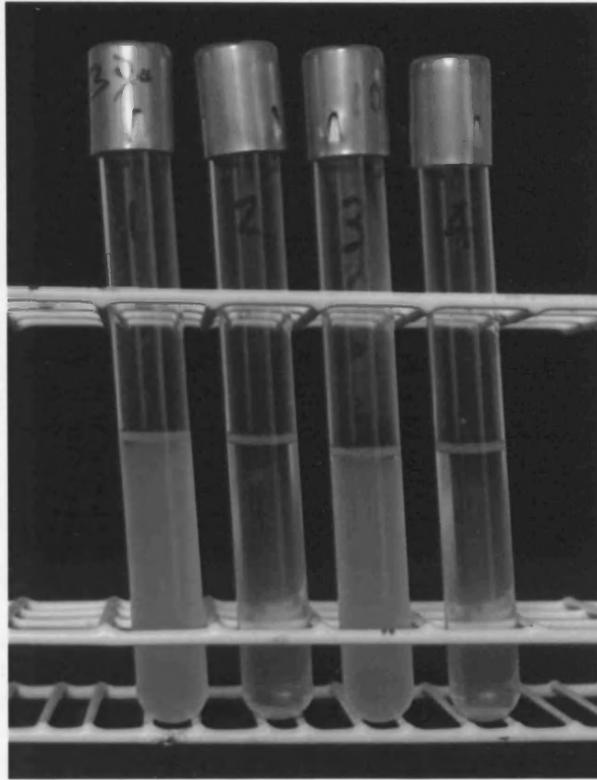
recombinations, as well as the real probability that there could be homologous recombination between the *Tn_{phoA}* sequence in pRDH39 and the CVD206 chromosomal *phoA* gene. Therefore six exconjugants were selected from each conjugation and the site of plasmid integration was confirmed by Southern hybridisation analysis. Chromosomal DNA was prepared using the miniprep method, digested with *Bam*HI and separated on 1% agarose gel electrophoresis. The fragments were transferred to Hybond-NTM by Southern blotting, and the blot hybridised with a probe made from the 6.3 kb *Bam*HI fragment of pRDH8. Each set of six exconjugants analysed gave identical bands, which were similar to those predicted for each merodiploid i.e. CVD206::pRDH38 5.8 kb and 11 kb, and CVD206::pRDH39 4.4 kb, 6.3 kb and 9.7 kb, respectively. The Southern hybridisation analysis of these twelve merodiploids gave no indication of either illegitimate recombination events or integration of pRDH39 into the *phoA* gene. It was observed, however, that a large proportion of the merodiploids obtained (in excess of 50% for pRDH39, greater than 20% for pRDH38) had acquired a new phenotype; after growth overnight in LUB these merodiploids would rapidly settle to the bottom of the tube upon standing (**Figure 5.6A**). Observation of the cultures by microscopy showed aggregation of the bacteria into irregular clumps (**Figure 5.6B**) which could be broken up by vigorous vortexing but then rapidly reformed. Similar autoaggregation has been observed with other EPEC strains (Tom Baldwin, personal communication) and a spontaneous autoaggregating variant of the nalidixic acid resistant E2348-69 strain has been isolated (RDH001; unpublished observation). It was unclear if the increased autoaggregation phenotype (IAP) observed was due to an effect of the *ompX* merodiploid construction and so both IAP⁺ and IAP⁻ merodiploids were chosen from the pRDH38 and pRDH39 exconjugants for the sucrose selection step.

5.2.2.2 Isolation of resolved merodiploids using sucrose selection. Resolved merodiploids were identified using the sucrose selection medium (Materials and Methods, section 2.3) developed by Blomfield *et al.* (1991). Overnight LUB cultures of the IAP⁺ and IAP⁻ merodiploids of CVD206::pRDH38 and CVD206::pRDH39 were

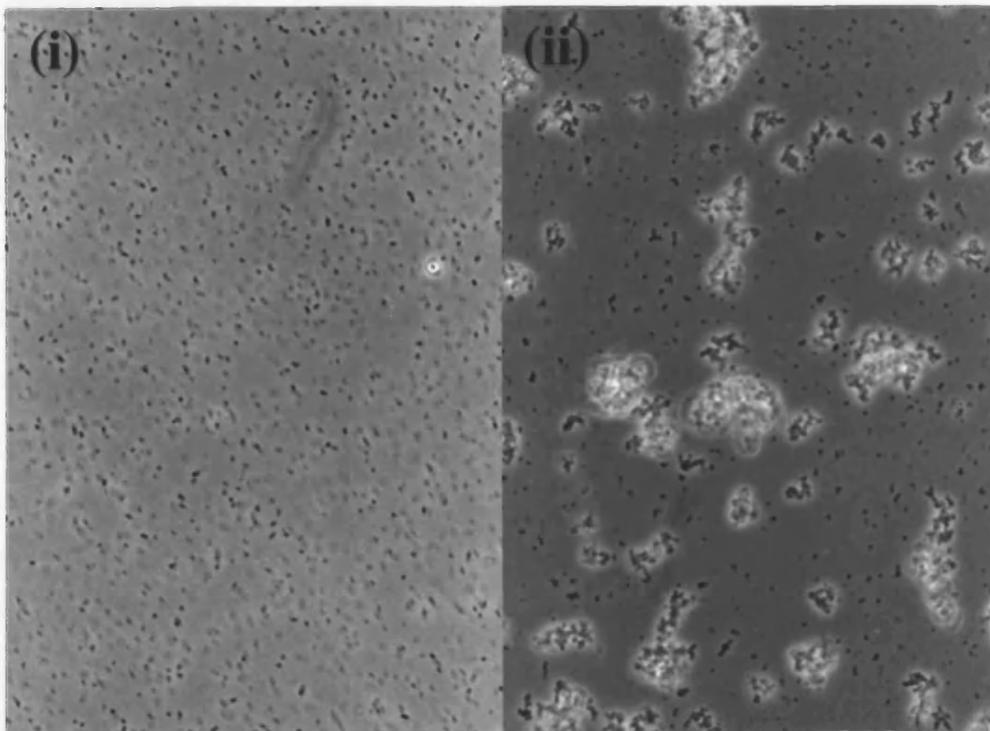
Figure 5.6 Increased autoaggregation phenotype (IAP) in EPEC. LUB cultures of EPEC were grown shaking vigorously at 37°C overnight and then allowed to stand for 1 hour. Cultures were examined for autoaggregation by (A) culture settling, or (B) phase contrast microscopy.

- (A) tube 1 - E2348-69 Strep^r
tube 2 - RDH006 (E2348-69 Strep^r IAP⁺)
tube 3 - E2348-69 Strep^r p6.8.1 (bfpA⁻)
tube 4 - RDH006 p6.8.1 (bfpA⁻)
- (B) (i) E2348-69 Strep^r
(ii) RDH006 (E2348-69 Strep^r IAP⁺)

A



B



diluted 100-fold into 6% sucrose broth and then grown overnight at 30°C. The resultant overnight cultures were diluted (10^6 - 10^7 -fold) and aliquots were spread onto 6% sucrose agar plates which were incubated overnight again at 30°C. The resultant colonies were patch-plated onto nalidixic acid or nalidixic acid and chloramphenicol containing LUA plates to test for chloramphenicol sensitivity. 100% of the 50 sucrose-resistant colonies tested for each merodiploid were identified as chloramphenicol sensitive indicating they had all undergone secondary recombination events to resolve the merodiploid.

Potential *ompX/TnphoA* CVD206 mutants were identified by patchplating the sucrose-resistant CVD206::pRDH39 colonies onto LUA plates containing kanamycin and XP. 1/50 of the IAP⁻ merodiploid-derived colonies and 2/50 of the IAP⁺ merodiploid-derived colonies were kanamycin resistant; none of the colonies appeared *phoA*-positive compared to the parent CVD206. Analysis of both the wild-type and *ompX/TnphoA* mutant CVD206 colonies indicated that the IAP⁻ and IAP⁺ phenotypes of the merodiploids were maintained irrespective of genotype indicating that the IAP phenotype was probably not due to *ompX*. To avoid complications with phenotype assignment IAP⁺ strains were not used in further experiments; the single kanamycin-resistant IAP⁻ colony was selected as the *ompX/phoA* reporter strain RDH004.

In order to identify CVD206 *ompXΔSphI* mutants twelve of the IAP⁻ sucrose-resistant colonies were examined for loss of the *SphI* site using a PCR based screen. Chromosomal template DNA was prepared using the boilate method (section 2.13) and standard 10 µl PCR reactions were performed in duplicate. Primers w11a and w9b (Chapter 4, section 4.2.1.5; Figure 4.2) were used to amplify a 2.3 kb PCR fragment, which includes the entire *ompX* sequence and approximately 300 bp of *orf2*. The PCR program optimised for the primer pair consisted of: one cycle of 95°C for 5 min, 55°C for 30 sec, and 72°C for 4 min; 35 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 4 min; and one cycle of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 6 min. One of each pair of PCR reactions was digested with *SphI* by addition of 0.1 µl of *SphI*

enzyme directly into the PCR reaction mixture and incubation at 37°C for one hour (Turbett and Sellner, 1996). The resultant DNA fragments and the uncut PCR products were separated by 1% agarose gel electrophoresis. Wild-type colonies were identified by *SphI* digestion of the 2.3 kb PCR fragment to give two bands of approximately 1.4 kb and 900 bp. A single colony of the twelve produced a PCR fragment which was uncut after digestion with *SphI* indicating it contained the *ompX*Δ*SphI* mutation; this strain was designated RDH003.

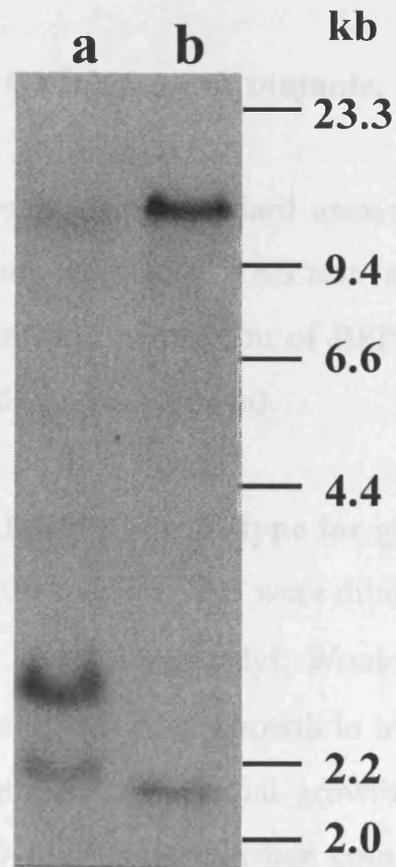
5.2.2.3 Confirmation of the CVD206 *ompX* mutants by Southern blotting analysis. Preliminary identification of the CVD206 mutants RDH003 (*ompX*Δ*SphI*) and RDH004 (*ompX*/*TnphoA*) had been made using a PCR screen and selection for kanamycin resistance respectively. To confirm that they actually contained the chromosomal mutations predicted, and to check that there were no other chromosomal alterations at the *ompX* locus, the two strains were compared to CVD206 by Southern hybridisation. Chromosomal DNA was prepared from strains RDH003, RDH004 and CVD206 using the miniprep method. RDH003 DNA was digested with *SphI*, RDH004 DNA was digested with *Bam*HI and CVD206 DNA was digested separately with both enzymes. The DNA fragments produced were separated by 1% agarose gel electrophoresis and transferred to Hybond-NTM by Southern blotting. The blots obtained were hybridised with a probe made from the 6.3 kb *Bam*HI fragment of pRDH8 (as previously described, section 4.2.4.2). As expected, from the sequence of pRDH8, the probe identified three *SphI* fragments in strain CVD206. In contrast, in strain RDH003 there were only two *SphI* fragments which hybridised with the probe indicating loss of an *SphI* site. The combined total size estimated for all *SphI* fragments in each strain was equivalent indicating that no chromosomal rearrangements had occurred (Figure 5.7A). In strain RDH004 the probe identified two *Bam*HI fragments compared to the single 6.3 kb *Bam*HI wild-type fragment seen in CVD206. The fragment sizes seen in RDH004 were estimated to be approximately equivalent to the 4.4 kb and 9.7 kb *ompX*/*TnphoA*-containing fragments which are seen upon digestion

Figure 5.7 Confirmation of the *ompX* mutations in RDH003 and RDH004 by Southern blot analysis. Chromosomal DNA prepared from strains CVD206, RDH003 (*ompX*) and RDH004 (*ompX/TnphoA*) were digested with *SphI* or *BamHI*, the fragments separated by 1% agarose gel electrophoresis and transferred to Hybond-N™ by Southern blotting. The blots were hybridised with a probe consisting of the 6.3 kb *BamHI* fragment of plasmid pRDH8.

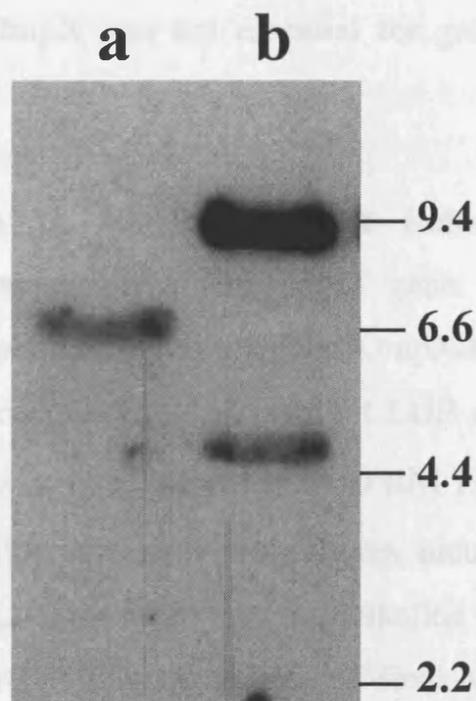
(A) CVD206 (lane a) and RDH003 (lane b) were digested with *SphI*. RDH003 gave two pRDH8 specific bands compared to the three bands found in CVD206 indicating the deletion of an *SphI* site.

(B) CVD206 (lane a) and RDH004 (lane b) were digested with *BamHI*. RDH004 had two pRDH8 specific bands compared to the single wild-type band in CVD206 indicating that *TnphoA* had been inserted within the *ompX* locus.

A



B



of plasmid pRDH8-12 with *Bam*HI; this indicated that the *TnphoA* insertion was correctly located within the *ompX* locus (Figure 5.7B).

5.2.3 Analysis of the phenotypes of the CVD206 *ompX* mutants.

Strains RDH003 and RDH004 were tested using standard assays for expression of EPEC virulence determinants, i.e. localised adherence, FAS test (after transformation with pCVD438), secretion of Esp proteins and expression of BFP; as expected, they were wild-type for all phenotypes assayed (data not shown).

5.2.3.1 The *ompX* frameshift mutant RDH003 is wild-type for growth in low iron.

Overnight LUB cultures of strains CVD206 and RDH003 were diluted into either LUB or an iron-limited medium (LUB + 300 μ M 2,2'-dipyridyl; Wooldridge *et al.*, 1992) and the cultures grown at 37°C with vigorous shaking. Growth in iron-limited medium resulted in a significant reduction in both the exponential growth rate and the final optical density in both CVD206 and RDH003 cultures when compared to growth in LUB (Figure 5.8). There was no apparent difference in growth rate or final culture density between the RDH003 *ompX* mutant and parent CVD206 during growth in iron-limited medium indicating that OmpX was not essential for growth under low iron conditions.

5.2.3.2 *ompX/phoA* expression in RDH004 is not induced by low iron concentrations.

The level of expression of the *ompX* gene was determined by comparison of the alkaline phosphatase activity of the OmpX/PhoA fusion protein carrying strain RDH004 to its parent CVD206. Overnight LUB cultures of CVD206 and RDH004 were inoculated into iron-replete (LUB + 50 μ M FeCl₃) or iron-limited medium (LUB + 300 μ M 2,2'-dipyridyl) and the cultures incubated at 37°C with vigorous shaking. The growth, OD₆₀₀ measurement, and alkaline phosphatase activity of the cells were determined hourly. The growth rate and final optical density of the

RDH004 strain were identical with that of the parental strain, CVD206, in both iron-limited and iron-replete media (**Figure 5.9A**) indicating that the *ompX::TnphoA* mutation, in agreement with the *ompXΔSphI* mutation (RDH003; see above), had no effect upon EPEC iron uptake *in vitro*. The alkaline phosphatase activity measured for RDH004 in iron-replete medium fluctuated over the time course from 3 to 8 units/ml; however, deduction of the similar fluctuations observed in background phosphatase activity, measured in strain CVD206, resulted in a average alkaline phosphatase activity over the time course of 2.02 U/ml (\pm 0.45 U) (**Figure 5.9B**). Growth of RDH004 in iron-limited medium resulted in a decrease in the total level of alkaline phosphatase activity; however, when the data were again corrected for the background phosphatase activity, as measured in CVD206, the average level of alkaline phosphatase activity over the time course was 1.24 U/ml (\pm 0.39 U) (**Figure 5.9B**). These data suggest a low constitutive expression level for the OmpX protein which is not inducible by growth in low iron concentrations.

5.2.4 Strategy for the construction of an *emtA* frameshift mutation.

As the upstream gene *orf4* is divergently transcribed it must be presumed that the *emtA* gene possesses its own promoter sequences, though these have not yet been identified. *emtA* also does not appear to form part of an operon; it is immediately followed by predicted Rho-independent terminator sequences (**Appendix 2**) and the downstream sequence encodes a predicted open reading frame, *orf3*, which is transcribed in the opposite direction (**Chapter 4, Figure 4.7**). It was therefore considered unlikely that a frameshift mutation would result in the altered expression of either of the adjacent genes and so would appear to be the best option for a non-polar mutation. *emtA* contains a single *HindIII* site approximately 250 bp downstream of the initiation codon; insertion of 4 bp at this point would result in a frameshift mutation whose modified sequence would encode a truncated EmtA polypeptide of 91 amino acids (processed polypeptide 76 amino acids) (**Figure 5.10B**). It was predicted from the lytic transglycosylase active site motif data (Dijkstra and Thunnissen, 1994) that even if the truncated polypeptide were stable it would have lost a significant proportion of its

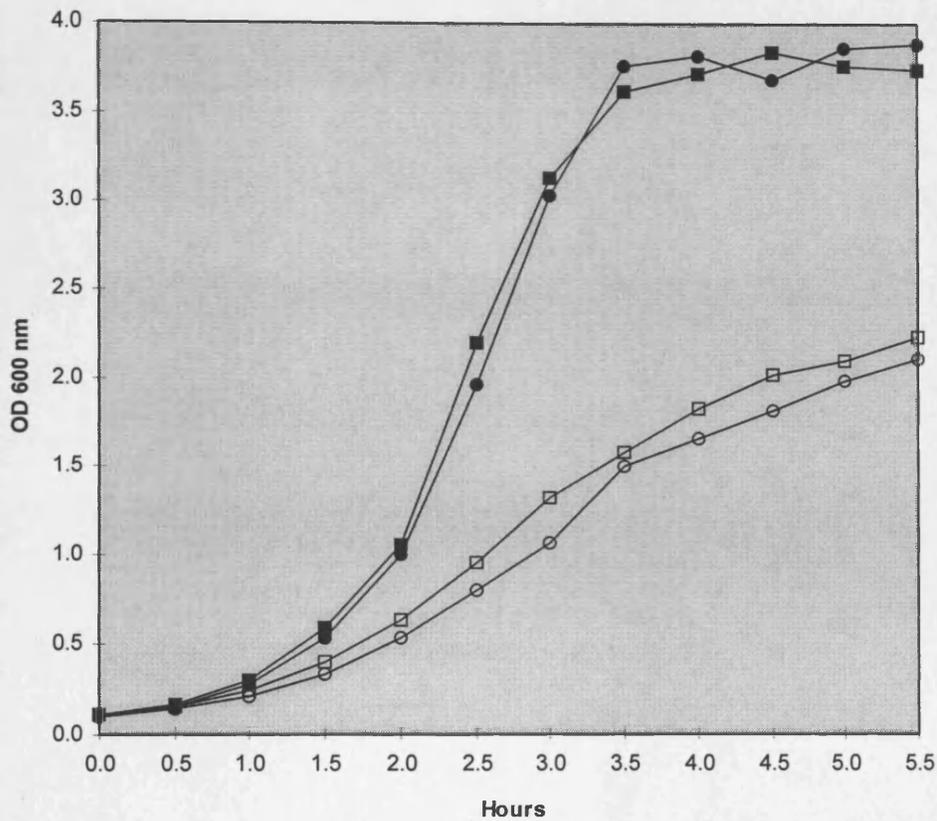
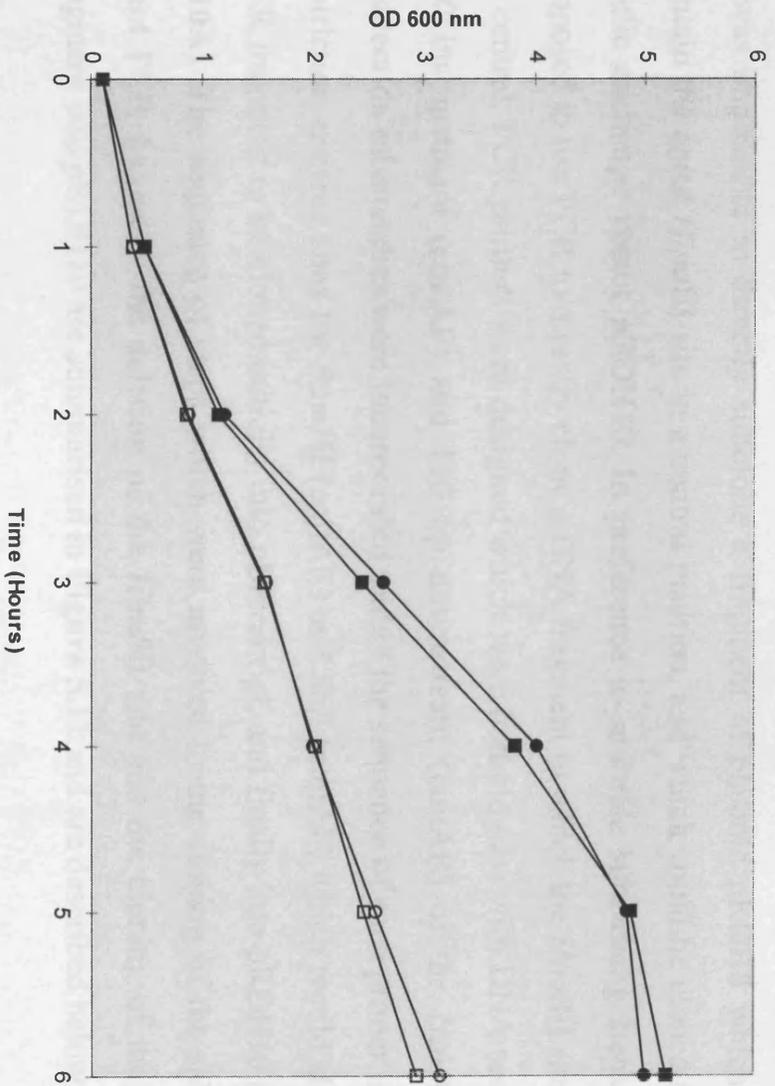
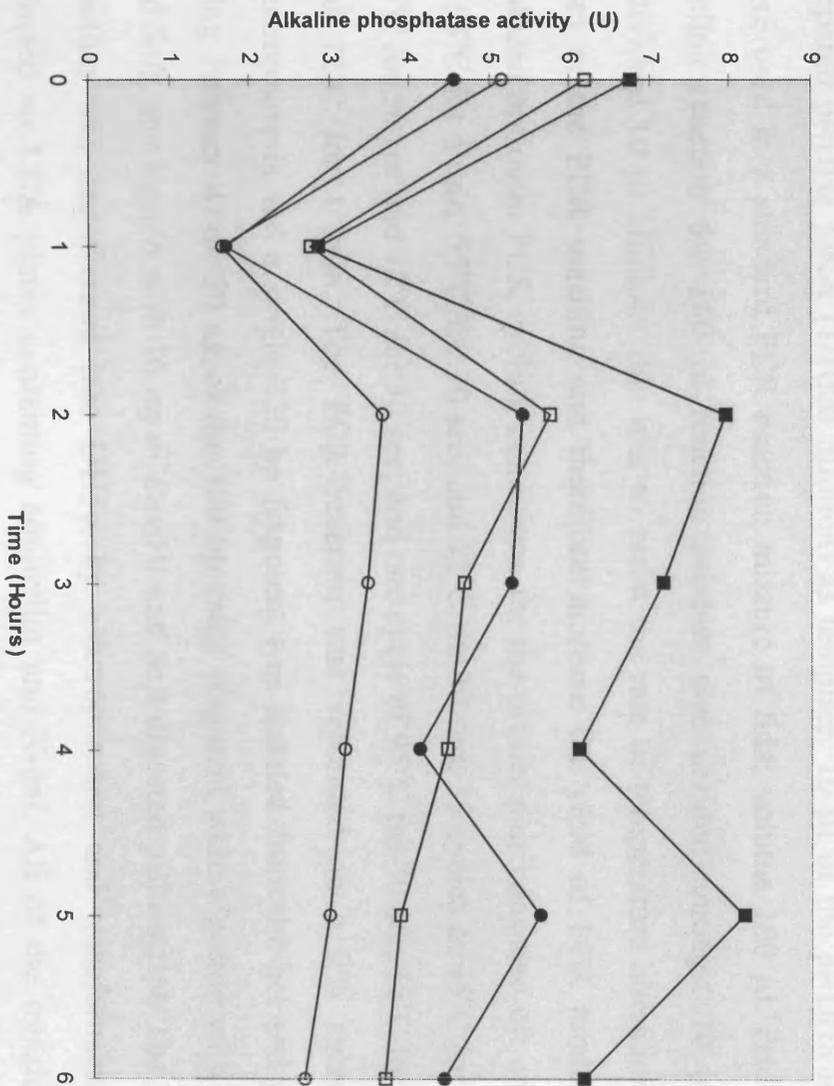


Figure 5.8 Growth of strain RDH003 (*ompX*) in iron limiting conditions. EPEC strains CVD206 (●,○) and RDH003 (■,□) were grown as 50 ml cultures at 37°C with vigorous shaking in either LUB (filled symbols) or iron restricted media (LUB + 300 μM 2,2'-dipyridyl; open symbols).

Figure 5.9 Investigation of *ompX* expression levels under iron limiting conditions. 50 ml cultures of EPEC strains CVD206 (●,○) and RDH004 (*ompX*::Tn*phoA*; ■,□) were grown at 37°C with vigorous shaking in either iron-replete medium (LUB + 50 μM FeCl₃; filled symbols) or iron-restricted medium (LUB + 300 μM 2, 2'-dipyridyl; open symbols). Samples were taken hourly and measured for OD₆₀₀ (A) and alkaline phosphatase activity (B). This is a representative example of three separate experiments.

A**B**

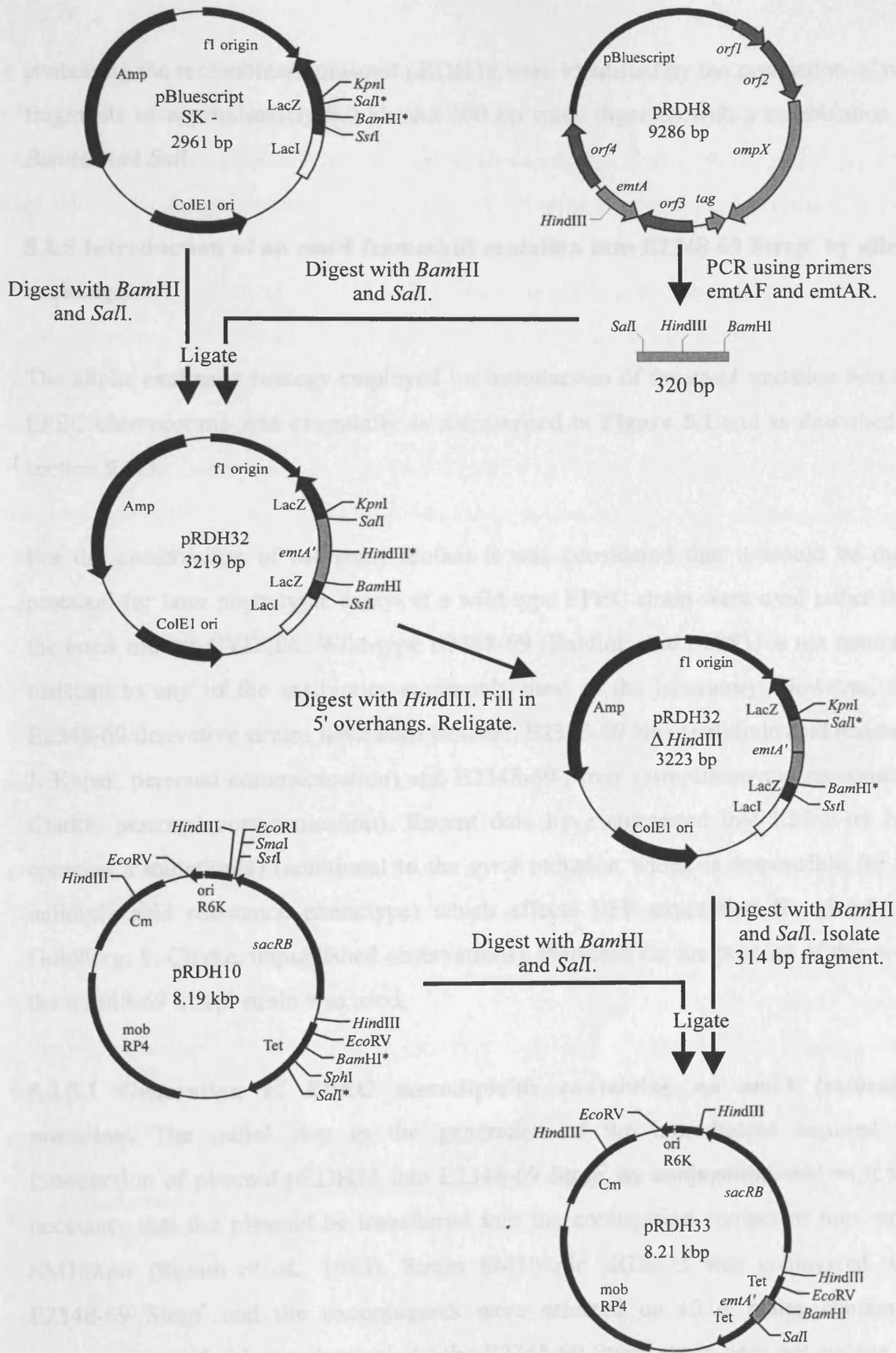
active site and therefore be inactive. Due to the lack of suitable restriction enzyme sites it was impossible to directly subclone a fragment of plasmid pRDH8 which would contain the *emtA* *Hind*III site in a central position, and which could be cloned into the allelic exchange vector pRDH10. In preference to multiple subcloning steps it was proposed to use PCR to directly clone a DNA fragment in which the *Hind*III site would be central. PCR primers were designed which were homologous with DNA sequences 160 bp upstream (*emtAF*) and 160 bp downstream (*emtAR*) of the *Hind*III site. Nucleotide mismatches were incorporated within the sequence of each primer to create restriction enzyme sites for *Bam*HI (*emtAR*) and *Sal*I (*emtAF*) which would allow the PCR fragment to be cloned initially into pBluescript, and finally into pRDH10 (**Figure 5.10A**). The sequence of steps which were involved in the cloning of the subgenetic *emtA* PCR fragment, the deletion of the *Hind*III site and the cloning of the mutant fragment into pRDH10 are summarised in **Figure 5.11** and are described below.

5.2.4.1 Cloning a subgenetic PCR fragment of *emtA* into pBluescript. Primers *emtAF* and *emtAR* were used to amplify a 320bp subgenetic *emtA* fragment using CsCl prepared pRDH8 DNA (1/1000 dilution) as template. 10 µl of the pRDH8 template were used in a standard PCR reaction mixture of final volume 100 µl. Prior to the cycling reaction the 100 µl reaction mixture was divided amongst 10 tubes as individual 10 µl aliquots; this was to assist the rate of temperature alteration (ramp rate) in the PCR machine and therefore increase the yield of PCR product. The deduced optimum PCR cycling conditions for the primer pair consisted of: one cycle of 95°C for 5 min, 55°C for 30 sec, and 72°C for 30 sec; 35 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec; and one cycle of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min. The PCR reaction was separated on a 2% agarose gel electrophoresis and a single 320 bp fragment was isolated from the gel and purified using Polymer wool. 50 ng of the 320 bp *emtA* fragment were digested with *Bam*HI and *Sal*I, and ligated with 10 ng of *Bam*HI and *Sal*I digested pBluescript. The ligation reaction was transformed into DH5α by electroporation and transformants were selected on LUA plates containing ampicillin and X-gal. All of the transformants

obtained from the initial transformation were blue, indicating, by continued expression of the *lacZ* gene, that they did not contain inserts. A repeated ligation and transformation also produced only blue colonies; however, it was observed that at an early stage in growth a proportion of the colonies were a much lighter blue colour than the rest. Plasmid DNA was prepared from a number of light-blue and dark blue colonies, the DNA digested with *Bam*HI and *Sal*I, and separated by 2% agarose gel electrophoresis. The dark blue colonies produced a single band of approximately 3 kb (indicating pBluescript), whilst the light blue colonies produced two bands of approximately 3 kb and 300 bp respectively, indicating they contained the recombinant plasmid pRDH32. It was speculated that the low level expression of β -galactosidase activity observed in pRDH32-carrying clones was due to transcriptional read-through because of the small size of the *emtA* insert. A single light blue clone carrying the plasmid pRDH32 was used to prepare CsCl plasmid DNA and the DNA sequence of the 320 bp insert was determined using the ABI PRISM™ BIG DYES Terminator Cycle Sequencing kit with the universal primers m13r and m13f (section 2.12). The DNA sequences obtained closely matched the sequence of pRDH8 and did not contain any basepair changes which would affect the predicted amino acid sequence of *emtA*.

5.2.4.2 Construction of a frameshift mutation in *emtA* by deletion of the unique *Hind*III site in pRDH32. During the construction of plasmid pRDH32 the *Hind*III site present in the pBluescript multiple cloning site was deleted such that the *Hind*III site within the cloned *emtA* subgenic fragment was now unique (Figure 5.11). 500 ng of plasmid pRDH32 were digested with *Hind*III to linearise the plasmid, the *Hind*III 5' overhanging ends were filled-in using Klenow and the blunt ended fragment was religated. The ligation mixture was transformed into DH5 α by electroporation and transformants were selected on LUA plates containing ampicillin. Even though Klenow is much more effective at filling in 5' overhangs than chewing back 3' overhangs, due to the low efficiency of ligation of the blunt-ended fragments compared with that of sticky-ended fragments it was still probable that the majority of clones would consist of religated pRDH32. In order to enrich for the clones carrying the

Figure 5.11 Summary of the construction of pRDH33; a positive-selection suicide vector containing an *emtA* frameshift mutation. (see sections 5.2.4.1-5.2.4.3 for details; see also **Figure 5.10**). Restriction enzyme sites which were used in each cloning step are indicated with asterisks.



containing the recombinant plasmid pRDH33 were identified by the production of two fragments of approximately 8.0 kb and 300 bp when digested with a combination of *Bam*HI and *Sal*I.

5.2.5 Introduction of an *emtA* frameshift mutation into E2348-69 Strep^r by allelic exchange.

The allelic exchange strategy employed for introduction of the *emtA* mutation into the EPEC chromosome was essentially as summarised in Figure 5.1 and as described in section 5.1.3.

For the construction of the *emtA* mutant it was considered that it would be more practical for later phenotypic assays if a wild-type EPEC strain were used rather than the *ea*eA mutant CVD206. Wild-type E2348-69 (Baldini *et al.*, 1983) is not naturally resistant to any of the antibiotics commonly used in the laboratory. However, two E2348-69 derivative strains have been isolated; E2348-69 Nal^r (nalidixic acid resistant; J. Kaper, personal communication) and E2348-69 Strep^r (streptomycin resistant; S. Clarke, personal communication). Recent data have suggested that E2348-69 Nal^r contains a mutation(s) (additional to the *gyrA* mutation which is responsible for the nalidixic acid resistance phenotype) which affects BFP expression (R. Haigh, M. Goldberg, S. Clarke, unpublished observations); therefore for the purpose of this work the E2348-69 Strep^r strain was used.

5.2.5.1 Generation of EPEC merodiploids containing an *emtA* frameshift mutation. The initial step in the generation of the merodiploid required the introduction of plasmid pRDH33 into E2348-69 Strep^r by conjugation, and so it was necessary that the plasmid be transferred into the conjugation competent host strain SM10 λ pir (Simon *et al.*, 1983). Strain SM10 λ pir pRDH33 was conjugated with E2348-69 Strep^r and the exconjugants were selected on LUA plates containing streptomycin and chloramphenicol. As the E2348-69 Strep^r strain does not encode the

π protein it is unable to maintain *oriR6K*-containing plasmids (such as pRDH33) and so selection upon chloramphenicol selected directly for recombination events where the pRDH33 plasmid had integrated into the chromosome. Due to the relatively small size of the *emtA* fragment, which was the only plasmid region homologous with chromosomal sequences, it was predicted that the yield of merodiploids would be low. Chloramphenicol and streptomycin resistant exconjugants were isolated at a rate of approximately $1/10^5$ streptomycin resistant colonies. Ten of these E2348-69 merodiploid colonies were selected and tested for the autoaggregation phenotype observed with the CVD206 *ompX* merodiploids; five of the ten E2348-69::pRDH33 colonies showed autoaggregation in LUB. Single IAP⁺ and IAP⁻ E2348-69::pRDH33 merodiploids were chosen for sucrose selection; however, the IAP⁺ merodiploid was only to be used to isolate wild-type E2348-69 Strep^r for investigation of the IAP phenotype.

5.2.5.2 Isolation of resolved merodiploids using sucrose selection. Resolved E2348-69::pRDH33 merodiploids were identified using the protocol modified from Blomfield *et al.*, (1991) as described in section 5.1.3. Overnight LUB cultures of both the IAP⁺ and IAP⁻ *emtA* merodiploids were diluted 100-fold into 6% sucrose broth and grown overnight at 30°C. This overnight culture was diluted (10^{-6} - 10^{-7}), aliquots spread onto 6% sucrose agar plates and incubated overnight at 30°C. The resultant sucrose-resistant colonies were patch plated onto streptomycin or chloramphenicol LUA plates to test for chloramphenicol sensitivity. 11/12 of the IAP⁻ colonies tested were chloramphenicol sensitive indicating excision and loss of the pRDH33 plasmid. In order to identify *emtA* mutants, the 11 chloramphenicol sensitive colonies obtained were screened for the loss of the *emtA* *Hind*III site using PCR. Chromosomal template was prepared from each of the colonies using the boilate method and standard 10 μ l PCR reactions were performed in duplicate. Primers *emtAF* and *emtAR* were used to amplify the 320 bp subgenic *emtA* fragment using the optimised PCR conditions described in section 5.2.4.1. One of each pair of PCR reactions was digested with *Hind*III by addition of 0.2 μ l of *Hind*III enzyme directly to the PCR reaction mixture

and incubation at 37°C for 1 hour (Turbett and Sellner, 1996). The resultant DNA fragments (a 160 bp doublet if digested with *HindIII*) and the uncut PCR fragments were separated by 2% agarose gel electrophoresis. 5/11 of the chloramphenicol sensitive colonies produced PCR fragments which were 320 bp after digestion with *HindIII* indicating that they contained the *emtA* frameshift mutation; one of these five colonies was selected and the strain designated RDH005. Concurrently a sucrose resistant, chloramphenicol sensitive IAP⁺ colony was identified by PCR which produced two 160 bp fragments on digestion with *HindIII* indicating it contained the wild-type *emtA* allele. This strain, which was effectively E2348-69 Strep^r IAP⁺, was designated RDH006 and retained for further analysis; see below section 5.2.8.

5.2.5.3 Confirmation of the *emtA*Δ*HindIII* mutation by Southern blot analysis and DNA sequencing. The preliminary identification of RDH005 was based upon a negative selection i.e. the inability of a PCR product to be digested with *HindIII*. In order to confirm the mutation and to check that there were no other chromosomal rearrangements at the *emtA* locus, the strain was compared with E2348-69 Strep^r by Southern hybridisation. Chromosomal DNA was prepared from E2348-69 Strep^r, E2348-69::pRDH33 and RDH005 using the miniprep method and 1 µg of each was digested with *HindIII*. The DNA fragments produced were separated on 1% agarose gel electrophoresis and transferred to Hybond-NTM by southern blotting. The blot was hybridised with the pRDH8 6.3 kb probe (section 4.2.4.2) at high stringency (Figure 5.12). E2348-69 Strep^r gave two pRDH8 specific bands of approximately 1 kb and 5 kb, whilst RDH005 gave a single larger band of approximately 6 kb indicating the loss of the *HindIII* site. As final confirmation of the *emtA* mutation in RDH005 the chromosomal DNA sequence was determined from a PCR fragment produced using the *emtAF* and *emtAR* primers. The 320 bp *emtA* subgenic fragment was generated (from a RDH005 boilate preparation) as described above and 200 ng of fragment were used in ABI PRISMTM BIG DYES Terminator Cycle Sequencing reactions with each of the primers *emtAF* and *emtAR*. The use of the same primer for both PCR and DNA sequencing reactions can often result in poor quality DNA sequence and this was the

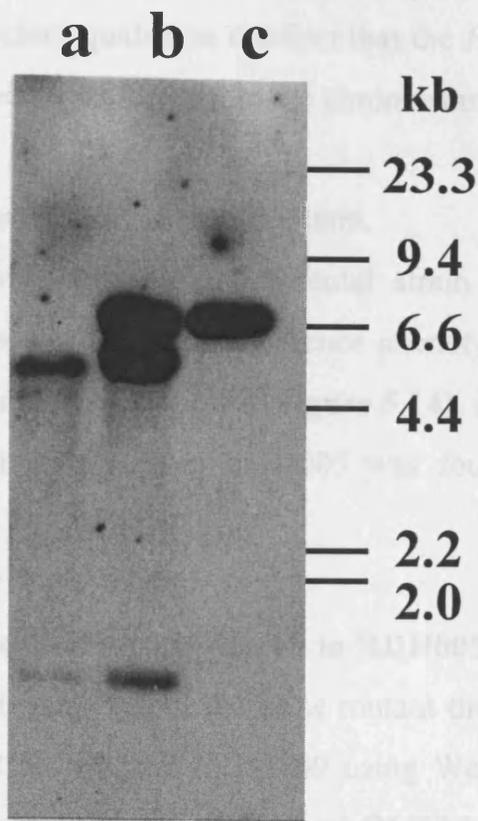


Figure 5.12 Confirmation of the *emtA* frameshift mutation in RDH005 by Southern blot analysis. Chromosomal DNA from EPEC strains E2348-69 *Strep*^r (lane a), E2348-69::pRDH33 (lane b), and RDH005 (*emtA*⁻; lane c) were digested with *Hind*III, the fragments separated by 1% agarose gel electrophoresis and transferred to Hybond-NTM by Southern blotting. The blot was hybridised with a probe consisting of the 6.3 kb fragment of pRDH8. Strain E2348-69 gave two pRDH8 specific bands whilst RDH005 gave a single band of increased size indicating the loss of a *Hind*III site.

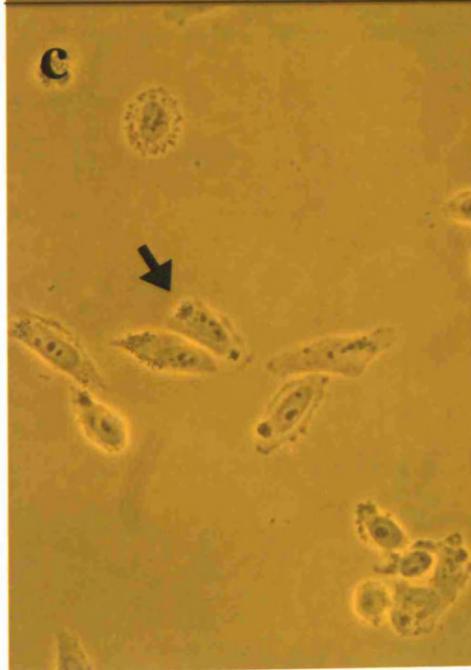
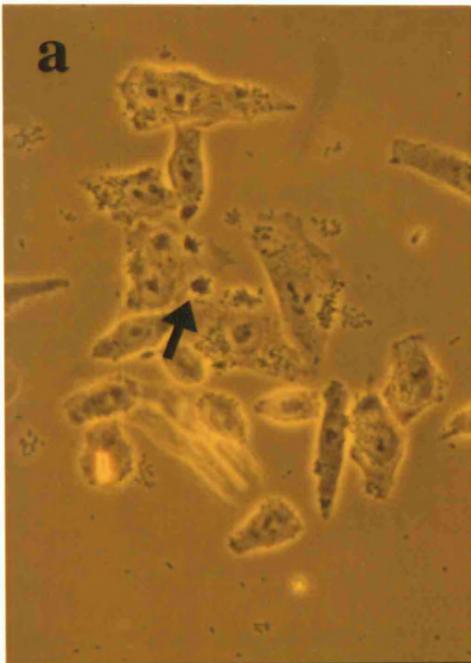
case with the *emtA* derived sequence. Fortunately the DNA sequence obtained from primer *emtAR* was of sufficient quality to confirm that the *Hind*III deletion identified in *pRDH32ΔHind*III had been transferred into the chromosome of RDH005.

5.2.6 Phenotypic analysis of EPEC strain RDH005.

Strain RDH005 (*emtA*⁻) was compared to parental strain E2348-69 Strep^r in the standard EPEC virulence assays; localised adherence phenotype and FAS test (**Figure 5.13**), adherence and invasion of HEp-2 cells (**Figure 5.14**), and expression of bundle-forming pili (**Figure 5.15**). In all cases RDH005 was found to be phenotypically identical to parent.

5.2.6.1 Altered time course of BfpA expression in RDH005. In order to confirm that BFP expression was indeed wild-type in the *emtA* mutant the expression of the BfpA protein was examined in RDH005 and E2348-69 using Western blotting. Overnight bacterial cultures were inoculated into pre-warmed DMEM medium in 6-well tissue culture trays and incubated at 37°C under 5% CO₂. Samples were taken hourly, the bacteria pelleted and resuspended in 2x SDS-PAGE loading buffer. Equal protein loadings of the cells (determined by OD₆₀₀ readings; **Figure 5.16A**) were separated by 15% SDS-PAGE gel electrophoresis and transferred to nitrocellulose filters by Western blotting. BfpA protein bands were identified using the Amersham ECL detection system with the rabbit anti-BFP antibody (Giron *et al.*, 1993) and a goat anti-rabbit IgG peroxidase conjugate (**Figure 5.16B**). Despite the similarity in growth curves of the two strains it was apparent that the expression of BfpA was slower in the *emtA* mutant; at one hour E2348-69 had a level of BfpA protein per cell which was not achieved in RDH005 until the three hour sample. However, in light of the results above, which indicated that levels of adherence and invasion were unaltered in RDH005 (at least at the three hour time point), it is unclear if the observed difference in BfpA expression is significant.

Figure 5.13 Localised adherence and AE lesion formation by RDH005 (*emtA*⁻). E2348-69 Strep^r (panels a and b) and RDH005 (*emtA*⁻) (panels c and d) were assayed for localised adherence and the formation of attaching and effacing lesions on HEp-2 cells using the FAS test as described in Materials and Methods. Attached bacteria were identified using phase contrast microscopy (panels a and c) and lesion formation was identified by bacterial-associated regions of bright fluorescence (aggregated actin stained with FITC-phalloidin) using incident light fluorescence microscopy (b and d). Adherent bacteria and lesions are indicated with arrows.



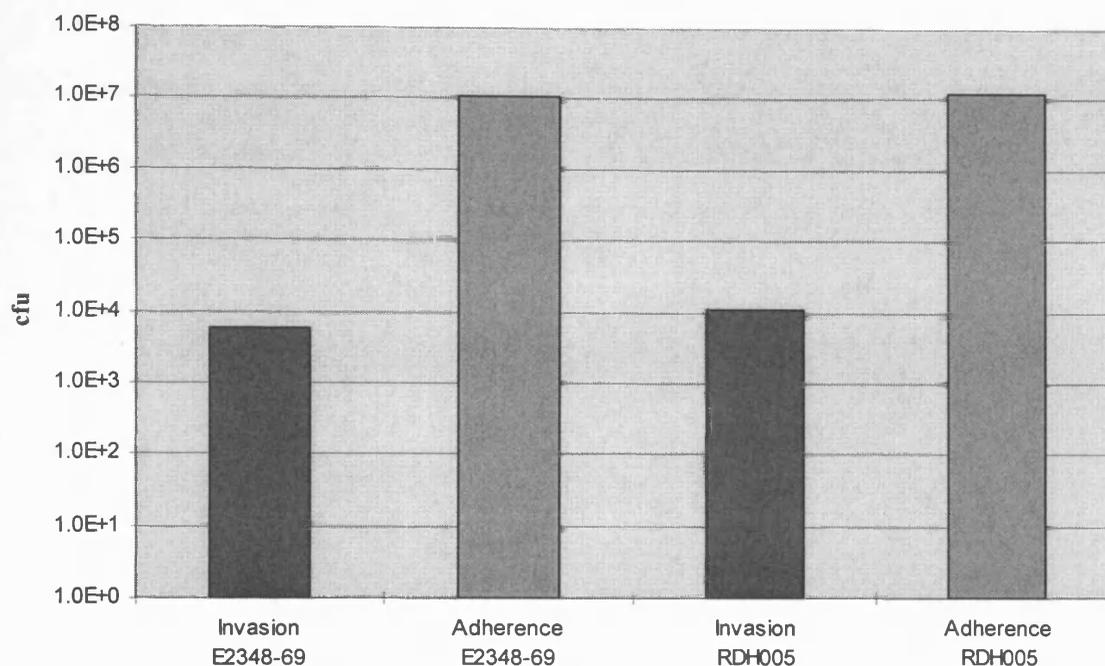


Figure 5.14 Characterisation of the adherence and invasion of RDH005 (*emtA*) in HEp-2 cells. Approximately 2×10^7 bacteria ($10 \mu\text{l}$ of an overnight LUB culture - 2×10^9 cfu/ml) were added to each well of a 24 well tissue culture plate containing a confluent HEp-2 cell monolayer. Cells were incubated at 37°C under $5\% \text{CO}_2$ for 3 hours. At this point nonadherent bacteria were removed by washing in PBS and adherent bacteria recovered; for invasion assays the monolayer was overlaid with DMEM containing $100 \mu\text{g/ml}$ gentamicin and incubated for a further 2 hours. Gentamicin-protected bacteria or adherent bacteria were recovered from the monolayer by lysing the cells in 1% Triton X-100 and enumerated by serial dilution. Data presented are the average of three separate experiments.

Figure 5.15 Production of BFP by RDH005 (*emtA*⁻). E2348-69 Strep^f (panels a and b) and RDH005 (*emtA*⁻) (panels c and d) were allowed to attach to HEp-2 cell and were then assayed for the production of BFP as described in Materials and Methods (section 2.5.3.2). Attached bacteria were identified using phase contrast microscopy (panels a and c) and BFP were visualised as bacterial-associated fluorescent strands (detected with rabbit anti-BFP and donkey anti-rabbit FITC conjugate antibodies) using incident light fluorescence microscopy (panels b and d). Adherent bacteria and pili are indicated with arrows.

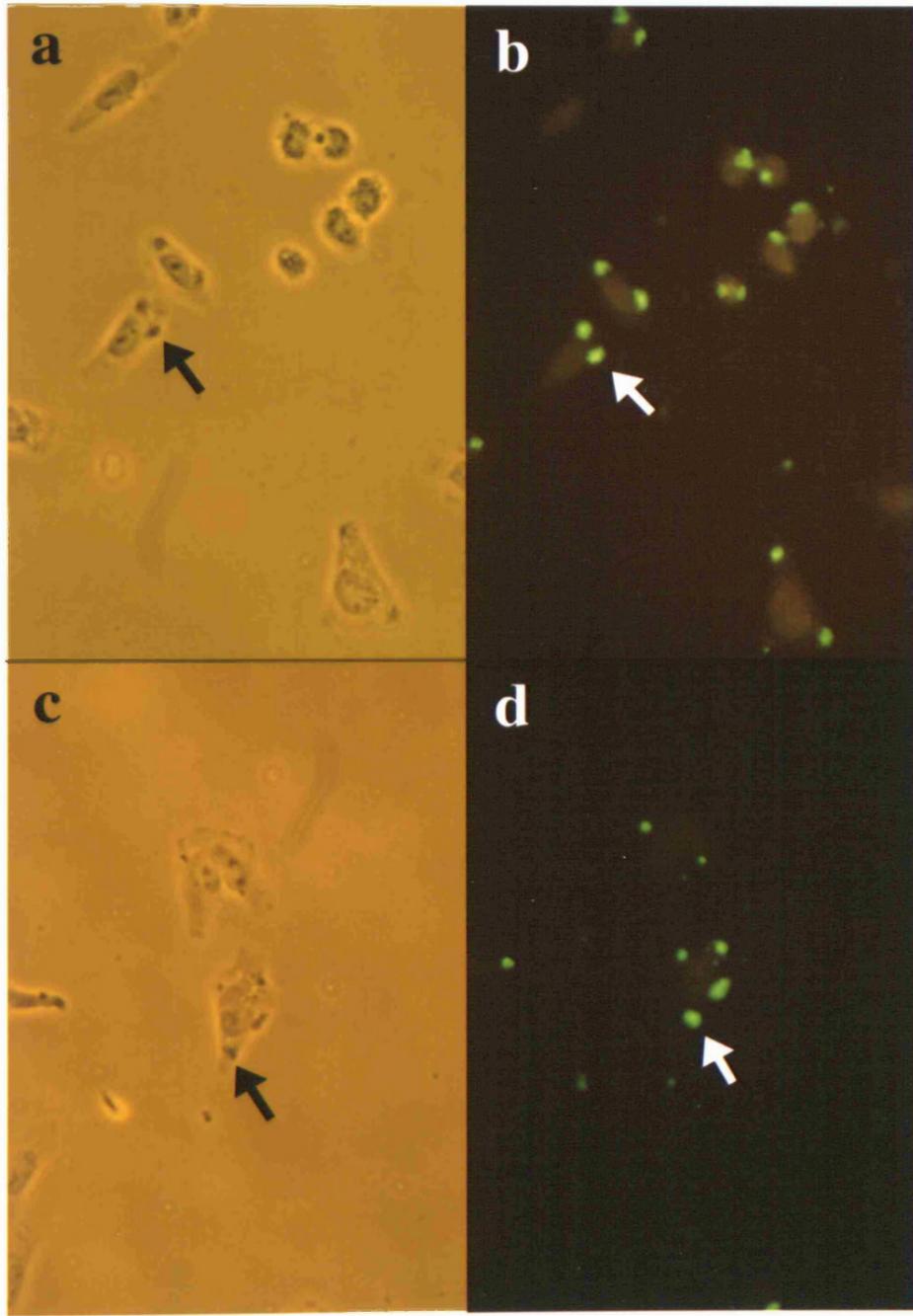


Figure 5.16 Western blot analysis of BfpA expression in RDH005 (*emtA*). 40 μ l aliquots of overnight LUB bacterial cultures of EPEC strains E2348-69 or RDH005 were inoculated, in triplicate, into individual wells of 6-well tissue culture plates containing 2 ml of DMEM (which had been preincubated for 90 mins at 37°C under 5% CO₂); the cultures were then incubated statically for up to 3 hours. Samples were taken hourly, the optical density measured, and the bacterial cells then pelleted by centrifugation and resuspended in 2x SDS-PAGE sample buffer (0.1 OD units/10 μ l). 0.05 OD units of each cell sample were separated on a 13% SDS-PAGE minigel and transferred to nitrocellulose by Western blotting (sections 2.16.1 and 2.16.4). BFP proteins (principally BfpA) were visualised using the polyclonal rabbit anti-BFP antibody (Giron *et al.*, 1993) and a goat anti-rabbit peroxidase conjugate secondary antibody; the blot was developed using the ECL system (Amersham LIFE SCIENCE; section 2.16.5).

(A) Growth curves of strains E2348-69 (■) and RDH005 (●).

(B) Western blot analysis of BfpA expression. The gel loading order was E2348-69 (lanes a, c, e and g) and RDH005 (lanes b, d, f, and h); 0 hour (lanes a and b), 1 hour (lanes c and d), 2 hours (lanes e and f) and 3 hours (lanes g and h). The doublet which represents both modified and unmodified (i.e. with signal sequence intact) BfpA is indicated by an arrow; other bands observed are constant between samples and result from non-specific binding of the secondary antibody.

5.2.7 Investigation of a proposed dominant-negative phenotype for the *emtA*/*TnphoA* mutants CVD206::*TnphoA* #13 and #137.

It was apparent from the phenotypic analysis of strain RDH005 that a non-polar mutation in *emtA* did not result in the complete BFP⁻ phenotype observed in the CVD206 *emtA*::*TnphoA* mutants #13 and #137. This result may be interpreted in two ways; (i) the specific loss of EmtA production is not responsible for the phenotype seen in *emtA*::*TnphoA* mutants, or (ii) the fact that the two mutations are different in genotype may explain their different phenotypes. The non-polar mutation, *emtA*Δ*HindIII*, was designed to encode an inactive truncated protein; as this partial polypeptide would be unable to fold correctly it was envisioned that it would be rapidly degraded by cell proteases. In contrast the *emtA*::*TnphoA* mutants were isolated on the basis that they formed highly active alkaline phosphatase protein fusions. The fact that the EmtA/PhoA protein fusions were *phoA*⁺ indicated that the EmtA lipoprotein signal sequence was functioning correctly, and the high level of alkaline phosphatase activity suggested that the fusion protein was stable; it was therefore proposed that the 91 amino acids of EmtA present in the fusion had folded correctly and that the fusion protein might retain an N-terminal EmtA protein domain (though as it lacked the lytic transglycosylase active site it would be inactive). It has been proposed that the *E. coli* lytic transglycosylases form part of a multi-enzyme complex or "holoenzyme", containing both peptidoglycan synthesis and degradation enzymes, which is responsible for replication of the murein cell wall (Höltje, 1996a; Höltje, 1996b). If similar complexes are also involved in the biogenesis of BFP it is possible to hypothesise that, whilst the loss of active EmtA protein (*emtA*Δ*HindIII*, RDH005) from the complex could be complemented by another *E. coli* transglycosylase function, the intercalation of a stable but inactive EmtA/PhoA fusion protein into the complex would prevent any possible complementation and therefore result in inactivation of the complex. If this dominant-negative inactivation of BFP biogenesis was the explanation of the loss of BFP production in the CVD206::*TnphoA* mutants #13 and #137, then

overexpression of a similar EmtA/PhoA fusion protein in wild-type EPEC would be expected to result in a similar phenotype.

During the transposon sequencing of pRDH8 four *TnphoA* insertions were isolated whose insertion sites were determined, by mapping and DNA sequencing, to be within the *emtA* gene (Chapter 4 section 4.2.1.5; Appendix 2). These included two *phoA*-positive insertions, identified in plasmids pRDH8-30 and pRDH8-15, which encoded protein fusions to the initial 91 (273 bp) and 175 amino acids (525 bp) of the EmtA protein, respectively. In the other two *TnphoA*-carrying plasmids, pRDH8-w26 and pRDH8-w32, the transposons had inserted into *emtA* in the opposite direction so that they did not form *phoA* fusions; these two insertions were both mapped to the same site 409 bp downstream of the *emtA* initiation codon. These plasmids therefore provided the necessary EmtA/PhoA protein fusions and *TnphoA* non-fusion controls which were required in order to test the dominant-negative hypothesis described above. Plasmid DNA preparations of pRDH8 and the four *TnphoA*-carrying pRDH8 derivatives were used to transform EPEC strain E2348-69 *Strep*^r by electroporation; transformants were selected on LUA plates containing streptomycin, kanamycin, ampicillin and XP. Each of the strains were tested for localised adherence, FAS test and for BFP production using the protocols described previously. In each case the plasmid containing strains demonstrated wild-type phenotypes identical to the parental E2348-69 strain (data not shown). Plasmid-carrying strains initially appeared identical to parent in all respects; however, during the course of the virulence assays described above it was observed that they exhibited altered bacterial cell morphology. E2348-69 strains carrying the *emtA::TnphoA* plasmids demonstrated a high incidence of cell filamentation, which was not observed with either E2348-69 parent or the strain carrying the pRDH8 plasmid (Figure 5.17).

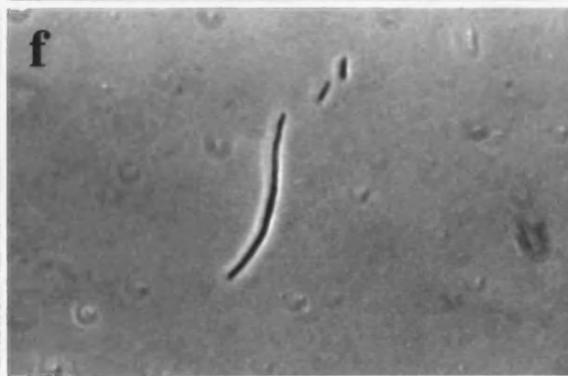
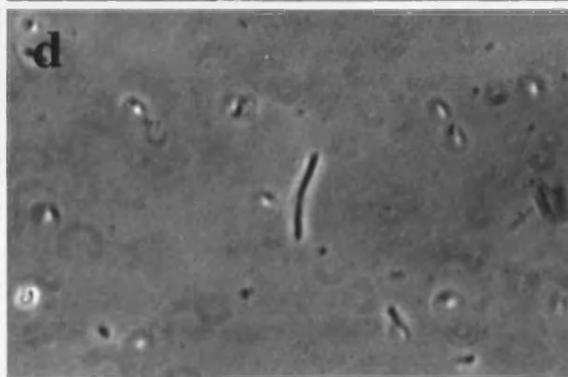
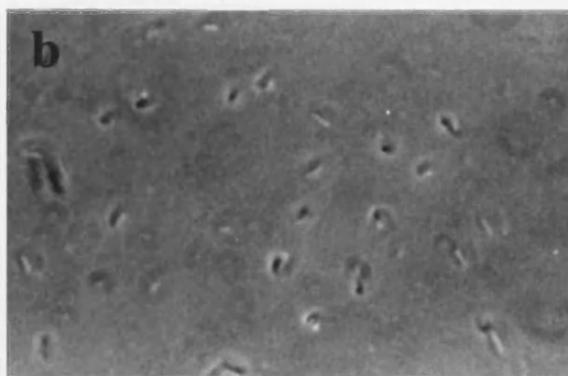
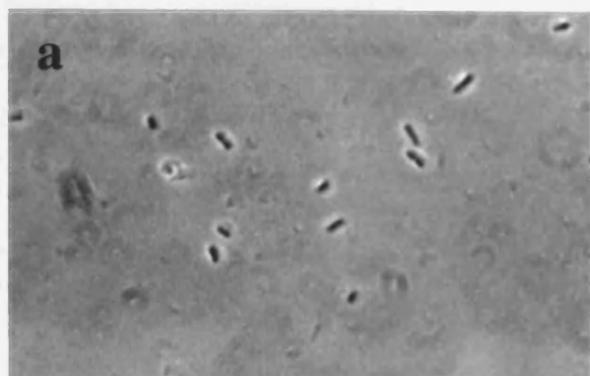
5.2.8 Phenotypic analysis of the IAP⁺ strain RDH006.

Generation of merodiploids in both CVD206 and E2348-69, using the suicide plasmids pRDH33, pRDH38 and pRDH39, had resulted in the isolation of strains which

exhibited a novel autoaggregation phenotype on growth in LUB. The IAP phenotype was maintained after resolution of the merodiploid and was independent of the outcome (i.e. mutant or wildtype) indicating that it was not a consequence of the mutations. In order to analyse the mechanism of the IAP phenotype, strain E2348-69 *Strep^r* was compared with the IAP⁺ strain RDH006 (E2348-69 *Strep^r* IAP⁺, section 5.2.5.2) for growth and in the standard EPEC virulence assays. RDH006 grown in LUB had an exponential growth rate equivalent to E2348-69 but this rate reproducibly slowed, in comparison with the parent strain, on entering stationary phase (**Figure 5.18**). Despite this slight growth defect, RDH006 appeared completely wild-type for localised adherence, FAS test and invasion assays (data not shown).

To assess the involvement of BFP in the IAP phenotype strains E2348-69 and RDH006 were transformed with p6.8.1, a derivative of the E2348-69 EAF plasmid pMAR2 (Nataro *et al.*, 1987); plasmid p6.8.1 contains a *TnphoA* insertion in the *bfpA* gene (p6.8.1 was isolated from strain 6.8.1; Donnenberg *et al.*, 1990; the *TnphoA* insertion site was determined to be identical to that in pRDH7-83; **Appendix1**). Transformation with p6.8.1 and selection for the kanamycin resistance encoded by *TnphoA* results in the isolation of EPEC strains which have lost the wild-type EAF plasmid pMAR2 and therefore are unable to make BFP. Strains E2348-69, RDH006 and their p6.8.1-carrying derivatives were tested for the production of BFP and the ability to autoaggregate in both LUB and DMEM (**Table 5.1**). As previously described E2348-69 was found to autoaggregate only in DMEM at 37°C, and this autoaggregation was dependent upon the expression of BFP. In contrast the loss of BFP production had no effect upon the ability of RDH006 to form aggregates in either LUB or DMEM, and RDH006 p6.8.1 was observed to retain the phenotype of rapid settling of bacteria in overnight LUB cultures on standing (**Figure 5.6A**). It was also observed that, unlike the expression of BFP, the IAP phenotype occurred at both 30°C and 37°C (**Table 5.1**).

Figure 5.17 Production of filamenting cells in EPEC strains carrying pRDH8 *emtA*::Tn*phoA* plasmids. EPEC strains grown overnight in LUB were examined for cell morphology by phase contrast microscopy. Panels (a) E2348-69, (b) E2348-69 pRDH8, (c) E2348-69 pRDH8-15, (d) E2348-69 pRDH8-w26, (e) E2348-69 pRDH8-30, (f) E2348-69 pRDH8-w32.



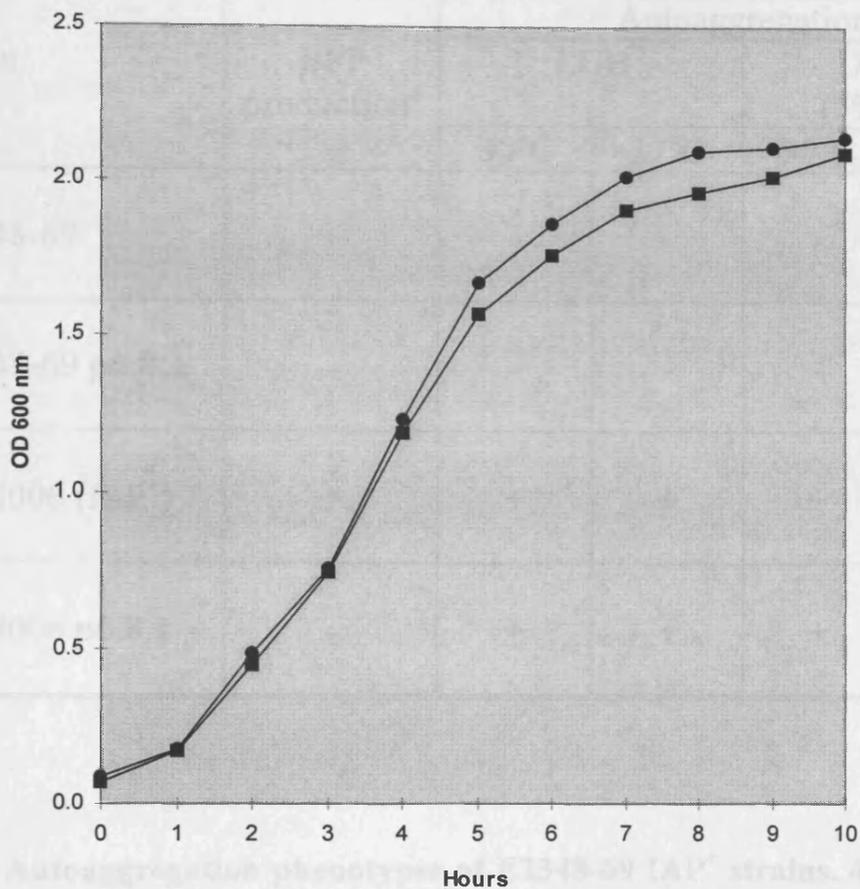


Figure 5.18 Growth characteristics of EPEC strain RDH006 (IAP⁺). Overnight LUB cultures of E2348-69 (●) and RDH006 (■) were diluted in LUB and grown with vigorous shaking. Measurements of OD₆₀₀ were taken hourly. This is a representative of three separate experiments.

Strain	BFP production ^a	Autoaggregation			
		LUB		DMEM (5% CO ₂)	
		30°C	37°C	30°C	37°C
E2348-69	+	-	-	-	+
E2348-69 p6.8.1	-	-	-	-	-
RDH006 (IAP ⁺)	+	+	+	+	+
RDH006 p6.8.1	-	+	+	+	+

Table 5.1 Autoaggregation phenotypes of E2348-69 IAP⁺ strains. 40 µl aliquots of overnight LUB cultures were inoculated into 1 ml of pre-warmed medium in individual wells of 24-well tissue culture plates and incubated for 2 hours. Autoaggregation was determined by microscopy. ^a production of BFP was assayed by immunofluorescence microscopy of bacteria adherent to HEp-2 cells using anti-BFP antibody (Giron *et al.*, 1993).

5.3 Discussion

The construction of non-polar mutations in genes identified by transposon mutagenesis strategies as bacterial virulence determinants is an essential requirement for formal proof of their role in pathogenesis. The difficulties inherent in the construction of non-polar non-selectable mutations in many Gram-negative bacteria have led to the recent emergence of high efficiency mutation strategies which are dependent on the use of counterselectable markers (Stibitz, 1994; Reyrat *et al.*, 1998). The *Bacillus subtilis* *sacB* gene, which confers sucrose sensitivity in many Gram-negative bacteria, has been used successfully as a counterselectable marker in EPEC in the construction of mutations in a number of virulence genes, including; *eaeA*, *espB*, *espA*, *bfpB*, *espD*, and recently *bfpC*, *bfpU* and *bfpH*, (Donnenberg and Kaper, 1991; Donnenberg *et al.*, 1993; Kenny *et al.*, 1996; Ramer *et al.*, 1996; Lai *et al.*, 1997; Anantha *et al.*, 1998). In this study, mutations were constructed in the *emtA* and *ompX* genes of EPEC strain E2348-69, or its derivative CVD206, by allelic exchange using the *sacB*-containing suicide vector pRDH10.

During the construction of the EPEC mutations significant numbers (up to 50%) of merodiploid strains were isolated which exhibited an increased autoaggregation phenotype (IAP) in both LUB and DMEM. Intriguingly these merodiploid strains were observed to retain the IAP⁺ phenotype even after resolution of the merodiploid and irrespective of the final outcome (i.e. wildtype or mutant) indicating that the mutation itself was not relevant to autoaggregation (similar results have also been observed when pRDH10 was used to create a mutation in the *typA* gene of strain E2348-69; S. Clarke, personal communication; Freestone *et al.*, 1998). IAP autoaggregation appeared visually distinct from the BFP-dependent EPEC autoaggregation described by Vuopio-Varkila and Skoolnik (1991); it also occurred in LUB at 30°C, when BFP are not normally expressed (Puente *et al.*, 1996). Furthermore autoaggregation of IAP⁺

strains was unaffected by the introduction of a *bfpA::TnphoA* mutation indicating that BFP are not required for this phenotype.

An IAP⁺ derivative of E2348-69 NaI^r (RDH001) has been isolated by chance from a culture of E2348-69 NaI^r streaked to single colonies (R. Haigh, unpublished data) and other EPEC strains exhibiting autoaggregation in LUB have been observed previously (T. Baldwin, personal communication) indicating IAP⁺ variants may also occur naturally. It was proposed that a low background level of IAP⁺ bacteria might exist within otherwise apparently phenotypically homogenous strains; however, attempts to quantitate a level of IAP⁺ expressing bacteria within stock EPEC strains, by the individual examination of 1000 separate colonies, have been unsuccessful (R. Haigh, unpublished data). Even if it is assumed that IAP⁺ bacteria occur naturally at a background level of approximately 1 in 1000, the rate of appearance of IAP⁺ strains during merodiploid construction still represents a massive enrichment and therefore indicates a strong selective pressure. It is unclear how this selection could occur; however, it does seem feasible that the increased aggregative ability of an IAP⁺ bacterial cell might increase conjugative ability and therefore the probability of the formation of merodiploids. This hypothesis has yet to be tested.

An autoaggregation phenotype similar to IAP has been previously reported in *E. coli* K-12 strains and has been associated with the metastable *flu* gene mapped to 43.6 min (Diderichsen, 1980). *flu* strains were identified by their ability to aggregate, "fluff" and sediment when liquid cultures were allowed to stand after shaking. Diderichsen (1980) also demonstrated that the *flu* mutation affected colony morphology in *E. coli* K-12 strains probably by an alteration in the production of type1 pili; *flu*⁺ strains were piliated and formed glossy colonies whereas *flu* strains were poorly piliated and had a frizzy colony morphology. It was observed that in *E. coli* K-12 strains the aggregation and fluffing property of strains could change at high frequency and that many strains contained both fluffing and nonfluffing variants. This "phase variation" was similar to that observed for the transition between piliated and non-piliated states in both *E. coli*

and *Salmonella typhimurium* (Brinton, 1959; Old *et al.*, 1968; Eisenstein, 1981; Eisenstein, 1988) and therefore Diderichson (1980) concluded that *flu* might be involved in type 1 pili regulation or biogenesis. More recently it has become apparent that the *flu* locus itself may encode proteins directly responsible for the autoaggregation observed (Owen *et al.*, 1996). Owen *et al.* (1987) identified an *E. coli* outer membrane protein, termed antigen 43 (Ag43), which exhibited reversible phase variation; the rates of variation in liquid media from Ag43⁺ to Ag43⁻ and from Ag43⁻ to Ag43⁺ states are 2.2×10^{-3} and 1×10^{-3} , respectively. Henderson *et al.* (1997) showed that *E. coli* K-12 strains containing mutations in deoxyadenosine methylase (Dam) were unable to express Ag43, whereas strains with mutations in the transcriptional regulator gene *oxyR* (Dorman, 1994) were "locked ON" for Ag43 expression; this has similarities to the work by Warne *et al.* (1990) who identified a gene they designated *mor* (*mor* \equiv *oxyR*) which locked *E. coli* into the *flu* phenotype. It has therefore been speculated that the phase variation of Ag43 is dependent upon a balance between the methylation of GATC sites in the Ag43 gene regulatory region by Dam, and the binding of the OxyR protein to non-methylated DNA as a transcriptional inhibitor (Henderson *et al.*, 1997). Ag43 is composed of two protein subunits α^{43} (apparent M_r 60,000) and β^{43} (apparent M_r 53,000) which are encoded by genes which have been mapped to the *flu* locus (Henderson and Owen, 1996; Henderson *et al.*, 1997). The β^{43} subunit is an integral outer membrane protein which anchors the α^{43} subunit to the cell surface via its C-terminus, and antibody studies have clearly demonstrated that α^{43} extends beyond the smooth LPS layer and is accessible on the bacterial surface (Owen *et al.*, 1996). The N-terminal sequence of α^{43} contains a six amino acid motif which is also present in the N-termini of the major subunits of several enterobacterial fimbriae (Henderson *et al.*, 1997), and the complete α^{43} amino acid sequence shows 30% identity (72% similarity) with AIDA-1, an outer-membrane-protein adhesin from diffuse adhering *E. coli* (Owen *et al.*, 1996; Benz and Schmidt, 1992). It has therefore been suggested that Ag43 might have a role in bacterial adhesion; this hypothesis is

supported by initial experiments which show Ag43⁺ bacterial strains exhibit a 6- to 8-fold increase in mannose-insensitive adhesion to HEp-2 cells (Owen *et al.*, 1996).

Despite the obvious similarity of the *flu* and IAP⁺ phenotypes it remains unclear if the appearance of IAP⁺ EPEC merodiploids is due to an enrichment for Ag43⁺ phase-variants. Outer membrane protein changes have not been observed in IAP⁺ strains (K. Sankaran, personal communication) and there are no obvious differences in colony morphology between IAP⁻ and IAP⁺ EPEC strains; however this may represent strain-dependent differences as Diderichsen (1980) also reported *E. coli* K-12 *flu* strains which autoaggregated but which did not exhibit differences in colony morphology. IAP⁺ EPEC strains do not exhibit massively increased or decreased adhesion to HEp-2 cells; however, these experiments have not been repeated in BFP⁻ strains, and therefore it is currently impossible to equate these results with the data obtained by Owen *et al.* (1996) which used poorly adherent *E. coli* K-12 strains. E2348-69 is the EPEC O127:H6 type strain, and as such has been used by several research groups to create defined mutations using *sacB*-dependent allelic exchange vector systems similar to pRDH10 (see above). An initial examination of the E2348-69 mutant strains CVD206 (*ea*eA; Donnenberg and Kaper, 1991) and UMD864 (*espB*; Donnenberg *et al.*, 1993) determined that these strains are IAP⁻; however, given the incidence of IAP⁺ merodiploids in our laboratory it seems remarkable that the IAP⁺ phenotype has not been reported previously by other groups. Whilst there is no current evidence that the IAP⁺ phenotype affects the expression of virulence genes it has been determined that our laboratory policy will be to ensure that all mutations constructed are in an IAP⁻ background in accordance with the original E2348-69 parental strain.

5.3.1 A defined non-polar *emtA* mutant produces BFP.

A defined non-polar frameshift mutation in the *emtA* gene in EPEC strain E2348-69 was assessed for its effect upon the production of BFP. Normally BFP production in EPEC is inhibited during growth in LUB but maximally induced during the

exponential growth phase when cultured in DMEM (Giron *et al.*, 1991; Puente *et al.*, 1996). On transfer to DMEM the *emtA* mutant RDH005 showed an initial lag in the induction of expression of the BfpA protein (bundlin) when compared to wild-type EPEC; however, after three hours the levels of bundlin in both strains were approximately equivalent. Despite the difference in bundlin expression in the mutant there was no apparent effect upon the levels of bacterial adherence and invasion of HEp-2 cells after three hours of infection. Furthermore the production of BFP by RDH005 bacteria adherent to HEp-2 cells, as determined by immunofluorescence microscopy using a BFP-specific antibody (Giron *et al.*, 1993) was identical to that of the E2348-69 parent. This was in complete contrast to the phenotypes of the CVD206 *emtA::TnphoA* mutants #13 and #137 which produced apparently normal levels of bundlin, but which were negative for BFP production and showed massively decreased levels of adherence and invasion when compared to the CVD206 parental strain.

The most probable explanation of the phenotypic differences observed between the *TnphoA* and non-polar mutations in *emtA* is that of polar influence by the transposon upon the expression of another gene(s) in the *emtA* locus. However a second possible explanation was proposed based upon the inherent differences between the mutations i.e. that an *emtA* Δ *HindIII* mutation results simply in gene inactivation, whereas an *emtA/TnphoA* mutant produces a stable and active fusion protein (albeit alkaline phosphatase activity) which might retain an active N-terminal EmtA protein domain. There is substantial circumstantial evidence to suggest that peptidoglycan metabolism is maintained by multi-enzyme complexes or "holoenzymes" (Höltje, 1996a; Höltje, 1996b), therefore it is feasible to assume that the assembly of trans-envelope structures through the cell wall may also involve such holoenzymes. It was proposed that whilst the complete loss of EmtA might be complemented in a holoenzyme by another similar enzyme activity this might not occur in an *emtA/TnphoA* mutant. An EmtA/PhoA fusion protein would be inactive because it lacked the transglycosylase catalytic domain, but it would still be able to interact with the holoenzyme via its N-terminal domain therefore preventing interactions with complementing enzymes and

resulting in a dominant-negative inactivation of the complex. The negative dominant hypothesis was tested by analysis of BFP expression in EPEC strains in which EmtA/PhoA fusions were overexpressed. These strains were constructed by introduction of derivatives of multicopy plasmid pRDH8 into wildtype EPEC; two of these plasmids contained *TnphoA* insertions within the *emtA* gene which encoded *emtA/phoA* fusions, whilst the other two merely disrupted the *emtA* gene. Unfortunately all of the plasmid-containing EPEC strains were wildtype for BFP production, therefore indicating that the EmtA/PhoA chimeric protein does not appear to exert a dominant negative effect. Intriguingly each of the strains containing pRDH8 derivatives with *TnphoA* insertions in the *emtA* gene (irrespective of orientation) showed alterations in bacterial cell morphology compared to the parent. Exponential phase LUB cultures contained a significant number (10-20%) of bacteria which were 2-4 times the normal bacterial length indicating that these strains were experiencing difficulties in cell division. This abnormal cell morphology was apparently due specifically to a *TnphoA* related effect as it was not observed in equivalent strains carrying the wildtype pRDH8 plasmid. However, as an EPEC strain carrying the plasmid pRDH8-12 (*ompX::TnphoA*) was shown to have normal cell morphology it was clear that the phenotype was not due to a *TnphoA* encoded gene. Overexpression of bacterial proteins from multicopy plasmids can frequently have deleterious effects upon cell growth (Tabor and Richardson, 1985; O'Connor and Timmis, 1987; Dong *et al.*, 1995). It is presumed therefore that the insertion of *TnphoA* into *emtA* results in the induced alteration in the expression of a pRDH8 gene product which is deleterious for growth.

It is apparent from the normal BFP production by both RDH005 (*emtA* Δ *HindIII*) and strains over-expressing the *emtA/phoA* fusion that the phenotypes observed in the *TnphoA* mutants #13 and #137 are not due to a loss of EmtA function. The only remaining explanation is that the insertion of the transposon into *emtA* has a polar effect upon the expression of a neighbouring gene(s) which then somehow affects the biogenesis of BFP. This hypothesis is consistent with the failure of the plasmid

pRDH8 to complement the *emtA::TnphoA* mutants. The failure of the pRDH8 *emtA::TnphoA* plasmids to affect pili biogenesis points clearly to the polar effect being one which results in the down-regulation of a gene whose product is essential for pili biogenesis, rather than one which results in the up-regulation of a gene product which is deleterious. (In contrast the cell division defect in strains carrying pRDH8 *emtA::TnphoA* plasmid must be due to the increased expression of a gene; this is a logical assumption because the opposite possibility, i.e. the reduction in the expression of additional plasmid encoded copies of a gene, would be expected to be less deleterious. Based upon the similarity of the phenotypes of the *tag::TnphoA* mutant #181 and the mutants #13/#137, an obvious candidate for the role of the repressed gene must be *tag*; however, the failure of plasmid pRDH8 to complement the mutation in #181 may be interpreted to suggest that the phenotypes of the *emtA::TnphoA* and *tag::TnphoA* mutations could also be due to polar effects of the transposon insertions upon a third gene, e.g. *orf3*.

5.3.2 Future work on the *emtA* locus.

The data presented in this chapter have clearly shown that the role proposed for EmtA in BFP biogenesis was incorrect and have demonstrated again the dangers inherent in the interpretation of phenotypes presented by transposon mutants. EmtA is undoubtedly involved in cell wall metabolism; this is demonstrated by the changes in the length of glycan strands and the increase in crosslinking within the murein observed in an *E. coli* K-12 strain overexpressing the *emtA* gene (Kraft *et al.*, 1998). Further understanding of the exact role of EmtA may come from the analysis of the peptidoglycan structure of an *emtA* Δ *HindIII* mutation (this work) of the *E. coli* K-12 strain MC1061 (Casbadan and Cohen, 1980; Kraft *et al.*, 1998) currently being undertaken in collaboration with the laboratory of Prof. Joachim-Volker Höltje in Tübingen.

Despite the wildtype production of BFP in an *emtA* Δ *HindIII* EPEC mutant, the observations made by Dijkstra and Keck (1996b) still hold and their hypothesis that lytic transglycosylases are involved in assembly and function of trans-envelope structures remains valid. Despite the recent report by Anantha *et al.* (1998a) which described a non-polar mutation in the lytic transglycosylase homologue gene *bfpH* (Kim and Kamono; 1997) to be unaffected in pili biogenesis, the possibility that BfpH is still somehow required for BFP function cannot be excluded. A second gene in the *bfp* operon, *bfpF*, has also been shown to be non-essential for pili biosynthesis (Anantha *et al.*, 1998b; Bieber *et al.*, 1988), yet in human volunteer studies *bfpF* mutants have been found to be 200-fold less virulent than the parent (Bieber *et al.*, 1988). This indicates that some BFP proteins are required for functions that occur post pili assembly. This is consistent with the role of the lytic transglycosylase homologue P19 of the conjugative plasmid R1; P19 is not required for assembly of the conjugative pili, but it is required for the efficient transfer of the plasmid DNA (Bayer *et al.*, 1995). Defining a role for BfpH in BFP biogenesis will first require the confirmation of its transglycolytic activity by overexpression, purification and analysis of its peptidoglycan metabolising activity.

The nature of the polar effect of the *emtA::TnphoA* insertion remains unclear, however the data suggest that it is due to the negative regulation of an adjacent gene. Therefore it should be possible to determine which gene is affected by construction of non-polar knockouts of the surrounding genes i.e. *tag*, *orf3* and possibly also *orf4*. The quick and simple method of non-polar mutant construction utilised in this study was dependent upon the use of conveniently positioned restriction enzyme sites to create frameshifts. The major limitations with this method are the requirement for an infrequently occurring restriction enzyme site and the necessity of being able to centrally position the mutated site within the homologous DNA fragment for efficient allelic exchange. Therefore for future work it is proposed to use the rapid inverse PCR mutagenesis (IPCRM) method of Wren *et al.* (1994). As IPCRM is not limited by the requirement

for pre-existing restriction enzyme sites it is possible to introduce frameshift or termination mutations at any specified point within the gene in question.

5.3.3 Regulation of the *ompX* operon.

It was proposed in Chapter 4 that the *orf5'-orf1-orf2-ompX* sequences form the distal genes of a novel *E. coli* siderophore uptake system. This hypothesis was based upon the observed sequence homology of the OmpX and Orf5' polypeptides with protein components of known high affinity iron uptake systems from *E. coli* and other Gram-negative bacteria. Whilst iron is a requirement for the synthesis of essential *E. coli* proteins an excess of iron can also be potentially deleterious to the cell. Hydroxyl free radicals generated through Haber-Weiss-Fenton chemistry, catalysed by iron, can accumulate resulting in damage to DNA and biological membranes, and ultimately in cell death (Flitter *et al.*, 1983; Weinberg, 1989; Wooldridge and Williams, 1993). Consequently bacterial cells have developed specific regulatory systems to control the production of proteins which are involved in both iron uptake and storage. Iron itself is directly involved in the principle regulatory system which controls iron uptake mechanisms in *E. coli*. In high iron concentrations the 17 kDa Fur protein (ferric uptake regulation; Hantke, 1984; Shäffer *et al.*, 1985) complexes with Fe^{2+} as a dimer and binds specifically to operator sequences of operons involved in iron uptake thereby repressing transcription; conversely in low iron concentrations the Fur/ Fe^{2+} complex cannot form resulting in the expression of iron uptake genes (Bagg and Neilands, 1987a; Bagg and Neilands, 1987b; de Lorenzo *et al.*, 1987; de Lorenzo *et al.*, 1988a; de Lorenzo *et al.*, 1988b; Griggs and Konisky, 1989). Analysis of the sequences of *E. coli* iron-regulated promoters has identified a 19 bp consensus Fur-binding sequence, which is generally located between the -35 and -10 regions (Fur box; GATAATGATAATCATTATC; de Lorenzo *et al.*, 1987; Pressler *et al.*, 1988). Furthermore this consensus sequence has been shown to function as an operator site by

insertion of a synthetic oligonucleotide containing the Fur box within the promoter sequence of a reporter gene (Calderwood and Mekalanos, 1988). *fur* gene homologues have been cloned from a variety of Gram negative bacteria and putative Fur box sequences have been identified in the promoters of numerous iron-regulated genes (Wooldridge and Williams, 1993; Earhart, 1996; Crosa, 1997; Byers and Arceneaux, 1998).

If the *ompX* gene encodes an outer membrane siderophore uptake protein, as predicted in Chapter 4, then it would be expected that its expression would be negatively regulated by the Fur repressor in response to high iron concentrations. Analysis of the pRDH8 and partial pRDH45 DNA sequences failed to identify sequences homologous with the consensus Fur box sequence; however there is also no evidence for the consensus -35 and -10 sequences indicative of *E. coli* promoters within these sequences. It has therefore been proposed that the promoter for *ompX* is upstream of *orf5*. Allelic exchange was used to introduce two *ompX* mutations into EPEC strain CVD206 in order to analyse both the role and regulation of the *ompX* operon. In the first mutant, RDH003, a non-polar *ompX* mutation was constructed by introduction of a frameshifted *ompX* allele which encoded a polypeptide truncated after 111 amino acids. In the second mutant, RDH004, the wildtype *ompX* allele was exchanged for an *ompX/phoA* fusion which had been previously isolated as a *phoA*-positive Tn*phoA* insertion derivative of plasmid pRDH8 (pRDH8-12).

Both RDH003 and RDH004 were completely unimpaired in their ability to grow in culture media containing limited iron. This result was consistent with the previous report that strain E2348-69, whilst aerobactin negative, is able both to produce and utilise the siderophore enterobactin (Law *et al.*, 1992). As predicted for an iron repressed gene, *ompX* expression was represented by a low level of *phoA* expression in both exponential and post-exponential growth phases during growth of the *ompX/phoA* reporter strain RDH004 in iron-replete media. The level of alkaline phosphatase activity in the bacterial cell fluctuated significantly during growth;

however, after subtraction of the background, the activity attributable to the OmpX/PhoA fusion was relatively constant (average 2.02 U/ml \pm 0.45). Intriguingly the growth of RDH004 in iron restricted media resulted in a decrease in alkaline phosphatase activity; this was partly due to a decrease in the background cellular alkaline phosphatase activity but was also apparently due to a reduction in the expression of the *ompX/phoA* fusion (average 1.24 U/ml \pm 0.39). There are several possible alternative explanations for the lack of induction of the *ompX* operon in low iron. It is possible that the *ompX* operon is cryptic; however, this seems unlikely due to the observed low but constitutive alkaline phosphatase activity in RDH004. It seems more probable that the results observed are due to a requirement for a specific positive regulatory signal, in addition to iron limitation, for the induction of transcription of the *ompX* operon. This secondary level of regulation is a feature which is being found amongst an increasing number of iron uptake systems in Gram-negative bacteria, and is usually dependent upon the presence of the siderophore (Braun, 1995; Braun, 1997; Crosa *et al.*, 1997).

The best characterised of the siderophore-inducible uptake systems is the ferric dicitrate (*fec*) operon of *E. coli* (Braun, 1997). The FecA siderophore receptor is only observed in the outer membrane of iron-starved bacteria if 0.1 mM citrate is added to the medium (Frost and Rosenberg, 1973; Hancock *et al.*; 1976; Wagegg and Braun, 1981; Zimmermann *et al.*, 1984; Pressler *et al.*, 1988). Whilst the induction of *fec* is dependent upon the function of the FecA, TonB, ExbB and ExbD proteins, it is apparent that the ferric dicitrate does not need to enter the cytoplasm because the function of the FecBCDE inner membrane permease proteins are not required (Hussein *et al.*, 1981; Wriedt *et al.*, 1995). Furthermore recent experiments with point mutations in *fecA* indicate that the binding of ferric dicitrate to the receptor, irrespective of whether it is transported to the periplasm, is sufficient for signal transduction across the bacterial envelope (Braun, 1995; Harle *et al.*, 1995; Braun, 1997). The inducible expression of FecA is dependent upon two further proteins encoded upstream of the *fecABCDE* operon, FecR and FecI (Pressler *et al.*, 1988, Staudenmaier *et al.*, 1989; van

Hove *et al.*, 1990; Ochs *et al.*, 1995). It is proposed that the binding of ferric dicitrate to FecA, through the activity of the TonB-ExbB-ExbD complex (Skare and Postle, 1991; Ahmer *et al.*, 1995; Braun, 1995), triggers a conformational change in the inner membrane protein FecR which then activates the cytoplasmic protein FecI (Crosa, 1997). FecI is a DNA binding protein, which has significant homology with the recently described extracytoplasmic function (ECF) alternative sigma factors (Missiakis and Raina, 1998), and has been shown to interact directly with the *fecA* promoter inducing transcription of the *fecABCDE* operon (Angerer *et al.*, 1995; Enz *et al.*, 1995; Ochs *et al.*, 1996; Kim *et al.*, 1997; Angerer and Braun, 1998).

In addition to *fec* there are also well characterised siderophore-inducible iron uptake systems reported from both *Pseudomonas* and *Vibrio* species; these systems demonstrate that the FecI/FecR regulators represent just one of a wide variety of positive regulation control mechanisms which are present in Gram-negative bacteria (Crosa, 1997). In the *Pseudomonas* species these regulatory mechanisms include: the uptake of pseudobactins BN7 and BN8 by PupB, which is controlled by PupI and PupR (homologues of FecI and FecR; Koster *et al.*, 1993; Koster *et al.*, 1994); the uptake of enterobactin by PfeA, which is controlled by the two component regulator homologues PfeS and PfeR (Dean and Poole, 1993; Dean *et al.*, 1996); the uptake of pyochelin by FptA, which is regulated by the AraC-like PchR (Heinrichs *et al.*, 1991; Heinrichs and Poole, 1993; Heinrichs and Poole, 1996); and the uptake of pyoverdine by FpvA which requires the σ^E -like factor PvdS (Cunliffe *et al.*, 1995; Miyazaki *et al.*, 1995). It is also apparent that in *Pseudomonas* species which contain multiple siderophore uptake systems there is a regulatory hierarchy; this regulatory system interaction ensures that under conditions in which multiple siderophores are available to the cell it is the uptake system whose siderophore has the highest affinity for iron which is preferentially expressed (Crosa, 1997).

It still remains to be proven whether the expression of *ompX* is under the control of a positive regulatory system and/or Fur; such information will only become available

with the complete determination of the DNA sequence of the *ompX* operon and the identification of consensus regulator binding sites within the promoter sequences.

5.3.4 Future work: Identification of a siderophore for the OmpX receptor.

Based upon the data determined in this study it is proposed that the *ompX* operon encodes an inducible ferrisiderophore uptake system. An important factor in the characterisation of this operon must therefore be the identification of the siderophore(s) which it can utilise. *E. coli* strains produce only two siderophores, enterobactin and aerobactin (Earhart, 1996); therefore it must be presumed that the siderophore(s) recognised by the OmpX receptor would have to be supplied exogenously as is the case for the hydroxamate siderophores utilised by receptors FhuA and FhuE (Koster, 1991).

OmpX has no obvious equivalents in other bacterial species; furthermore phylogenetic comparison of the OmpX polypeptide sequence with those of the other *E. coli* siderophore receptors failed even to clearly assign it within the hydroxamate or catechol receptor subfamilies. It is therefore proposed to construct assay strains which will allow the screening of large numbers of potential siderophore candidates in collaboration with the laboratory of Dr Rolf Reissbrodt in Wernigerode. The screen can take two possible forms: (i) growth induction under iron limiting conditions of an *E. coli* K-12 *fepA* strain, which cannot utilise enterobactin (LG1466; Williams, 1979), carrying the *ompX* operon-encoding plasmid pRDH45; or (ii) induction of alkaline phosphatase activity in the *ompX/phoA* reporter strain RDH004 under iron limiting conditions. For the latter screen *ompX* will also be provided *in trans* (by introduction of plasmid pRDH8 or pRDH45) as it is predicted that, in common with FecA and the outer membrane receptors for the *Pseudomonas* siderophore uptake systems, OmpX will be essential for the transduction of the trans-envelope signal which up-regulates its own transcription.

Chapter 6. General Discussion.

Rapid developments in molecular techniques in the twenty years since the unequivocal confirmation of the diarrhoeagenic potential of EPEC strains (Levine *et al.*, 1978) have allowed significant progress in the elucidation of the underlying pathogenic mechanisms of EPEC disease (Kaper, 1994; Donnenberg *et al.*, 1997a; Kaper, 1998; Nataro and Kaper, 1998). It is increasingly apparent, as the virulence determinants of more pathogens are analysed, that type III secretion systems (Lee, 1997; Hueck, 1998) and type IV pili (Manning and Meyer, 1997) represent elements common to many Gram-negative bacterial virulence strategies (Finlay and Falkow, 1997; Elliot *et al.*, 1998b). The cloning and complete determination of the DNA sequences of both the *bfp* operon (Sohel *et al.*, 1996; Stone *et al.*, 1996) and the LEE pathogenicity island (Elliot *et al.*, 1998b) have allowed the demonstration that these regions contain all of the genes required either for BFP production (Stone *et al.*, 1996), or for AE lesion formation (McDaniel and Kaper, 1997). However, it still remains to be determined if there are additional genes, not encoded within these regions, which are required for the full process of EPEC disease.

The initial aim of this work was the isolation and characterisation of transposon mutants in EPEC which were unable to induce the serine/threonine phosphorylation of myosin light chain (MLC) which is observed during infection of cultured human epithelial cells. MLC phosphorylation, which has been implicated in the opening of tight junctions resulting in increased intestinal permeability, occurs due to the activation of MLC kinase (MLCK) in a manner which is independent of PKC activity, tyrosine phosphorylation or AE lesion formation (Haigh *et al.*, 1995; Hecht *et al.*, 1996; Yuhan *et al.*, 1997); therefore it was hoped that such mutants would identify genes involved in other aspects of EPEC pathogenesis.

6.1 EPEC MLC phosphorylation-negative mutants.

Each of the five EPEC MLC phosphorylation-negative transposon mutants identified and characterised in this work were found to be deficient in the assembly of the bundle-forming pilus. Whilst BFP have recently been shown to be essential for EPEC virulence (Beiber *et al.*, 1998), there are no published data to suggest that any bacterial pili are responsible for induction of host signal transduction processes of the type which might lead to MLCK activation. Two of the mutations were found to be localised within the gene encoding the outer membrane protein BfpB (Sohel *et al.*, 1996; Stone *et al.*, 1996) which has been shown to be required for production of BFP (Ramer *et al.*, 1996). The other three mutations defined two genes, in a chromosomal locus, which encode a novel endo-specific lytic transglycosylase, EmtA (Kraft *et al.*, 1998), and an uncharacterised predicted inner membrane protein, Tag.

It was proposed, based upon the hypothesis of Dijkstra and Keck (1996b) which predicts the involvement of lytic transglycosylase homologues in the construction of transenvelope structures in Gram-negative bacteria, that EmtA was required for assembly of BFP through the peptidoglycan cell wall. However, the construction of a non-polar mutation within the *emtA* gene of EPEC strain E2348-69 was observed to have no effect upon pili biosynthesis. It is now proposed that the phenotypes observed in the original EPEC *emtA* transposon mutants were the result of polar effects upon the expression of an adjacent gene(s), e.g. the *tag* gene. Despite this failure to confirm the hypothesis of Dijkstra and Keck (1996b), recent DNA sequence data have also suggested a role for lytic transglycosylase homologues in the assembly or function of transenvelope structures involved in EPEC virulence. As discussed in chapter 5, the gene for BfpH, encoded within the *bfp* operon (Sohel *et al.*, 1996; Stone *et al.*, 1996), shows significant homology with the *E. coli* lytic transglycosylase family (Kim and Kamono, 1997). It is unclear as yet exactly what role BfpH plays in BFP synthesis or function (Anantha *et al.*, 1998a). In addition Elliot *et al.* (1998b) report the identification of an open reading frame, *rorf3*, within the LEE whose predicted protein

product shows significant homology to the virulence plasmid-encoded, protein secretion-associated proteins IagB and IpgF, from *Salmonella* and *Shigella* respectively (Allaoui *et al.*, 1993; Miras *et al.*, 1995); these proteins were previously identified by Dijkstra and Keck (1996b) as putative lytic transglycosylases. Furthermore the predicted *rorf3* gene product contains the 3 amino acid motifs which are present in all characterised lytic transglycosylases and which are predicted to be located within the active site (Dijkstra and Thunnisen, 1994). Based upon its location adjacent to the *escRSTU* genes (previously *sepIGHF*; Jarvis *et al.*, 1995), which encode components of the type III secretion apparatus, it seems probable that this lytic transglycosylase homologue is involved in either the construction or the function of the EPEC type III secretion system.

6.2 Is there a role for OmpX in *E. coli* virulence?

All prokaryotic organisms, with the exception of the lactic acid bacteria (Archibald, 1983), have an absolute requirement for micromolar concentrations of iron for growth (Weinberg, 1978). In the hostile environment of the host tissues the level of extracellular iron is maintained well below micromolar concentrations by the action of the iron binding proteins transferrin and lactoferrin (Weinberg, 1984; Bezkorovainy, 1987; Brock 1989; Weinberg, 1989), therefore presenting a major nutritional difficulty for pathogenic bacteria during infection. In addition, in response to the presence of bacteria, the host can further reduce the already low iron concentration by a series of mechanisms collectively termed the hypoferraemic response (Weinberg, 1978; Weinberg, 1984). As the ability to acquire iron must obviously play a crucial role in the establishment and progression of bacterial infections it might be expected that the siderophore-based high affinity iron uptake systems found in many bacterial species (Matzanke, 1991; Winkelmann; 1991) would be obvious candidate virulence determinants.

Whilst there is evidence for a role for siderophores in bacterial infections there are apparent differences in the importance of individual siderophore uptake systems between bacterial species (Wooldridge and Williams, 1993). It has been unequivocally demonstrated the presence of exogenous siderophore in animal hosts infected with strains of *Salmonella*, *Vibrio vulnificus*, *Yersinia* or *Pseudomonas aeruginosa* results in enhanced bacterial virulence (Jones *et al.*, 1977; Write *et al.*, 1981; Cox, 1982; Melby *et al.*, 1982; Robins-Browne and Prpic, 1985); however, several *in vivo* studies using bacterial strains deficient in their endogenous siderophore uptake systems present a more complicated picture. The presence of a plasmid encoding the genes for the synthesis and uptake of the siderophore aerobactin correlates well with virulence in pathogenic *Klebsiella pneumoniae* (Nasif and Sansonetti, 1986) and in *E. coli* UTI strains (Montgomerie *et al.*, 1984; Carbonetti *et al.*, 1986; Linggood *et al.*, 1987; Jacobson *et al.*, 1988). Furthermore introduction of a plasmid encoding aerobactin synthesis/uptake into laboratory strains of *E. coli* has been shown to enhance their virulence in mice (Williams, 1979; Roberts *et al.*, 1989). However, whilst deletion of the aerobactin synthesis/uptake genes in *Shigella* inhibited growth in iron limiting conditions *in vitro* it was found to have minimal effects upon virulence in *in vivo* models (Lawlor *et al.*, 1987; Nassif *et al.*, 1987). It has been proposed that this difference in the requirement for aerobactin is because *Shigella* is a predominantly intracellular pathogen and may therefore utilise host haem compounds (probably via the *shu* haem uptake system; Wyckoff *et al.*, 1988) for their iron requirements while they are within the host cells (Lawlor *et al.*, 1987; Wooldridge and Williams, 1993). There are also conflicting data about the role of enterochelin in bacterial infections. Yancey *et al.* (1979) reported that *Salmonella typhimurium* enterochelin synthesis (*ent*) mutants were less virulent for mice; however, in contrast, later investigations found that whilst *ent* mutants were unable to grow in mouse serum their virulence in mouse models was not massively altered (Benjamin *et al.*, 1985). The extent of the role of enterochelin in *E. coli* pathogenesis is currently unknown, though enterochelin synthesis has been reported in experimental *E. coli* infections (Griffiths and Humphreys, 1980). It is still unclear to what extent iron acquisition systems contribute

to the virulence of all pathogenic bacteria; however, it has been observed that expression of a number of bacterial envelope proteins is enhanced *in vivo* in the iron limiting conditions of the host and that many of these proteins are ferric-siderophore receptors (Griffiths *et al.*, 1983; Sciortino and Finkelstein, 1983; Brown *et al.*, 1984; Lam *et al.*, 1984; Shand *et al.*, 1985).

E. coli has seven iron-regulated TonB-dependent siderophore receptors which recognise a range of endogenous and exogenous siderophores of both the catecholate and hydroxamate families (Hantke, 1981; Earhart, 1996). In addition it has recently been reported that some pathogenic *E. coli* contain a chromosomally encoded haem uptake system (*chu*) analogous to that encoded by the *shu* operon found in *Shigella* (Mills and Payne, 1995; Torres and Payne, 1997; Wyckoff *et al.*, 1988). There are at present no uncharacterised *E. coli* siderophore uptake activities which are documented in the extensive literature on this subject. However, the EPEC OmpX protein identified in this work shows significant homology with the other *E. coli* siderophore receptors; furthermore the other genes in the *ompX* operon have homologues in iron acquisition systems thereby indicating a likely role for OmpX in ferric-siderophore uptake. As it has been demonstrated that EPEC strain E2348-69 can only synthesise enterochelin (Law *et al.*, 1992) it must be concluded that any siderophore taken up by OmpX would have to be supplied exogenously. At first sight a system for the uptake of exogenous ferric-siderophore might appear unlikely to be an obvious virulence determinant. There is however a precedent; it has recently been reported that the outer membrane receptor for exogenously supplied ferrioxamine in *Salmonella enterica* serovar Typhimurium, FoxA (Tsolis *et al.*, 1995), is essential for virulence in an LD₅₀ mouse model (Kingsley, 1997).

Future work will require the determination of the complete DNA sequence of the *ompX* operon in EPEC strain E2348-69 and an investigation of the mechanism by which gene expression is regulated. These sequence data, combined with further analyses of the distribution of the *ompX* locus in phylogenetically aligned strains of *E.*

coli, should help to clarify the current confusion with regard to the existence of an *ompX*-containing pathogenicity island. Ultimately, confirmation of a role for OmpX in *E. coli* virulence will require testing an *ompX* mutant in a suitable animal model.

6.3 Concluding remarks.

Recent work has caused speculation that the phosphorylation of MLC by MLCK (Baldwin *et al.*, 1990; Manjarrez-Hernandez *et al.*, 1991; 1996) may result in the increased permeability of tight junctions and therefore contribute to the promotion of diarrhoea during EPEC infection (Yuhan *et al.*, 1997; Kaper, 1998). Whilst the data presented in this study demonstrate that efficient adhesion of bacteria to tissue culture cells is important in the induction of MLC phosphorylation, they do not identify the specific mechanism by which it occurs. It had been previously observed that there was a strict correlation between the ability of an *E. coli* strain to cause AE lesion formation and its ability to induce phosphorylation of P29 and MLC (Baldwin *et al.*, 1990); however, it was apparent from further work (Haigh *et al.*, 1995) that lesion formation itself was not specifically required for EPEC-induced protein phosphorylation. AE lesion formation also correlates with the presence of the LEE pathogenicity island and indeed it has been demonstrated that transfer of the E2348-69 LEE region to a laboratory strain of *E. coli* is sufficient in itself to allow lesion formation (Kaper and McDaniel, 1997). Based upon these observations it seems reasonable to propose that the MLCK activity which occurs during EPEC infection of host cells is induced by a virulence determinant encoded within the LEE. As the LEE encodes 41 predicted orfs, of which approximately 20 are currently of unconfirmed function (Elliot *et al.*, 1998b), there are many possible candidate genes. Preliminary data have suggested that the secreted protein EspB is not required for MLC phosphorylation (R. Haigh, unpublished observations); however, it is still unknown if the other secreted proteins or the EPEC type III secretion pathway are required. The recently discovered secreted protein EspF is unlike the other known EPEC secreted proteins in that it is not required for AE lesion formation (McNamara and Donnenberg, 1998) and therefore may be an obvious

candidate for involvement in MLCK activation (which is also independent of AE lesion formation). EspF has been identified to contain proline-rich protein motifs indicative of SH3 (Src homology 3) binding domains (Mayer and Eck, 1995; McNamara and Donnenberg, 1998); these domains are normally found on cytoskeletal proteins and are involved in interactions between cytoplasmic tyrosine kinases and their substrates. Proteins with proline-rich SH3-binding domains have been previously implicated in the subversion of eukaryotic intracellular signalling systems by a number of pathogenic microorganisms (Bliska, 1996). However, as activation of MLCK by EPEC has been determined to be independent of host tyrosine kinase activity (Yuhan *et al.*, 1997) this casts some doubt on a role for EspF in this process.

Proof of the role of determinants encoded within the LEE in MLCK activation and the resultant increases in the permeability of epithelial cell tight junctions will require the demonstration of such abilities in a laboratory *E. coli* strain containing the entire cloned region (McDaniel and Kaper, 1997). Armed with this result the current strategy of systematic mutagenesis being employed with all of the LEE open reading frames (Elliot *et al.*, 1998a) should rapidly identify the genes responsible.

Appendix 1. The annotated sequence of pRDH7.

The 5347 bp DNA sequence of the *Bam*HI fragment from pRDH7 is presented in the same orientation as the open reading frames are depicted in Figure 4.4. The predicted Shine-Delgarno sequences, start codons and termination codons of each open reading frame are indicated in bold and underlined. The deduced amino acid sequence of each open reading frame is shown below the nucleotide sequence; where appropriate signal sequences are indicated by underlining. The exact location of the transposon insertion sites in the CVD206::*TnphoA* mutants #55 and #138, and the pRDH7::*TnphoA* plasmids used in the DNA sequencing (see also Figure 4.2) are indicated by vertical arrows (↓ indicates the nucleotide before the insertion).

BamHI
GGATCCTTTTCTGCTCAAATCCTCCATGAAGCCAGTCATGACGCGGGCTGAGGGCTGGT
1 -----+-----+-----+-----+-----+-----+-----+ 60
CCTAGGAAAAAGACGAGTTTAGGAGGTA CTTTCGGTCAGTACTGCGCCCGACTCCCGACCA
TTACGGTTTTTGGGAGAAATGGGGTAATGCCAGGAGTACCGGAAGTCAAATTCATGGGGA
61 -----+-----+-----+-----+-----+-----+ 120
AATGCCAAAAACCCTCTTTACCCCATACGGTCCTCATGGCCTTCAGTTTAAGTACCCCT
TATTGACTGTTTATGTACTGGGGGGGACGGAAACACACAAAAAAGAAAAAAGAT
121 -----+-----+-----+-----+-----+-----+ 180
ATAACTGACAAATACATGACCCCCCTGCCTTTGTGTGTTTTTTTTTTCTTTTTTTCTA
TATTTTTTTTCTTGGTGCTTGCGTGTCTTTTTTAGTTTTAAGATTATTCCGTGACCTATT
181 -----+-----+-----+-----+-----+-----+ 240
ATAAAAAAAGAACCACGAACGCACAGAAAAAATCAAATTCATAAAGGCACTGGATAA
SD
AATACGGGGGTTTTATAAGG**AAACAG**TTTT**TATG**TTTCTAAAATCATGAATAAGAAAT
241 -----+-----+-----+-----+-----+-----+ 300
TTATGCCCCCAAATATTCCTTTTGTCAAATAACCAAAGATTTTAGTACTTATTCTTTA

BfpA M V S K I M N K K Y
ACGAAAAAGGTCTGTCTTTGATTGAATCTGCAATGGTGCTTGCGCTTGCTGCCACCGTTA
301 -----+-----+-----+-----+-----+-----+ 360
TGCTTTTTCCAGACAGAACTAACTTAGACGTTACCACGAACGCGAACGACGGTGGCAAT
E K G L S L I E S A M V L A L A A T V T
CCGCAGGTGTGATGTTTTACTACCAGTCTGCGTCTGATTCCAATAAGTCGCAGAATGCTA
361 -----+-----+-----+-----+-----+-----+ 420
GGCGTCCACACTACAAAATGATGGTCAGACGCAGACTAAGGTTATTCAGCGTCTTACGAT
A G V M F Y Y Q S A S D S N K S Q N A I

TTTCAGAAGTAATGAGCGCAACGTCTGCAATTAATGGTCTGTATATTGGGCAGACCAGTT
 421 -----+-----+-----+-----+-----+-----+-----+ 480
 AAAGTCTTCATTACTCGCGTTGCAGACGTTAATTACCAGACATATAACCCGTCTGGTCAA
 S E V M S A T S A I N G L Y I G Q T S Y
 ATAGTGGATTGGACTCAACGATTTTACTTAAACACATCTGCAATTCCGGATAATTACAAAG
 481 -----+-----+-----+-----+-----+-----+-----+ 540
 TATCACCTAACCTGAGTTGCTAAAATGAATTGTGTAGACGTTAAGGCCTATTAATGTTTC
 S G L D S T I L L N T S A I P D N Y K D
 ATACAACAAACAAAAAATAACCAACCCATTTGGGGGGGAATTAATGTAGGTCCAGCAA
 541 -----+-----+-----+-----+-----+-----+-----+ 600
 TATGTTGTTTGTTTTTTTTATTGGTTGGGTAAACCCCCCTTAATTTACATCCAGGTCGTT
 T T N K K I T N P F G G E L N V G P A N
 ACAATAACACCGCATTTGGTTACTATCTGACGCTTACCAGGTTGGATAAAGCGGCATGTG
 601 -----+-----+-----+-----+-----+-----+-----+ 660
 TGTTATTGTGGCGTAAACCAATGATAGACTGCGAATGGTCCAACCTATTTGCGCGTACAC
 N N T A F G Y Y L T L T R L D K A A C V
 TTAGTCTTGCAACCTTGAACCTAGGTACTTCAGCGAAAGGCTACGGTGTTAATATCTCTG
 661 -----+-----+-----+-----+-----+-----+-----+ 720
 AATCAGAACGTTGGAACCTGAATCCATGAAGTCGCTTTCGATGCCACAATTATAGAGAC
 S L A T L N L G T S A K G Y G V N I S G
 GCGAAAATAACATTACATCATTGGTAATAGCGCTGATCAGGCTGCTAAATCGACTGCTA
 721 -----+-----+-----+-----+-----+-----+-----+ 780
 CGCTTTTATTGTAATGTAGTAAACCATTATCGCGACTAGTCCGACGATTTAGCTGACGAT
 E N N I T S F G N S A D Q A A K S T A I
 TTACTCCTGCTGAAGCGGCAACTGCATGTAAAAATACTGATTCAACCAATAAAGTTACAT
 781 -----+-----+-----+-----+-----+-----+-----+ 840
 AATGAGGACGACTTCGCCGTTGACGTACATTTTTATGACTAAGTTGGTTATTTCAATGTA
 T P A E A A T A C K N T D S T N K V T Y
 ATTTTATGAAGTAATCTACATAATACGCCCGTTTTGGCGTATTATATGGGAGGTATATG
 841 -----+-----+-----+-----+-----+-----+-----+ 900
 TAAAATACTTCATTAGATGTATTATGCGGGCAAACCCGCATAATATACCCTCCATATAC
 F M K * BfpG M
 TGAGGACAGTAATCCTTTTTTTTTTGTCTTTTGGTATCCTGTTTTACTCAGGCACAGGAAA
 901 -----+-----+-----+-----+-----+-----+-----+ 960
 ACTCCTGTCATTAGGAAAAAAAACAGAAAACCATAGGACAAAATGAGTCCGTGTCCTTT
R T V I L F F C L L V S C F T Q A Q E S
 ↓ -86/-88 PstI
 GTGCAAATAAAAACGAGAACTTTTTTCAGGGGACACAACGTCTGTTGCACTGCAGGCAA
 961 -----+-----+-----+-----+-----+-----+-----+ 1020
 CACGTTTATTTTTGCTCTTTGAAAAAAGTCCCCTGTGTTGCAGACAACGTGACGTCCGTT
 A N K N E K L F S G D T T S V A L Q A S
 GTGAAAGAACAATTAATAAATGCCAAAGCTGACTGGTCTGTAATAAAGGACAGACTTTAT
 1021 -----+-----+-----+-----+-----+-----+-----+ 1080
 CACTTTCTTGTTAATTTTTACGGTTTCGACTGACCAGACATTATTTTCCTGTCTGAAATA
 E R T I K N A K A D W S V I K G Q T L Y
 ATGGCGTTCTTTCCGACTGGGCGCAACAGAGCGGATGGACTGTATTATGGCATGCAAAAT
 1081 -----+-----+-----+-----+-----+-----+-----+ 1140
 TACCGCAAGAAAGGCTGACCCGCTTGTCTCGCTACCTGACATAATACCGTACGTTTAA
 G V L S D W A Q Q S G W T V L W H A N Y

*Pst*I ↓ -25

1141 ATAACTATACTTTAAATGCAT**CTGCAG**TGTTTAAAGGATATGTGCGCTGGATGATGCAGCAA 1200
 -----+-----+-----+-----+-----+-----+-----+
 TATTGATATGAAATTTACGTAGACGTCACAAATTCCTATACAGCGACCTACTACGTCGTT
 N Y T L N A S A V F K D M S L D D A A K

1201 AGAAACTACTAGAATCGATGGGAAATGTGTCACCCAGGGTCTTTATTAAATTTTCATAACG 1260
 -----+-----+-----+-----+-----+-----+-----+
 TCTTTGATGATCTTAGCTACCCTTTACACAGTGGGTCCAGAAATAATTTAAAGTATTGC
 K L L E S M G N V S P R V F I K F H N G

SD

1261 GGAATCATGTCATTCTCGTTACACAATCACCC**GAAGGATAA**TTTAAAATGAAACTTGGCA 1320
 -----+-----+-----+-----+-----+-----+-----+
 CCTTAGTACAGTAAGAGCAATGTGTTAGTGGGCTTCCTATTAATTTTACTTTGAACCGT
 N H V I L V T Q S P E G * **BfpB M K L G R**

1321 GGTATTCACTTCCCTTATTGTGTCCCTCCTGGCATCTTGCTCGGGTAATGGATTTTATA 1380
 -----+-----+-----+-----+-----+-----+-----+
 CCATAAGTGAAGGGAATAACACAGGGGAGGACCGTAGAACGAGCCCATTACCTAAAATAT
Y S L P L L C P L L A S C S G N G F Y K

*Sph*I

1381 AAGATAATCTTGGCGTAATCGATAAAAATATTCTGCAT**GCTGAC**ACCTCATTGCTGAAGT 1440
 -----+-----+-----+-----+-----+-----+-----+
 TTCTATTAGAACCGCATTAGCTATTTTTATAAGACGTACGACTGTGGAGTAACGACTTCA
 D N L G V I D K N I L H A D T S L L K S

1441 CGAAGAATAAAGAACATTATAAAAGCAGTGATCTGGTAAGCAAAACAGACAGTATTTATA 1500
 -----+-----+-----+-----+-----+-----+-----+
 GCTTCTTATTTCTTGTAATATTTTCGTCACTAGACCATTTCGTTTTGTCTGTCATAAATAT
 K N K E H Y K S S D L V S K T D S I Y I

1501 TAGGAAATAGCAGTTTTTCAGACATATCATGGAGAGCCCCTGCCCGGTAAATTAGAAGGTG 1560
 -----+-----+-----+-----+-----+-----+-----+
 ATCCTTTATCGTCAAAGTCTGTATAGTACCTCTCGGGGACGGGCCATTTAATCTTCCAC
 G N S S F Q T Y H G E P L P G K L E G V

1561 TACATGGAATTATTCTGAGATCAAGTACACCTCTGGGATTTGATGAAGTACTCTCTATGA 1620
 -----+-----+-----+-----+-----+-----+-----+
 ATGTACCTTAATAAGACTCTAGTTCATGTGGAGACCCTAAACTACTTCATGAGAGATACT
 H G I I L R S S T P L G F D E V L S M I

1621 TTCAGGATTCCAGTGGTATTCCCATTGTA AAAACATAACCACGAAAGATGTGATCAGTGGAG 1680
 -----+-----+-----+-----+-----+-----+-----+
 AAGTCCTAAGGTCACCATAAGGGTAACATTTTGTATGGTGCTTTTCTACTAGTCACCTC
 Q D S S G I P I V K H T T K D V I S G G
 ↓ #138/-w3

1681 GGGTGTCCAGCAAAGTCTGGCCGCAACTGTGCGGAAAAAATGAACAGTGCAACAGGCG 1740
 -----+-----+-----+-----+-----+-----+-----+
 CCCACAGGTCGTTTTTACAGACCGGCGTTGACAGCGCCTTTTTTACTTGTCACGTTGTCCGC
 V S S K S L A A T V A E K M N S A T G G
 ↓ -16 ↓ -13

1741 GAAAAAGTACTGATCAATTTGATCATCTACTGCTTGAGGTATCATCAGAACATCAGCTTA 1800
 -----+-----+-----+-----+-----+-----+-----+
 CTTTTTTCATGACTAGTTAAACTAGTAGATGACGAACTCCATAGTAGTCTTGTAGTCGAAT
 K S T D Q F D H L L L E V S S E H Q L M

TGGATGTTAATTATCAGGGCGCGCTTCCACGTTTCTTGATAAAGTTGCGGCAAACCTACA
 1801 -----+-----+-----+-----+-----+-----+-----+ 1860
 ACCTACAATTAATAGTCCCGCGCGAAAGGTGCAAAGAACTATTTCAACGCCGTTTGATGT
 D V N Y Q G A L S T F L D K V A A N Y N
 ACCTGTACTGGACATATGAGTCTGGCAGGATTGCTTTTTCAAACGAGGAACTAAACGCT
 1861 -----+-----+-----+-----+-----+-----+-----+ 1920
 TGGACATGACCTGTATACTCAGACCGTCCTAACGAAAAAGTTTGCTCCTTTGATTTGCGA
 L Y W T Y E S G R I A F S N E E T K R F
 TCAGTATAAGTATATTACCGGGTGGTAAATATACATCAAAAACTCTATCAGTTCGGACA
 1921 -----+-----+-----+-----+-----+-----+-----+ 1980
 AGTCATATTCATATAATGGCCACCATTATATGTAGTTTTTTGAGATAGTCAAGCCTGT
 S I S I L P G G K Y T S K N S I S S D S
 GCAATAGCAGTTCGGGAAGTTCGGGAAGTTCGGGCAGTAGTCCCAGTGATTCTGGAGCAG
 1981 -----+-----+-----+-----+-----+-----+-----+ 2040
 CGTTATCGTCAAGCCCTTCAAGCCCTTCAAGCCCGTCATCAGGGTCACTAAGACCTCGTC
 N S S S G S S G S S G S S P S D S G A E
 AGTTGAAATTCGATTCAGATGTGGATTTCTGGAAAGATATTGAGAACTCGATAAACTGA
 2041 -----+-----+-----+-----+-----+-----+-----+ 2100
 TCAACTTTAAGCTAAGTCTACACCTAAAGACCTTTCTATAACTCTTGAGCTATTTTGACT
 L K F D S D V D F W K D I E N S I K L I
 TACTGGGCAGCGATGGTTCGTATTCAATATCGACAAGTACATCGTCTGTTATTGTCAGAA
 2101 -----+-----+-----+-----+-----+-----+-----+ 2160
 ATGACCCGTCGCTACCAAGCATAAGTTATAGCTGTTTCATGTAGCAGACAATAACAGTCTT
 L G S D G S Y S I S T S T S S V I V R T
 CATCTTCTCCAAACATGAAGAAGATAAACGAATATATTAATACACTGAATGCACAGCTTG
 2161 -----+-----+-----+-----+-----+-----+-----+ 2220
 GTAGAAGAGGTTTGTACTTCTTCTATTTGCTTATATAATTATGTGACTTACGTGTGCAAC
 S S P N M K K I N E Y I N T L N A Q L E
 AAAGACAGGTTACCATTGATGTGCGGATCTATAATGTTACAACAACACTGACAGCAGTGATC
 2221 -----+-----+-----+-----+-----+-----+-----+ 2280
 TTTCTGTCCAATGGTAACTACAGCGCTAGATATTACAATGTTGTTGACTGTGCTCACTAG
 R Q V T I D V A I Y N V T T T D S S D L
 TGGCAATGTCTCTGG**AAGCTT**TGCTAAAGCATAACGGAGGTGTTCTGGGCTCTGTTTCAA
 2281 -----+-----+-----+-----+-----+-----+-----+ 2340
 ACCGTTACAGAGACCTTCGAAACGATTTTCGTATTGCCTCCACAAGATTTGAGACAAAGTT
 A M S L E A L L K H N G G V L G S V S T
 CCTCAAATTTTTCGGCAACAAGCGGCACCCCTTCATTTACAC**GATATC**TGAATGGGAACG
 2341 -----+-----+-----+-----+-----+-----+-----+ 2400
 GGAGTTTAAACGCCGTTGTTGCGCCGTGGGGAAGTAAATGTGCTATAGACTTACCCTTGC
 S N F A A T S G T P S F T R Y L N G N G
 -w2 ↓↓ -87
 GCGACGCAAGCAACCAGGTTTTGCTTAACTTGTTAGCCGAAAAAGGAAAAGTTAGTGTGG
 2401 -----+-----+-----+-----+-----+-----+-----+ 2460
 CGCTGCGTTTCGTTGGTCCAAAACGAATTGAACAATCGGCTTTTTCTTTTCAATCACACC
 D A S N Q V L L N L L A E K G K V S V V
 ↓ -85
 TGACATCTGCATCTGTAACACTACGATGTCAGGTGTCAGCCGGTTCCTTTGAAAGTAGGAAATG
 2461 -----+-----+-----+-----+-----+-----+-----+ 2520
 ACTGTAGACGTAGACATTGATGCTACAGTCCAGTCCGGCCAAGGAACTTTTCATCCTTTAC
 T S A S V T T M S G Q P V P L K V G N D

↓ -w25

2521 ACAGAACATATGTATCTGAGATAGGGACTGTATTAAGCCAGTCATCAACCAGTACAACAG 2580
 -----+-----+-----+-----+-----+-----+-----+
 TGTCTTGTATACATAGACTCTATCCCTGACATAATTCGGTCAGTAGTTGGTCATGTTGTC
 R T Y V S E I G T V L S Q S S T S T T A

2581 CCAGTACGTCAACTGTTACCAGCGGTTTTCTGATGAACCTGTTACCGCAGGTTGCTGATG 2640
 -----+-----+-----+-----+-----+-----+-----+
 GGTCATGCAGTTGACAATGGTCGCCAAAAGACTACTTGGACAATGGCGTCCAACGACTAC
 S T S T V T S G F L M N L L P Q V A D D

↓ -7

2641 ACGGGAATATATTGTTGCAGTATGGTGTAACGCTTTCAGAACTTGTGGCTCAAATAATG 2700
 -----+-----+-----+-----+-----+-----+-----+
 TGCCCTTATATAACAACGTCATACCACATTGCGAAAGTCTTGAACAACCGAGTTTATTAC
 G N I L L Q Y G V T L S E L V G S N N G

2701 GTTTCGATCAGGCAACTGTTAATGGAACCGTAATTCAATTACCTAATGTTGATTCCACAA 2760
 -----+-----+-----+-----+-----+-----+-----+
 CAAAGCTAGTCCGTTGACAATTACCTTGGCATTAAAGTTAATGGATTACAACCTAAGGTGTT
 F D Q A T V N G T V I Q L P N V D S T T

SphI

2761 CCTTTGTGCAATCAAGCATGCTTCGCAATGGTAATACTCTGGTTCTGGCAGGATATGAGA 2820
 -----+-----+-----+-----+-----+-----+-----+
 GGAAACACGTTAGTTCGTACGAAGCGTTACCATTATGAGACCAAGACCGTCCTATACTCT
 F V Q S S M L R N G N T L V L A G Y E K

2821 AGAAGAGAAAACGAAAGTGTGGATCAGGGTGTGGGACAACCAGTTTTAAACTCTTGGGTG 2880
 -----+-----+-----+-----+-----+-----+-----+
 TCTTCTCTTTGCTTTCACACCTAGTCCCACAACCCTGTTGGTCAAATTTGAGAACCCAC
 K R N E S V D Q G V G T T S F K L L G G

2881 GCGCACTGAATGGAAGTGCTTCCCGAACAGTCACTGTCATTTGTATAACCCCGAGAATTA 2940
 -----+-----+-----+-----+-----+-----+-----+
 CGCGTGACTTACCTTACGAAGGGCTTGTGAGTGACAGTAAACATATTGGGGCTCTTAAT
 A L N G S A S R T V T V I C I T P R I I

SD

2941 TTGATCTCAAGGCTTCTGGCGAATGATAAAGAATAATCTTGGCGTAGCGGTAATTGGTTC 3000
 -----+-----+-----+-----+-----+-----+-----+
 AACTAGAGTTCGAAGACCGCTTACTATTTCTTATTAGAACCGCATCGCCATTAACCAAG
 D L K A S G E * **BfpC** M I K N N L G V A V I G S

3001 GAAACAATATGCCGTGAATTTATTATGGGGAAGTTCTCAGGATACAGAAACAACAAATCA 3060
 -----+-----+-----+-----+-----+-----+-----+
 CTTTGTATACGGCACTTAAATAATACCCCTTCAAGAGTCCTATGTCTTTGTTGTTAGT
 K Q Y A V N L L W G S S Q D T E T T N Q

3061 GGCTCTGAATAAGAGCCTGACGTTGATGTCCTCAAAGCTATATTCCGTTATTGGACGATT 3120
 -----+-----+-----+-----+-----+-----+-----+
 CCGAGACTTATTCTCGGACTGCAACTACAGGAGTTTCGATATAAGGCAATAACCTGCTAA
 A L N K S L T L M S S K L Y S V I G R F

3121 TCAGGGTGAACAGTTTGCCGTAGGAGATAAAAATATTGGCCACAAAAGAGGCCAGGTGAC 3180
 -----+-----+-----+-----+-----+-----+-----+
 AGTCCCACTTGTCAAACGGCATCCTCTATTTTTATAACCGGTGTTTTCTCCGGTCCACTG
 Q G E Q F A V G D K N I G H K R G Q V T

↓ -w18

3181 CTTACTGTCTGCAATTGATTTTCGATGGCAGTTCTTTTTGTGGGCTGTTTCCTGCTGATAA 3240
 -----+-----+-----+-----+-----+-----+-----+
 GAATGACAGACGTTAACTAAAGCTACCGTCAAGAAAAACACCCGACAAAGGACGACTATT
 L L S A I D F D G S S F C G L F P A D N

3241 TGAATTATGGCTTGTTCATCGGTGTTGATAAAGATGGAATGGTTCACTTTGACAAGTCCTT 3300
 -----+-----+-----+-----+-----+-----+-----+
 ACTTAATACCGAACAGTAGCCACAACACTATTTCTACCTTACCAAGTGAACTGTTTCAGGAA
 E L W L V I G V D K D G M V H F D K S F

3301 TCACAGTAAGGATGATGCAAAAAAATTCTTTTTTCGACCATGTGGCATATGGATACCCGTG 3360
 -----+-----+-----+-----+-----+-----+-----+
 AGTGTCATTCTACTACGTTTTTTTAAGAAAAAGCTGGTACACCGTATACCTATGGGCAC
 H S K D D A K K F F F D H V A Y G Y P W

↓ -w17

3361 GGACAGAACATACAGTCCTTCAGATGTTGGAGTCGGTGAAAGCAGGAGTATTTCTGAGCT 3420
 -----+-----+-----+-----+-----+-----+-----+
 CCTGTCTTGTATGTCAGGAAGTCTACAACCTCAGCCACTTTCGTCTCATAAAGACTCGA
 D R T Y S P S D V G V G E S R S I S E L

3421 TTCACTAATCAAAGGGAAGAAATTAAGGAAAAAGGTGCCGGTCAACTAATGCCATTGCT 3480
 -----+-----+-----+-----+-----+-----+-----+
 AAGTGATTAGTTTCCCTTCTTTAATTCCTTTTTCCACGGCCAGTTGATTACGGTAACGA
 S L I K G K K L K E K G A G Q L M P L L

3481 GTTATCTGGCGTTGGAATTATAATATTCTTTGTTCATCATATATAATTTTATGAACTTATG 3540
 -----+-----+-----+-----+-----+-----+-----+
 CAATAGACCGCAACCTTAATATTATAAGAAACAGTAGTATATATTTAAATACTTGAATAC
 L S G V G I I I F F V I I Y N F M N L W

↓ -73

3541 GTTAAGTGGCAGAGATAAGGATATTCTTAACGAAGAACCATCTCCGGCCGTCAGTGAGGA 3600
 -----+-----+-----+-----+-----+-----+-----+
 CAATTCACCGTCTCTATTCTATAAGAATTGCTTCTTGGTAGAGGCCGGCAGTCACTCCT
 L S G R D K D I L N E E P S P A V S E E

3601 AGTCAGAAATATCCCGTGGGAAGGAAAAGCAAACCGGGCTCATTGTTTTCTAAATGTAT 3660
 -----+-----+-----+-----+-----+-----+-----+
 TCAGTCTTTATAGGGCACCTTCCCTTTTTCGTTTTGGCCCGAGTAACAAAAGATTTACATA
 V R N I P W E G K S K P G S L F S K C I

3661 AAAAGAAATGGATTTCATTCCGGTTCCAAGCAGCATCCGTTCCAGGCTGGATGCCTGAAAA 3720
 -----+-----+-----+-----+-----+-----+-----+
 TTTTCTTTACCTAAGTAAGGCCAAGGTTTCGTTCGTAGGCAAGGTCCGACCTACGGACTTTT
 K E M D S F R F Q A A S V P G W M P E N

3721 TAAAGTAACCTGTAATGACTCTGATATAAACTTTAGTGTTTATCGTTCTGGCGGACTTAA 3780
 -----+-----+-----+-----+-----+-----+-----+
 ATTTCAATGGACATTACTGAGACTATATTTGAAATCACAATAGCAAGACCGCCTGAATT
 K V T C N D S D I N F S V Y R S G G L N

↓ -89

3781 TATATGGTATGAAAATCCGGGCCTGTTTTTTTACAGTGGAGACATCCCAGACATTACCAG 3840
 -----+-----+-----+-----+-----+-----+-----+
 ATATACCATACTTTTAGGCCCGGACAAAAAAGTGTACCTCTGTAGGGTCTGTAATGGTC
 I W Y E N P G L F F H S G D I P D I T R

3841 AAATGATAATGATAAAAATAACATTTACATGGCCTCTTGAAACGGTCAGGTATGACAGTTC 3900
 -----+-----+-----+-----+-----+-----+-----+-----+
 TTTACTATTACTATTTTTATTGTAAATGTACCGGAGAACCTTGCCAGTCCATACTGTCAAG
 N D N D K I T F T W P L E T V R Y D S S
 3901 GATACTCGACGTGCATACCTTGAATAAGACTCAAGAAATAATAAGGTATCTTTCACGTCA 3960
 -----+-----+-----+-----+-----+-----+-----+-----+
 CTATGAGCTGCACGTATGGAACCTTATTCTGAGTTCTTTATTATTCCATAGAAAGTGCAGT
 I L D V H T L N K T Q E I I R Y L S R Q
 3961 GTTTGAGAATAGCTGGATAGAGATAAAAATTTGAACCGTCCGTGCCGGACGGCAATACAGA 4020
 -----+-----+-----+-----+-----+-----+-----+-----+
 CAAACTCTTATCGACCTATCTCTATTTTAAACTTGGCAGGCACGGCCTGCCGTTATGTCT
 F E N S W I E I K F E P S V P D G N T E
 ACGAGTGGGTTTTTCTTTTTCATTTAACCATGATCCGAGAATCTTACTCCAATTCTTTC
 4021 -----+-----+-----+-----+-----+-----+-----+-----+ 4080
 TGCTCACCCAAAAGAAAAGTAAATTGGTACTAGGCTCTTAGAATGAGGGTTAAGAAAG
 R V G F S F S F N H D P R I L L P I L S
 TGAAGTAAATGGACTGATTATCACTAATATCGATTATGATTTTTTCAGGAAGTAACTGGAA
 4081 -----+-----+-----+-----+-----+-----+-----+-----+ 4140
 ACTTCATTTACCTGACTAATAGTGATTATAGCTAATACTAAAAGTCCTTCATTGACCTT
 E V N G L I I T N I D Y D F S G S N W N
 CGTGAAAGGTGTTTTTTGGGGGAGAGCTTAGTTTGAAAAAATATTGTTTACTTTTTTGT
 4141 -----+-----+-----+-----+-----+-----+-----+-----+ 4200
 GCACCTTCCACAAAACCCCTCTCGAATCAAACCTTTTTTTATAACAAATGAAAAACA
 V K D V F W G R A *
 BfpU M K K I L F T F L F
 ↓ -90 ↓ -84
 4201 TTGTCGTTTTCTTCAGTCGTGAATGCTGAGGTCAATACAAGTGTTGATGTTCCGGTAAGA 4260
 -----+-----+-----+-----+-----+-----+-----+-----+
 AACAGCAAAGAAGTCAGCACTTACGACTCCAGTTATGTTCACTAACAAGGCCATTCT
 V V S S V V N A E V N T S V D V S G K N
 XbaI
 4261 ATGACTCTAGAAAAGAAGAAAATAAATTCCTGATGTTGATTTAGAAAAGTATAAAGATG 4320
 -----+-----+-----+-----+-----+-----+-----+-----+
 TACTGAGATCTTTTCTTCTTTTATTTAACGGACTACAATAAATCTTTTCATATTTCTAC
 D S R K E E N K L P D V D L E K Y K D E
 AATTGCTTTCCTATCAGGAACCTCAGCGTAAGAATATGATGCTTAAATTACAGTCTGAAA
 4321 -----+-----+-----+-----+-----+-----+-----+-----+ 4380
 TTAACGAAAGGATAGTCCTTGAAGTCGCATTCTTATACTACGAATTTAATGTCAGACTTT
 L L S Y Q E L Q R K N M M L K L Q S E N
 ATCAGAAACTGGAGAGTGAAATAAAAAAAGTGGATCATTTATCTCATATGATACAAATG
 4381 -----+-----+-----+-----+-----+-----+-----+-----+ 4440
 TAGTCTTTGACCTCTCACTTTATTTTTTTTTCACCTAGTAAATAGAGTATACTATGTTTAC
 Q K L E S E I K K S G S F I S Y D T N E
 4441 AAAACACGATATACTTAATGTCTATAATGACAACCTCCTAAAAAAGGATTATGTGCAACCA 4500
 -----+-----+-----+-----+-----+-----+-----+-----+
 TTTTGTGCTATATGAATTACAGATATTACTGTTGAGGATTTTTTTCCTAATACACGTTGGT
 N T I Y L M S I M T T P K K G L C A T I

↓ -w31 **EcoRV**

4501 TATTTAATGGGGCTATAAGCAGAGTCTGTAAGGGCGATATCATTAGCGAACATTCTGTGG 4560
 -----+-----+-----+-----+-----+-----+-----+-----+
 ATAAATTACCCCGATATTCGTCTCAGACATTCCCGCTATAGTAATCGCTTGTAAACACC
 F N G A I S R V C K G D I I S E H S V V

4561 TGAGTGACATTACCGGTAATTATGTGGAGATCAGGAGTATTAATAACATATCCGATCCCG 4620
 -----+-----+-----+-----+-----+-----+-----+
 ACTCACTGTAATGGCCATTAATACACCTCTAGTCCTCATAATTATTGTATAGGCTAGGGC
 S D I T G N Y V E I R S I N N I S D P V

4621 TGCAGAAAATTTTTCTGAGATAGGTAAGTTGATGCTAACCAAAACAGAAAAAACTTCAG 4680
 -----+-----+-----+-----+-----+-----+-----+
 ACGTCTTTTAAAAAGACTCTATCCATTCAAACACTACGATTGGTTTTGTCTTTTTTTGAAGTC
 Q K I F L R * **BfpD'** M L T K T E K T S D

BglII

4681 ATCTGATGTTTGAACGATTTAAGCGCAATGTATCTGAAATCGTCACGGGAGATGGGGGGG 4740
 -----+-----+-----+-----+-----+-----+-----+
 TAGACTACAACTTGCTAAATTTCGCGTTACATAGACTTTAGCAGTGCCCTCTACCCCCC
 L M F E R F K R N V S E I V T G D G G E

↓ -2

4741 AGCTTGAGCTTACTGTTGAACAGAGAAAATACTTCCTGATATTCAAAAATGGTGATTTTC 4800
 -----+-----+-----+-----+-----+-----+-----+
 TCGAACTCGAATGACAACCTGTCTCTTTTATGAAGGACTATAAGTTTTTACCACTAAAAG
 L E L T V E Q R K Y F L I F K N G D F L

4801 TTGTTTCTTCATGTCATATGAAGCATCATCTGGTACAGATGTTACGTGAAATCGCAACAA 4860
 -----+-----+-----+-----+-----+-----+-----+
 AACAAAGAAGTACAGTATACTTCGTAGTAGACCATGTCTACAATGCACTTTAGCGTTGTT
 V S S C H M K H H L V Q M L R E I A T R

↓ -w4 **EcoRV**

4861 GAAAAGGGTATCCAAACCTCACTATTTATGAGGTGAACTTAAAAGATATCAGGTTATTGT 4920
 -----+-----+-----+-----+-----+-----+-----+
 CTTTTCCCATAGGTTTGGAGTGATAAATACTCCACTTGAATTTTCTATAGTCCAATAACA
 K G Y P N L T I Y E V N L K D I R L L Y

4921 ATGAAGCATCTCTCAAACCGTACAGAATAATGGACAGGATCTATTACCAGTCGAAAAAA 4980
 -----+-----+-----+-----+-----+-----+-----+
 TACTTCGTAGAGAGTTTTGGCATGTCTTATTACCTGTCCTAGATAATGGTCAGCTTTTTT
 E A S L K T V Q N N G Q D L L P V E K R

4981 GAGCTTCAATGCTACTTTTCGAATGCGCCGAGATGAGAGTAAGTGATTTACACATCAAAG 5040
 -----+-----+-----+-----+-----+-----+-----+
 CTCGAAGTTACGATGAAAAGCTTACGCGGCTCTACTCTCATTCACTAAATGTGTAGTTTC
 A S M L L F E C A E M R V S D L H I K V

5041 TGTATGATGCTGAGGCTGATATTTATATACGTAAAGATGGGGATATGGAACCTCTGCGTC 5100
 -----+-----+-----+-----+-----+-----+-----+
 ACATACTACGACTCCGACTATAAATATATGCATTTCTACCCCTATACCTTGAGAACGCAG
 Y D A E A D I Y I R K D G D M E L L R Q

5101 AGATTGAGTCTAATACGGCGCATTCTATCCTTGCATCATTATATAATAATGCAGATGATT 5160
 -----+-----+-----+-----+-----+-----+-----+
 TCTAACTCAGATTATGCCGCGTAAGATAGGAACGTAGTAATATATTATTACGTCTACTAA
 I E S N T A H S I L A S L Y N N A D D S

5161 CTGATGCAACCTATAAAATTAATGCTTATCAGGCTGCTCGTATAGTTGCTTCAAAATCGA 5220
 -----+-----+-----+-----+-----+-----+-----+
 GACTACGTTGGATATTTTAATTACGAATAGTCCGACGAGCATATCAACGAAGTTTTAGCT

 D A T Y K I N A Y Q A A R I V A S K S R

 GGCTGGCGTTACCTCCCGTAATTCAGGCGGTGCGCCTTCAGTTTAACCCCTCTTGGTCAAG
 5221 -----+-----+-----+-----+-----+-----+ 5280
 CCGACCGCAATGGAGGGCATTAAAGTCCGCCACGCGGAAGTCAAATTGGGAGAACCAGTTC

 L A L P P V I Q A V R L Q F N P L G Q G

 GTGGGCGATATTTAATCGCAGATTCCCT**GTATAC**AGACAAAAGTGAAAAACAAAAGAAA
 5281 -----+-----+-----+-----+-----+ 5340
 CACCCGCTATAAATTAGCGTGCTAAGGACATATGTCTGTTTTCACTTTTTGTTTTCTTT

 G R Y L I A R F L Y T D K S E K Q K E M

BamHI
TGGATCC
 5341 ----- 5347
 ACCTAGG

D

Appendix 2. The annotated sequence of pRDH8.

The 6334 bp DNA sequence of the *Bam*HI fragment from pRDH8 is presented in the same orientation as the open reading frames are depicted in **Figure 4.7**. The predicted Shine-Delgarno sequences, start codons and termination codons of each open reading frame are indicated in bold and underlined. The deduced amino acid sequence of each open reading frame is shown below the nucleotide sequence; where appropriate signal sequences are indicated by underlining. The exact location of the transposon insertion sites in the CVD206::*TnphoA* mutants #13, #137 and #181, and the pRDH8::*TnphoA* plasmids used in the DNA sequencing (see also **Figure 4.2**) are indicated by vertical arrows (↓ indicates the nucleotide before the insertion).

```

BamHI
GGATCCTAAGATGAAAAGGCTAATTATTGTTGGTATCCAACGGTCGGTTACTCCTGAT
1  -----+-----+-----+-----+-----+-----+-----+ 60
  CCTAGGATTCTACTTTTCCGATTAATAACAACACCATAGGTTGCCAGCCAATGAGGACTA

  TACGCTGTCTCTTATATAAACCTACCATATTAGCGGCAGAGACAGCGTTTTACTTAGCC
61  -----+-----+-----+-----+-----+-----+-----+ 120
  ATGCGACAGAGAATATATTTGGGATGGTATAATCGCCGTCTCTGTGCGAAAATGAATCGG
  ↓ -w2
121  -----+-----+-----+-----+-----+-----+-----+ 180
  CTGAAGCGTGATTTTTTTTACATTTTAAGAACAGGATGACCGCTAATCGCCAACCTGAGTGC
  GACTTCGCACTAAAAAAAATGTAATAATTCTTGTCTACTGGCGATTAGCGGTTGACTCACG
  * M K L V P H G S I A L Q T

  CTTCCTGGGTGTTATTCAGAATGGCATGTGCGCCCAACGGCAGCGTAACCGTGCGTTGTT
181  -----+-----+-----+-----+-----+-----+-----+ 240
  GAAGGACCCACAATAAGTCTTACCGTACACGCGGGTTGCCGTGCGATTGGCAGCAACAA
  G E Q T N N L I A H A G L P L T V T R Q

  CATGACCAAAATCGAGGCCGGTAATGAGTGGAAATCGACAGGCGGGAACGCAAAAACGCGT
241  -----+-----+-----+-----+-----+-----+-----+ 300
  GTACTGGTTTTAGCTCCGGCCATTACTCACCTTAGCTGTCCGCCCTTGCGTTTTTGGCGA
  E H G F D L G T I L P I S L R S R L F A
  ↓ -w6
301  -----+-----+-----+-----+-----+-----+-----+ 360
  AACTGACTCCAGGTTGTAACCTGCGTCATAATCATTAGGCGTGCTGCCGCTAAAGCTAC
  TGTGACTGAGGTCCAACATTGGACGCGAGTATTAGTAATCCGCACGACGGCGATTTCGATG
  Y V S E L N Y G A D Y D N P T S G S F S
  
```


1081 ATGTATATCATGCCGCTTAGGTGTGCCGTTGTACACCTCAACGCGGATTCCAGGCTATTAG 1140
-----+-----+-----+-----+-----+-----+-----+
TACATATAGTACGGCGAATCCACACGGCAACAGTGGAGTTGCGCCTAAGGTCCGATAATC

SD

1141 GATAGAAGAAGTGAAATTGAGATGGTTTGCCTTTTGGATTGTGTTATTAGCGGGTTGTTC 1200
-----+-----+-----+-----+-----+-----+-----+
CTATCTTCTTCACTTTAACTCTACCAAACGGAAAACTAACACAATAATCGCCCAACAAG

E m t A M K L R W F A F L I V L L A G C S

1201 ATCAAAGCATGACTATACGAACCCGCCGTGGAACGCGAAAGTTCCGGTGCAACGTGCGAT 1260
-----+-----+-----+-----+-----+-----+-----+
TAGTTTCGTACTGATATGCTTGGGCGGCACCTTGCCTTTCAAGGCCACGTTGCACGCTA

S K H D Y T N P P W N A K V P V Q R A M

1261 GCAGTGGATGCCAATAAGCCAGAAAGCCGGTGCAGCCTGGGGCGTCGATCCACAATTGAT 1320
-----+-----+-----+-----+-----+-----+-----+
CGTCACCTACGGTTATTCGGTCTTTTCGGCCACGTCGGACCCCGCAGCTAGGTGTTAACTA

Q W M P I S Q K A G A A W G V D P Q L I

1321 CACGGCGATTATCGCTATCGAATCGGGTGGTAATCCGAACGCGGTGAGTAAATCGAATGC 1380
-----+-----+-----+-----+-----+-----+-----+
GTGCCGCTAATAGCGATAGCTTAGCCACCATTAGGCTTGCGCCACTCATTAGCTTACG

T A I I A I E S G G N P N A V S K S N A

HindIII ↓ #13/#137/-30

1381 CATTGGTTTGTATGCAGATAAAAGCTTCAACCTCCGGACGTGATGTATATCGTCGTATGGG 1440
-----+-----+-----+-----+-----+-----+-----+
GTAACCAAACACTACGTCTATTTTCGAAGTTGGAGGCCTGCACTACATATAGCAGCATACCC

I G L M Q I K A S T S G R D V Y R R M G

1441 CTGGAGTGGTGAGCCGACGACCAGCGAGCTGAAAAATCCGGAGCGTAATATTTCAATGGG 1500
-----+-----+-----+-----+-----+-----+-----+
GACCTCACCCTCGGCTGCTGGTCGCTCGACTTTTATAGGCCTCGCATTATAAAGTTACCC

W S G E P T T S E L K N P E R N I S M G

-w32/-w26 ↓

1501 AGCGGCTTACCTGAATATTCTGGAAACCGGCCCGCTGGCAGGCATTGAAGATCCGAAGGT 1560
-----+-----+-----+-----+-----+-----+-----+
TCGCCGAATGGACTTATAAGACCTTTGGCCGGGCGACCGTCCGTAACCTTCTAGGCTTCCA

A A Y L N I L E T G P L A G I E D P K V

1561 ACTGCAATATGCGCTGGTGGTGTGCATACGCTAACGGGGCAGGTGCGCTGCTACGGACTTT 1620
-----+-----+-----+-----+-----+-----+-----+
TGACGTTATACGCGACCACCACAGTATGCGATTGCCCGTCCACGCGACGATGCCTGAAA

L Q Y A L V V S Y A N G A G A L L R T F

↓ -15

1621 CTCGTCAGATCGGAAAAAGGCGATCACCAAAATCAACGATCTGGATTCTGATGAGTTCCT 1680
-----+-----+-----+-----+-----+-----+-----+
GAGCAGTCTAGCCTTTTTCGCTAGTGGTTTTAGTTGCTAGACCTAAGACTACTCAAGGA

S S D R K K A I T K I N D L D S D E F L

1681 CGAACACGTAGCGCGAAATCATCCTGCGCCGCGAGGCTCCGCGCTATATCTACAACTTGA 1740
-----+-----+-----+-----+-----+-----+-----+
GCTTGTGCATCGCGCTTTAGTAGGACGCGGCGTCCGAGGCGCGATATAGATGTTTGAAC

E H V A R N H P A P Q A P R Y I Y K L E

**rho independent
terminator**

1741 GCAGGCACTGGACGCGATGTAAATCAGTCGCGCACATTGTCGGCTTTTTTCCGGGCTTCT 1800
-----+-----+-----+-----+-----+-----+-----+
CGTCCGTGACCTGCGCTACATTTAGTCAGCGCGTGTAACAGGCGAAAAAAGGCCCGAAGA
Q A L D A M * * D R V N D A K K R A E
1801 CGCTCGAGAGAGAAAATAATCCGCTGTAATTGCCGCTCCACCGTCGGGCTGACGTTAAGA 1860
-----+-----+-----+-----+-----+-----+-----+
GCGAGCTCTCTCTTTTATTAGGCGACATTAACGGCGAGGTGGCAGCCCGACTGCAATTCT
R E L S F I I R Q L Q R E V T P S V N L
1861 AAACGGAAGCTCAGACGGGGAGTGGTGATGGTTTTTCAATTTTTGCCATCAATCACTTTGCGC 1920
-----+-----+-----+-----+-----+-----+-----+
TTTGCCTTCGAGTCTGCCCTCACCACTACCAAAGTAAAAACGGTAGTTAGTGAAACGCG
F R F S L R P T T I T E N K G D I V K R
1921 TCGCTGATGGATATTAAGTGGGCGTCAAAGTGAAAAACACCCCATGCCCCATGTTGACT 1980
-----+-----+-----+-----+-----+-----+-----+
AGGCACTACCTATAATTGACCCGCGAGTTTCACTTTTTGTGGGGTAACGGGGTACAAGTGA
E S I S I L Q A D F H F V G W Q G M N V
SphI
1981 TCAATCTGAGCGAAGTGCATGCCCTTCGTGTAATCCGGCAGGCTTTGCTGTTTCCAGTAAT 2040
-----+-----+-----+-----+-----+-----+-----+
AGTTAGACTCGCTTCACGTACGGAAGCACATTAGGCCGTCCGAAACGACAAAGGTCATTA
E I Q A F H M G E H L G A P K A T E L L
2041 GCGCCCATGCCGCTAACGACAAATCATAACAGGCGGAAACGTAACGTACTGTTATCCGCC 2100
-----+-----+-----+-----+-----+-----+-----+
CGCGGGTACGGCGGATTGCTGTTTAGTATGTCCGCCTTTGCATTGCATGACAATAGGCGG
A G M G G L S L D Y L R F R L T S N D A
2101 AGTTTGGTCTGGCAAAAATAAGGCGGATGGAGTGGGGCGGTGATGCGGAAATATCGGCGT 2160
-----+-----+-----+-----+-----+-----+-----+
TCAAACCAGACCGTTTTTATTCCGCCTACCTACCCCGCCACTACGCCTTTATAGCCGCA
L K T Q C F Y P P H L P A T I R F Y R R
2161 CGTTGTACAAACCATAAGGTGGGAGGCGGTACGGTAATAAATGCCGGAAGCTGCAAGTAT 2220
-----+-----+-----+-----+-----+-----+-----+
GCAACATGTTTGGTATTCCACCCTCCGCCATGCCATTATTTACGGCCTTCGACGTTTATA
R Q V F W L T P P P V T I F A P L Q L Y
↓ -w28
2221 TCACTCTGCTGTAGTTGTTCAACAGTAAACTCGACTTTTCGCACCCTGAGTTTTCGGCGGTA 2280
-----+-----+-----+-----+-----+-----+-----+
AGTGAGACGACATCAACAAGTTGTCATTTGAGCTGAAAGCGTGGGACTCAAAGCCGCCAT
E S Q Q L Q E V T F E V K A G Q T E A T
2281 ATGGTAATGTGCTGCGCCTTTAGCACGGCGTTGTTATCTTCGGCCTGGCTGTCGAAATCC 2340
-----+-----+-----+-----+-----+-----+-----+
TACCATTACACGACGCGGAAATCGTGCCGCAACAATAGAAGCCGGACCGACAGCTTTAGG
I T I H Q A K L V A N N D E A Q S D F D
2341 AGCACCAGTTTATCCGGGGTTATTACCAGTATTTGGTGATCAGCTGTCCGCCATTCCAG 2400
-----+-----+-----+-----+-----+-----+-----+
TCGTGGTCAAATAGGCCCAATAATGGTCATAAAACCACTAGTCGACAGGCGGTAAGGTC
L V L K D P T I V L I K T I L Q G G N W

2401 CTGATACGCAAAGGAATTGCGGCTTTGTTCAAATCGCGTAACACGCCAGGACGGCTAAT 2460
 -----+-----+-----+-----+-----+-----+-----+-----+
 GACTATGCGTTTTCTTAACGCCGAAACAAGTTAGCGCATTGTGCGGGTCCTGCCGATTA
 S I R L P I A A K N L D R L V G L V A L
 2461 GGATTTTGTTCAGGAACTGCTCATGGTAATGACTCACGCGGAAAACTCCTGACTGACA 2520
 -----+-----+-----+-----+-----+-----+-----+-----+
 CCTAAAACAAGTCCTTGACGAGTACCATTACTGAGTGCCTTTTTTGAGGACTGACTGT
 P N Q K L F Q E H Y H S M Orf3
 ↓ -w27
 2521 AACTGTCTTAGGTTTATCGGTCACAGTTAACAAAACCTTAATACAAATGCGTGAATATTTT 2580
 -----+-----+-----+-----+-----+-----+-----+-----+
 TTGACAGAATCCAAATAGCCAGTGTCAATTGTTTTGAATTATGTTTACGCACTTATAAAA
 2581 TTACATGTCGTTCTTAAATCAGCCGACATACGCCAGCATCGATCCCGCCCTTGCCTATA 2640
 -----+-----+-----+-----+-----+-----+-----+-----+
 AATGTACAGCAAGAATTTAGTCGGCTGTATGCGGGTCGTAGCTAGGGCGGGAACGGATAT
 CTTAGAGCGTTGACGTAAGCATTCTTTCGCTCCGATTCATCGAAACGAGGGCCTGAACAT
 2641 -----+-----+-----+-----+-----+-----+-----+-----+ 2700
 GAATCTCGCAACTGCATTTCGTAAAGAACGCAGGCTAAGTAGCTTTGCTCCCGGACTTGTA
 Tag M
 ↓ -w11
 2701 GGGAATTATTGCCTGGATTATTTTTGGCCTGATAGCCGGCATTATCGCCAAGCTAATCAT 2760
 -----+-----+-----+-----+-----+-----+-----+-----+
 CCCTTAATAACGGACCTAATAAAAACCGGACTATCGGCCGTAATAGCGGTTTCGATTAGTA
 G I I A W I I F G L I A G I I A K L I M
 2761 GCCGGGGCGTGATGGTGGTGGATTTTTCTGACCTGTATTCTTGGGATAGTCGGTGCGGT 2820
 -----+-----+-----+-----+-----+-----+-----+-----+
 CGGCCCGCACTACCACCACCTAAAAAGGACTGGACATAAGAACCCTATCAGCCACGCCA
 P G R D G G G F F L T C I L G I V G A V
 ↓ #181
 2821 GGTCGGCGGCTGGCTGGCGACCATGTTTGGCATTGGCGGCTCCATCAGCGGTTTTAATTT 2880
 -----+-----+-----+-----+-----+-----+-----+-----+
 CCAGCCGCGGACCGGCTGGTACAAACCGTAACCGCCGAGGTAGTCGCCAAAATTA
 V G G W L A T M F G I G G S I S G F N L
 2881 GGACAGCTTCCTGGTGGCGGTGGTGGGCGCTATCCTGGTGCTGGGCGTATTCCGCCTCCT 2940
 -----+-----+-----+-----+-----+-----+-----+-----+
 CCTGTGGAAGGACCACCGCCACCACCGCGATAGGACCACGACCCGCATAAGGCGGAGGA
 D S F L V A V V G A I L V L G V F R L L
 Inverted Repeats
 2941 GCAAAGAGAATAAGATTTTCATAAGGCGGATAGCGGATACAGATGCGCTATCCGCTTTCA 3000
 -----+-----+-----+-----+-----+-----+-----+-----+
 CGTTTTCTCTTATTCTAAAAGTATTCGCGCTATCGCTATGTCTACGGCGATAGGCGAAAGT
 Q R E *
 AccI ↓ K-12 cut off
 3001 CATCAGAACGTATACTCGATACCTGCCAGAAATTGCGACCTTCTTCAATAAACCTTCA 3060
 -----+-----+-----+-----+-----+-----+-----+-----+
 GTAGTCTTGCATATGAGCTATGGACGGGTCTTTAACGCTGGAAGAAGTTATTTGGGAAGT
 * F T Y E I G A W F N R G E E I F G E
 3061 CTGTAGGCATATTTGGTATCAAACAGGTTATTGACCGACGCATTAACGCTGAAGCCATGA 3120
 -----+-----+-----+-----+-----+-----+-----+-----+
 GACATCCGTATAAACCATAGTTTGTCCAATAACTGGCTGCGTAATTGCGACTTCGGTACT
 S Y A Y K T D F L N N V S A N V S F G H

3121 CCTAAGGTGTAATCGGCTCGAATGTGGGTACCCCAAACCGGCGGCTTTTTGTGAACCG 3180
 -----+-----+-----+-----+-----+-----+-----+
 GGATTCCACATTAGCCGAGCTTACACCCAGTGGGGTTTTGGCCGCGGAAAAACACTTGGC
 G L T Y D A R I H T V G F G A A K Q S G
 3181 TCACTGTTGCTGTAGCTGGAGGAACGCGCCTCTTCCGACAGCGTTACGCTTAACGGCTCC 3240
 -----+-----+-----+-----+-----+-----+-----+
 AGTGACAACGACATCGACCTCCTTGCGCGGAGAAGGCTGTGCAATGCGAATTGCCGAGG
 D S N S Y S S S R A E E S L T V S L P E
 3241 CACGGTTTGAGAGTCATCCATGCGGTCATTGTCTGCGTTGGCAGATCGTTTATCTTGCCG 3300
 -----+-----+-----+-----+-----+-----+-----+
 GTGCCAAACTCTCAGTAGGTACGCCAGTAACAGACGCAACCGTCTAGCAAATAGAACGGC
 W P K L T M W A T M T Q T P L D N I K G
 3301 ATGTCTTTACGTTTGGCATCAGCGTGGATCAGGGCGTAGCTCAATCCTACATCCAGTATA 3360
 -----+-----+-----+-----+-----+-----+-----+
 TACAGAAATGCAAACCGTAGTCGCACCTAGTCCCGCATCGAGTTAGGATGTAGGTCATAT
 I D K R K A D A H I L A Y S L G V D L I
 3361 TTGCTGATTTTCCCCTTAATACCGGCATCCAGACCGCTGTAATCCACCGTGCCGCTGTTC 3420
 -----+-----+-----+-----+-----+-----+-----+
 AACGACTAAAAGGGGAATTATGGCCGTAGGTCTGGCGACATTAGGTGGCACGGCGACAAG
 N S I K G K I G A D L G S Y D V T G S N
 3421 TGATTTTGAATGGTATCGGCATCGATATTGTGCGAGAGGATGGCATCACTCACCCGGTTA 3480
 -----+-----+-----+-----+-----+-----+-----+
 ACTAAAACCTTACCATAGCCGTAGCTATAACACGCTCTCCTACCGTAGTGAGTGGGCCAAT
 Q N Q I T D A D I N H S L I A D S V R N
 3481 TAGTAAACGCTGACCTCAAATCCCCAGTCGTGCGTGAAGGCACCATTCCAGGTTAAATCC 3540
 -----+-----+-----+-----+-----+-----+-----+
 ATCATTGCGACTGGAGTTTAGGGGTCAGCACGCACCTTCCGTGGTAAGGTCCAATTTAGG
 Y Y V S V E F G W D H T F A G N W T L D
 3541 ACCCCGCGCGCGGCTTCCGGCTTGAGCTGCGGGTTAACAATCGCTATCTGGTTATACGCA 3600
 -----+-----+-----+-----+-----+-----+-----+
 TGGGGCGCGCGCAAGGCCGAACCTCGACGCCCAATTGTTAGCGATAGACCAATATGCGT
 V G R A R E P K L Q P N V I A I Q N Y A
 3601 GGTTTGGACGTGGTATAGCGTTCTTTCAGCGTCGGAAAGCGTGTGCGGTCATAGTACGAA 3660
 -----+-----+-----+-----+-----+-----+-----+
 CCAAACCTGCACCATATCGCAAGAAAGTCGCAGCCTTTCGCACACGCCAGTATCATGCTT
 P K S T T Y R E K L T P F R T R D Y Y S
 3661 AGCGCCAGCGTGTCTTCATTGCCAAAGTGGTATTTCCCATCACCTGCCAGTTAAAAGCT 3720
 -----+-----+-----+-----+-----+-----+-----+
 TCGCGGTGCGCACAGAAGTAACGGTTTCACCATAAAGGGGTAGTGGACGGTCAATTTTCGA
 L A L T D E N G F H Y K G M V Q W N F A
 3721 GACTGATTGTTGTCGTCATAATGGGTGATACTGCCATCTTTCTCATGTTTTTTTCGCTTCT 3780
 -----+-----+-----+-----+-----+-----+-----+
 CTGACTAACCAACAGCAGTATTACCCACTATGACGGTAGAAAGAGTACAAAAAAGCGAAGA
 S Q N N D D Y H T I S G D K E H K K A E
 3781 ACGCTATCGCGCCAGTCATAGCTGATTCCAGCCACGACATCGACATTATCGGCAGCAGCC 3840
 -----+-----+-----+-----+-----+-----+-----+
 TGCGATAGCGCGGTGAGTATCGACTAAGGTGCGGTGCTGTAGCTGTAATAGCCGTGTCGG
 V S D R W D Y S I G A V V D V N D A A A

↓ -w10

3841 CATTGATATTCACTGGCGAGCGACCAGGTACGATCTTCATAGCGATCGTAAGCGGCGTG 3900
 -----+-----+-----+-----+-----+-----+-----+
 GTAACATAAGTGACCGCTCGCTGGTCCATGCTAGAAGTATCGCTAGCATTCCGCCGACG
 W Q Y E S A L S W T R D E Y R D Y A A H

3901 GGCGCACCTTTTTCCCGGTGTACGTCATCTTTCCAGTTAACGGCAAACGACAGCAGATCG 3960
 -----+-----+-----+-----+-----+-----+-----+
 CCGCGTGGAAAAAGGGCCACATGCAGTAGAAAGGTCAATTGCCGTTTGCTGTCTAGC
 P A G K E R H V D D K W N V A F S L L D

3961 TTTTCACGCACATCGGCTGCCAGTTGTAATCCGGCACCGTCCGCTGTAATCGGAATAATGG 4020
 -----+-----+-----+-----+-----+-----+-----+
 AAAAGTGCCTGTAGCCGACGGTCAACATTAGGCCGTGGCAGCGACATTAGCCTTATTACC
 N E R V D A A L Q L G A G D S Y D S Y H

4021 CTGTAGCTGCCTTTTTTATTTTTCAAATCAGCCAGCGAGTTGTACATCATTAGCGTATTT 4080
 -----+-----+-----+-----+-----+-----+-----+
 GACATCGACGGAAAAATAAAAAGTTTAGTCCGGTCCGCTCAACATGTAGTAATCGCATAAA
 S Y S G K K N K L D A L S N Y M M L T N
 ↓ -w31

4081 TCAAAGGTGTCCGATAACAGCCGACTTTTCAGGGTAAAACGATCGTTTAGTTGGGTCGTT 4140
 -----+-----+-----+-----+-----+-----+-----+
 AGTTTCCACAGCGCTATGTCCGGCTGAAAAGTCCCATTTTGCTAGCAAATCAACCCAGCAA
 E F T D R Y L R S K L T F R D N L Q T T

4141 CCCTGATAATAAAAACTTTCTTTGTCATACTCTGGCCACTGCCAGTAGCGTGATTTCTGA 4200
 -----+-----+-----+-----+-----+-----+-----+
 GGGACTATTATTTTTGAAAGAAACAGTATGAGACCGGTGACGGTCATCGCACTAAAGACT
 G Q Y Y F S E K D Y E P W Q W Y R S K Q

4201 CCACTATTTCTACTGTATGGCGGTTATCTTTTTACCATCCTGCTTAATGTAAGTCAAT 4260
 -----+-----+-----+-----+-----+-----+-----+
 GGTGATAAAGATGACATAACCGCCAATAGAAAAAGTGGTAGGACGAATTACATTCAGTTA
 G S N R S Y P P N D K E G D Q K I Y T L
 ↓ -4

4261 GTGTATTCATCGTTTTTCGCGTGGTGTAAAACCGAGCTTCACAATGCCGCGTTTATCATCA 4320
 -----+-----+-----+-----+-----+-----+-----+
 CACATAAGTAGCAAAAAGCGCACCATTTTTGGCTCGAAGTGTACGGCGCAAATAGTAGT
 T Y E D N E R P T F G L K V I G R K D D

4321 GCCGATGAATTAATCATCTTGCCGTGTTTGGCTGCAATATCATTATTTACGCCATGCGGC 4380
 -----+-----+-----+-----+-----+-----+-----+
 CGGCTACTTAATTAGTAGAACGGCACAAACCGACGTTATAGTAATAAATGCCGGTACGCCG
 A S S N I M K G H K A A I D N N V G H P

4381 AGCCCCGAGAAAATCCTGTTTTAGCTGGCTACCGCTGACTTGCAAATACCCCAGATCGCTG 4440
 -----+-----+-----+-----+-----+-----+-----+
 TCGGGCTCTTTTAGGACAAAATCGACCGATGGCGACTGAACGTTTATGGGGTCTAGCGAC
 L G L F D Q K L Q S G S V Q L Y G L D S

SphI

4441 CTGACGGCAAATGAAGCATGCATATCGTAGGCATTGTCCTGGCTACGGCTCCATCCCTGG 4500
 -----+-----+-----+-----+-----+-----+-----+
 GACTGCCGTTTACTTCGTACGTATAGCATCCGTAACAGGACCGATGCCGAGGTAGGGACC
 S V A F S A H M D Y A N D Q S R S W G Q

EcoRV

4501 **CGATATC**CCAGACTGGCTTCCAGAGGTTTTGTTGGCTTCTGGGTGGTGATATTAATGGCT 4560
 GCTATAGGGTCTGACCGAAGGTCTCCAAAACAACCGAAGACCCACCACTATAATTACCGA
 R Y G L S A E L P K T P K Q T T I N I A
 ↓ -w12
 4561 CCGCCCATCTGATTAGGCCCTGAAGCAGCGACGAATACCCTTTGGAACTTCAACTGCT 4620
 GGCGGGTAGACTAATCCGGGGACTTCGTGCTGCTTATGGGAAACCTTTGAAGTTGACGA
 G G M Q N P G Q L L S S Y G K S V E V A
 4621 CCCAAATTGTTGGTCAGAATCCGCGCCAGATCGAGGTTGCCGTCATAGGGAACATAAATG 4680
 GGGTTTAACAACCGAGTCTTAGGCGCGGTCTAGCTCCAACGGCAGTATCCCTTGTATTTAC
 G L N N T L I R A L D L N G D Y P V Y I
 ↓ -12
 4681 GGCACACCGTCGAAATAGACCGGCACCTGACGACTATCAAAGCCACGAACTTTGACCTGT 4740
 CCGTGTGGCAGCTTTATCTGGCCGTGGACTGCTGATAGTTTTCGGTGCTTGAAACTGGACA
 P V G D F Y V P V Q R S D F G R V K V Q
 4741 TCTTCGTTGCGGCTACCTGACTTTTGCAGTACCACGCCGGGGACGACACTTAATGCCTGA 4800
 AGAAGCAACGCCGATGGACTGAAAACGTGATGGTGGCGCCCTGCTGTGAATTACGGACT
 E E N R S G S K Q L V V G P V V S L A Q
 4801 GCGACATTCTGTTTATCCAGGGCCTTCATGGTGGGTTGATCGAGAACGGTGGTCGTCGTT 4860
 CGCTGTAAGACAAATAGGTCCCGGAAGTACCACCCAAGTACTGCTTGGCCACCAGCAGCAA
 A V N Q K D L A K M T P Q D L V T T T T
 ↓ -w33
 4861 GATGATACCGGACTGGACCATACGGTCAGCGTGTGCTCTCCTGAGCTGCTACGGCCTGC 4920
 CTACTATGGCCTGACCTGGTATGCCAGTCGCACAGCGAGAGGACTCGACGATGCCGGACG
 S S V P S S W V T L T D S E Q A A V A Q
 4921 TGGGTAAAGGCAAGTGATAATGCTGTGCAGAGATAACGTTTTTTGAGTCTCATAGAATAA 4980
 ACCCATTTCCGTTCACTATTACGACACGTCTCTATTGCAAAAAACTCAGAGTATCTTATT
Q T F A L S L A T C L Y R K K L R M **OmpX**
 4981 ATTCCGTTTAAACAAAGAATTAATAGAGGCTGCGGGTTGCAGACTGACTTTAATGTCAGC 5040
 T**AAGG**CAAATTGTTTCTT**AAT**TTATCTCCGACGCCAACGTCTGACTGAAATTACAGTCG
SD
 * I S A A P Q L S V K I D A
 5041 AGGTGCCGCGTAGTAAGGCGCGGAGGTAATAAAAAGGCGAATGCCGCAGTCGAGGTAATT 5100
 TCCACGGCGCATCATTCCGCGCCTCATTATTTTTCCGCTTACGGCGTCAGCTCCATTA
 P A A Y Y P A S T I F L R I G C D L Y N
 5101 TTTGAGTGTGTCAGATTAATTCGCGCGGTGAGCCAGCGTGCAGTGGGGAGCCAGCGA 5160
 AAACCTCACAACAGTCTAATTAAGGCGGCCAGTCTCGGTGCGACGTCACCCCTCGGTGCT
 K L T T L N I G G T L A L T C H P A L S

5161 CGGTGCTATTTGGGCAATTTTCAGTTGCCTGTTGCGGACTAAATTTGTTCGAGCTGAAGTAC 5220
 -----+-----+-----+-----+-----+-----+-----+
 GCCACGATAAACCCGTTAAAGTCAACGGACAACGCCTGATTTAAACAGCTCGACTTCATG
 P A I Q A I E T A Q Q P S F K D L Q L V
 5221 GTCTGGTTGCGCGCGTAACGCGGCGATTGCCTCTTTCGGCGCGTCGGCTTCAACAACATAT 5280
 -----+-----+-----+-----+-----+-----+-----+
 CAGACCAACGCGCGCATTGCGCCGCTAACGGAGAAAGCCGCGCAGCCGAAGTTGTTGATA
 D P Q A R L A A I A E K P A D A E V V I
 5281 TTTCTTCTCGGGTGCCTGGCGACGTAACGATTGATTGCGCCTGACCAGTCTGATTGTC 5340
 -----+-----+-----+-----+-----+-----+-----+
 AAAGAAGAGCCCACGCACCGCTGCATTGACTAACTAACGCGGACTGGTCAGGACTAACAG
 K K E P A H R R L Q N I A G S W D Q N D
 5341 ATGAAGAAAATGGCGGTGGTTGGCAAACAATAATATGGTTTCCGCACATCCGGCACGATG 5400
 -----+-----+-----+-----+-----+-----+-----+
 TACTTCTTTTACCGCCACCAACCGTTTGTATTATACCAAAGCGTGTAGGCCGTGCTAC
 H L F H R H N A F L L I T E A C G A R H
 5401 AATCAGTCCTCCGGCAGCCAGAATTGCCTGCGAGGGCAGTAAGCGGGTACC CGGAATTGC 5460
 -----+-----+-----+-----+-----+-----+-----+
 TTAGTCAGGAGGCCGTCGGTCTTAACGGACGCTCCCGTCATTCGCCCATGGGCCTTAACG
 I L G G A A L I A Q S P L L R T G P I A
 5461 TTTTCGGGTGCAGGCGAATATTGCCATCAGGGTAACGTTACGAAGTAACGCCAGCATT 5520
 -----+-----+-----+-----+-----+-----+-----+
 AAAAGCCCACGTCCGCTTATAACGGTAGTCCCATTGCAAGTGCTTCATTGCGGTGCTAAA
 K R T C A F I A M L T V N V F Y R W C K
 5521 GAGCGAGATAATCAGAAACAACCGCAAACCTCCACTCCAGCACATTCTGGACTGCCTTCCA 5580
 -----+-----+-----+-----+-----+-----+-----+
 CTCGCTCTATTAGTCTTTGTTGGCGTTTGGAGTGAGGTGCGTGTAAAGACCTGACGGAAGGT
 L S I I L F L R L S W E L V N Q V A K W
 5581 TCCCTGATGAAGTGCTGCGGCATTACCTTGCGCACGGATTAGCCGCTGACCCGCGTTTCGC 5640
 -----+-----+-----+-----+-----+-----+-----+
 AGGGACTACTTCACGACGCCGTAATGGAACCGGTGCCTAATCGGCGACTGGGCGCAAGCG
 G Q H L A A A N G Q A R I L R Q G A N A
 5641 TTGTGAACCGTCGCTGACCGCGTCATCAATAGTTAATCCCAGCGTAGTTAACATCTTACA 5700
 -----+-----+-----+-----+-----+-----+-----+
 AACACTTGGCAGCGACTGGCGCAGTAGTTATCAATTAGGGTCGCATCAATTGTAGAATGT
 Q S G D S V A D D I T L G L T T L M K C
 5701 CGCGACAGAAATGCCGCTGACGCAGCCGCCCTGACGGAGAAAAAACTCTATATAGCCATG 5760
 -----+-----+-----+-----+-----+-----+-----+
 GCGCTGTCTTTACGGCGACTGCGTCGGCGGGACTGCCTCTTTTTTGAGATATATCGGTAC
 A V S I G S V C G G Q R L F F E I Y G H
 5761 CTGGTGTCCAATATTTAATGCTCGCGTGGTCAGGTCACCGCCCTGGATATCTTCCAGCAG 5820
 -----+-----+-----+-----+-----+-----+-----+
 GACCACAGGTTATAAATTACGAGCGCACCAGTCCAGTGGCGGGACCTATAGAAGGTCGTC
 Q H G I N L A R T T L D G G Q I D E L L
 5821 TAACGCATCGATTGTTGCCTGAGAAAGGAAAATCATCGCAGTGTCTCCTGCGGTACGCTA 5880
 -----+-----+-----+-----+-----+-----+-----+
 ATTGCGTAGCTAAACACGGACTCTTTCCTTTTAGTAGCGTCACAGAGGACGCCATGCGAT
 L A D I Q A Q S L F I M * R L T E Q P V S
 SD
 Orf2

↓ -w9

5881 TCCCACCAGATGAGCGCCCAGTCGCGGCTGGCAGTGGCGAGGGCACGAGCATCTTGTTGT 5940
 -----+-----+-----+-----+-----+-----+-----+
 AGGGTGGTCTACTCGCGGGTCAGCGCCGACCGTCACCGCTCCCGTGCTCGTAGAACACA
 D W W I L A W D R S A T A L A R A D Q Q

5941 TGATACCAGCGATATAACCGCTCACGCTCATCGTCGTTAAGCGCGCCCAGACTCCAGCTG 6000
 -----+-----+-----+-----+-----+-----+-----+
 ACTATGGTCGCTATATTGGCGAGTGCGAGTAGCAGCAATTCGCGCGGGTCTGAGGTCGAC
 Q Y W R Y L R E R E D D N L A G L S W S

6001 ACGTTCTGCTCAAAACGTTCCCAGGTGCTGTTATCCTGTTGGCAATTTTGTCCGCGGATA 6060
 -----+-----+-----+-----+-----+-----+-----+
 TGCAAGACGAGTTTTGCAAGGGTCCACGACAATAGGACAACCGTTAAAACAGGCGCCTAT
 V N Q E F R E W T S N D Q Q C N Q G R I

6061 AAATCAACGTGAGCATAAATTCCCATCTGATACAGGACGTTGAGCGCAAAGATATAGTTG 6120
 -----+-----+-----+-----+-----+-----+-----+
 TTTAGTTGCACTCGTATTTAAGGGTAGACTATGTCCTGCAACTCGCGTTTCTATATCAAC
 F D V H A Y I G M Q Y L V N L A F I Y N

↓ -w18

6121 GGTAATTCGATCACTTCCCGACCAGCCGCCCCGCTGAATGGCTGGAGAGACGAACGAGGTG 6180
 -----+-----+-----+-----+-----+-----+-----+
 CCATTAAGCTAGTGAAGGGCTGGTCGGCGGGCGACTTACCGACCTCTCTGCTTGCTCCAC
 P L E I V E R G A A R Q I A P S V F S T

6181 CTGACCAGATGAGTCGTATAGACGCGAAGCCGTGCCTGATTATTCAGTTTGCTCATAGCC 6240
 -----+-----+-----+-----+-----+-----+-----+
 GACTGGTCTACTCAGCATATCTGCGCTTCGGCACGGACTAATAAGTCAAACGAGTATCGG
 S V L H T T Y V R L R A Q N N L K S M A

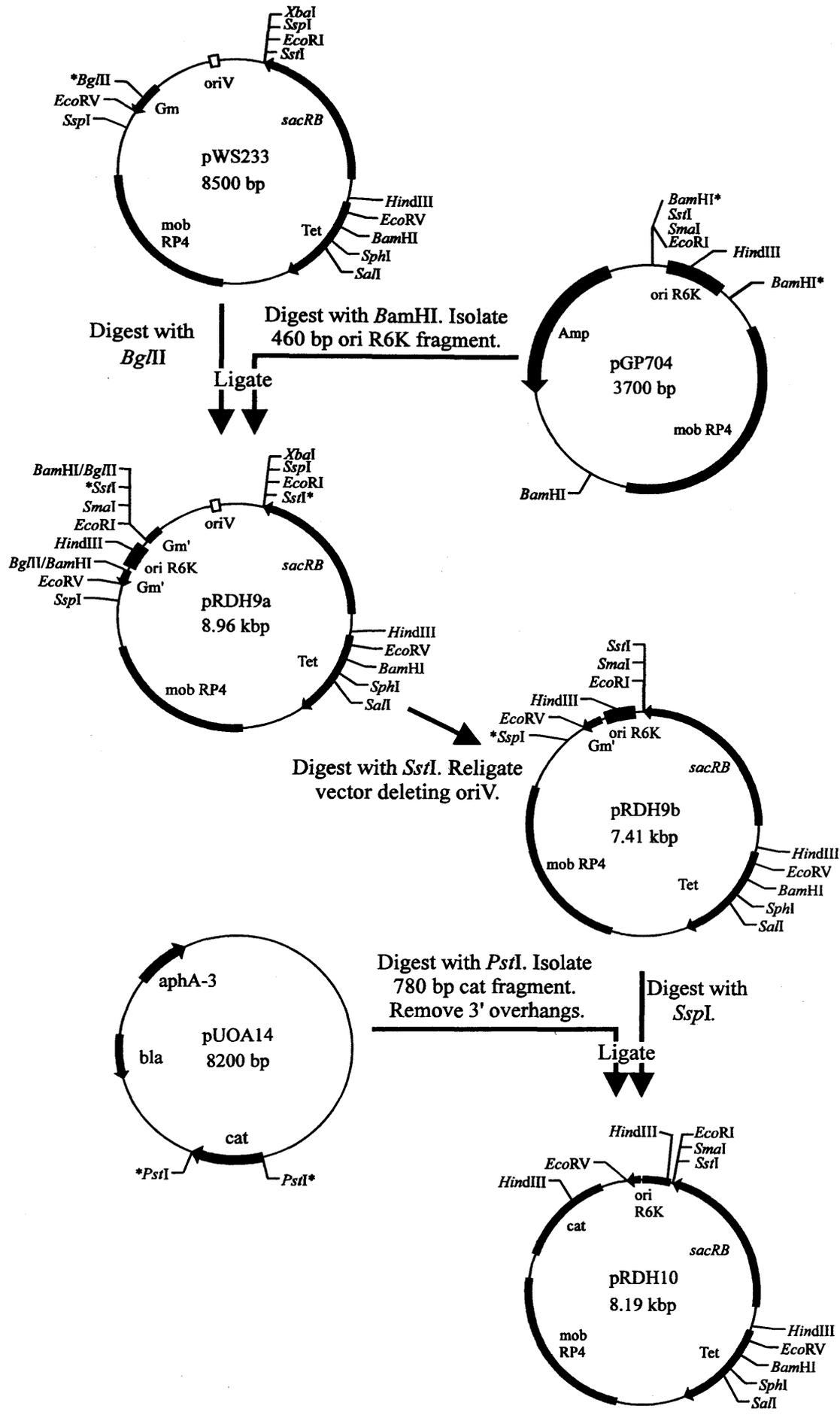
6241 TGGCGCATATCAGCCACCAGAGTGGAGCGGGAAGCCACGGCAATATCGCAGCGTGGCAGG 6300
 -----+-----+-----+-----+-----+-----+-----+
 ACCGCGTATAGTCGGTGGTCTCACCTCGCCCTTCGGTGCCGTTATAGCGTCGCACCGTCC
 Q R M D A V L T S R S A V A I D C R P L

BamHI

6301 TCACTCCAGTCCTCTTCCCAGGCGCGCTGGATCC 6334
 -----+-----+-----+-----+-----+
 AGTGAGGTCAGGAGAAGGGTCCGCGCGACCTAGG
 D S W D E E W A R Q I **Orf1'**

Appendix 3. Construction of pRDH10.

pRDH10 was based on the *sacRB*-dependent counterselectable allelic exchange vector pSW233 which was designed to construct unmarked mutations in *Rhizobium* (Selbitschka *et al.*, 1993). A 460 bp *Bam*HI fragment containing the *oriR6K* origin of replication was isolated from plasmid pGP704 (Miller and Mekalanos, 1988) and ligated into the *Bg*III site within the gentamicin resistance gene of pSW233. Ligations were transformed into *E. coli* strain DH5 α and transformants selected on LUA plates containing tetracycline. Colonies carrying the correct recombinant plasmid pRDH9a were differentiated from those carrying religated pSW233 by selection for gentamicin sensitivity. The *oriV* of pRDH9a was deleted by digestion with *Sst*I and religation of the remaining plasmid fragment. The religated vector was then transformed into strain SY327 λ pir (the π protein encoded by λ pir is essential for the replication of plasmids containing *oriR6K*; Kolter *et al.*, 1978) and transformants were selected on LUA plates containing tetracycline. Large colonies containing the recombinant plasmid pRDH9b were easily identifiable from the small colonies containing religated pRDH9a vector; the poor growth of the pRDH9a containing colonies was attributed to impaired plasmid replication resulting from *oriV* and *oriR6K* both being active when the plasmid is in SY327 λ pir. A blunt ended 780 bp fragment containing the chloramphenicol resistance gene (*cat*) was isolated from the *Campylobacter/E. coli* shuttle vector pUOA14 (Wang and Taylor, 1990) by digestion with *Pst*I and filling in the 3' overhangs using Klenow (Materials and Methods, section 2.9.3.1); this fragment was then ligated into pRDH9b vector linearised with *Ssp*I. The ligation was transformed into SY327 λ pir and colonies carrying recombinant plasmid pRDH10 were selected by growth on LUA plates containing tetracycline and chloramphenicol.



Literature Cited

- Adam, A. (1923) Über die Biologie der Dyspepsie coli und ihre Beziehungen zur Pathogenese der Dyspepsie und Intoxikation. *Jahrbuch für Kinderheilkunde und physische Erziehung* **101**:295-314.
- Adam, A. (1927) Dyspepsiecoli: zur Frage der bakteriellen Ätiologie der sogen. alimentären Intoxikation. *Jahrbuch für Kinderheilkunde und physische Erziehung* **106**:8-40.
- Adam, T., Arpin, M., Prevost, M. C., Gounon, P., and Sansonetti, P. J. (1995) Cytoskeletal rearrangements and the functional role of T-plastin during entry of *Shigella flexneri* into HeLa cells. *J. Cell Biol.* **129**:367-381.
- Adu-Bobie, J., Frankel, G., Bain, C., Goncalves, A. G., Trabulsi, L. R., Douce, G., Knutton, S., and Dougan, G. (1998) Detection of intimins α , β , γ , and δ , four intimin derivatives expressed by attaching and effacing microbial pathogens. *J. Clin. Microbiol.* **36**:662-668.
- Agin, T. S., Cantey, J. R., Boedeker, E. C., and Wolf, M. K. (1996) Characterization of the *eaeA* gene from rabbit enteropathogenic *Escherichia coli* strain RDEC-1 and comparison to other *eae* genes from bacteria that cause attaching effacing lesions. *FEMS Microbiol. Lett.* **144**:249-258.
- Agin, T. S., and Wolf, M. K. (1997) Identification of a family of intimins common to *Escherichia coli* causing attaching-effacing lesions in rabbits, humans and swine. *Infect. Immun.* **65**:320-326.
- Ahmer, B. M. M., Thomas, M. G., Larsen, R. A., and Postle, K. (1995) Characterization of the *exbBD* operon of *Escherichia coli* and the role of ExbB and ExbD in TonB function and stability. *J. Bact.* **177**:4742-4747.
- Albert, M. J., Alam, K., Islam, M., Montanaro, J., Rahman, A. S. M. H., Haider, K., Hossain, M. A., Kibriya, A. K. M. G., and Tzipori, S. (1991) *Hafnia alvei*, a probable cause of diarrhea in humans. *Infect. Imm.* **59**:1507-1513.
- Albert, M. J., Faruque, S. M., Ansaruzzaman, M., Islam, M. M., Haider, K., Alam, K., Kabir, I., and Robins-Browne, R. (1992) Sharing of virulence associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. *J. Med. Microbiol.* **37**:310-314.
- Allaoui, A., Menard, R., Sansonetti, P. J., and Parsot, C. (1993) Characterisation of the *Shigella flexneri* *ipgD* and *ipgF* genes, which are located in the proximal part of the *mxi* locus. *Infect. Immun.* **61**:1707-1714.
- Allaoui, A., Sansonetti, P. J., Menard, R., Barzu, S., Mounier, J., Philipon, A., and Parsot, C. (1995) MxiG, a membrane protein required for secretion of *Shigella* spp. Ipa invasins: involvement in entry into epithelial cells and in intercellular dissemination. *Mol. Microbiol.* **17**:461-470.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Altschul, S. F., Madden T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic. Acids Res.* **25**:3389-3402.

- Ames, G. F.-L. (1986) Bacterial periplasmic uptake systems: structure, mechanism, and evolution. *Annu. Rev. Biochem.* **55**:397-425.
- Ames, G. F.-L., and Joshi, A. K. (1990) Energy coupling in bacterial periplasmic permeases. *J. Bact.* **172**:4133-4137.
- Anantha, R. P., Stone, K. D., and Donnenberg, M. S. (1998a) Non-polar mutagenesis of the *bfp* gene cluster from enteropathogenic *Escherichia coli*. *Abstracts of the 98th General Meeting of the American Society for Microbiology*. B-155.
- Anantha, R. P., Stone, K. D., and Donnenberg, M. S. (1998b) Role of BfpF, a member of the PilT family of putative nucleotide-binding proteins, in type IV pilus biogenesis and in interactions between enteropathogenic *Escherichia coli* and host cells. *Infect. Immun.* **66**:122-131.
- Andrews, G. P., Hromockyj, A. E., Coker, C., and Maurelli, A.T. (1991) Two novel virulence loci, *mxiA* and *mxiB*, in *Shigella flexneri* C 2a facilitate excretion of invasion plasmid antigens. *Infect. Immun.* **59**:1997-2005.
- Angerer, A., and Braun, V. (1998) Iron regulates transcription of the *Escherichia coli* ferric citrate transport genes directly and through the transcription initiation proteins. *Arch. Microbiol.* **169**:483-490.
- Angerer, A., Gaisser, S., and Braun, V. (1990) Nucleotide sequence of the *sfuA*, *sfuB*, and *sfuC* genes of *Serratia marcescens* suggest a periplasmic-binding-protein-dependent iron transport mechanism. *J. Bact.* **172**:572-578.
- Archibald, F. (1983) *Lactobacillus plantarum*, an organism not requiring iron. *FEMS Microbiol. Lett.* **19**:29-32.
- Arduino, R. C., and DuPont, H. L. (1993) Traveler's diarrhoea. *Ballieres Clin. Gastroenterol.* **7**:365-385.
- Arnold, L. (1927) Diarrhea in infants. *Arch. Pediatr.* **44**:71-85.
- Arriaga, Y. L., Harville, B. A., and Dreyfus, L. A. (1995) Contribution of individual disulfide bonds to biological action of *Escherichia coli* heat-stable enterotoxin B. *Infect. Immun.* **63**:4715-4720.
- Azad, A. K., Sirakova, T. D., Fernandes, N. D., and Kolattukudy P. E. (1997) Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. *J. Biol. Chem.* **272**:16741-16745.
- Bachmann, B. J. (1983) Linkage map of *Escherichia coli* K-12 edition 7. *Microbiol. Rev.* **47**:180-230.
- Bachmann, B. J. (1996) Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. 2460-2488. In Neidhardt, F. C., Curtis III, R., Ingraham, J. L., Lin, E. E. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schacchter, M., and Umberger, H. E. (Ed). *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D. C.
- Bagg, A., and Neilands, J. B. (1987a) Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiol. Rev.* **51**:509-518.
- Bagg, A., and Neilands, J. B. (1987b) Ferric uptake regulation protein acts as a repressor, employing iron(II) as a cofactor to bind to the operator of an iron transport operon in *Escherichia coli*. *Biochemistry* **26**:5471-5477.

- Bahr, L. (1912) Untersuchungen über die Aetiologie der Cholera infantum. *Centralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten* 66:335-386.
- Bain, C., Keller, R., Collington, G. K., Trabulsi, L. R., and Knutton, S. (1998) Increased levels of intracellular calcium are not required for the formation of attaching and effacing lesions by enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect. Immun.* 66:3900-3908.
- Baldini, M. M., Kaper, J. B., Levine, M. M., Candy, D. C. A., and Moon, H. W. (1983) Plasmid mediated adhesion in enteropathogenic *Escherichia coli*. *J. Pediatr. Gastroenterol. Nutr.* 2:534-538.
- Baldini, M. M., Nataro, J. P., and Kaper, J. B. (1986) Localization of a determinant for HEp-2 adherence by enteropathogenic *Escherichia coli*. *Infect. Immun.* 52:334-336.
- Baldwin, T. J. (1990) Disease mechanisms of enteropathogenic *Escherichia coli*. Thesis. University of Leicester.
- Baldwin, T. J., Brooks, S. F., Knutton, S., Hernandez, H. A. M., Aitken, A., and Williams, P. H. (1990) Protein phosphorylation by protein kinase-C in HEp-2 cells infected with enteropathogenic *Escherichia coli*. *Infect. Immun.* 58:761-765.
- Baldwin, T.J., Knutton, S., Sellers, L., Hernandez, H. A. M., Aitkin, A., and Williams, P. H. (1992) Enteroaggregative *Escherichia coli* strains secrete a heat-labile toxin antigenically related to *E. coli* hemolysin. *Infect. Immun.* 60:2092-2095.
- Baldwin, T. J., LeeDelaunay, M. B., Knutton, S., and Williams, P. H. (1993) Calcium-calmodulin dependence of actin accretion and lethality in cultured HEp-2 cells infected with enteropathogenic *Escherichia coli*. *Infect. Immun.* 61:760-763.
- Baldwin, T. J., Ward, W., Aitken, A., Knutton, S., and Williams, P. H. (1991) Elevation of intracellular free calcium levels in HEp-2 cells infected with enteropathogenic *Escherichia coli*. *Infect. Immun.* 59:1599-1604.
- Bally, M., Filloux, A., Akrim, M., Ball, G., Lazdunski, A., and Tommassen, J. (1992) Protein secretion in *Pseudomonas aeruginosa*: characterisation of seven *xcp* genes and processing of secretory apparatus components by prepilin peptidase. *Mol. Microbiol.* 6:1121-1131.
- Bardwell, J. C., McGovern, K., and Beckwith, J. (1991) Identification of a protein required for disulphide formation *in vivo*. *Cell* 67:581-589.
- Barenberg, L. H., Levy, W., and Grand, M. J. H. (1936) An epidemic of infectious diarrhea in the new-born. *JAMA* 106:1256-1260.
- Baudry, B., Maurelli, A. T., Clerc, P., Sadoff, J. C., and Sansonetti, P. J. (1987) Localization of plasmid loci necessary for the entry of *Shigella flexneri* into HeLa cells, and characterisation of one locus encoding four immunogenetic polypeptides. *J. Gen. Microbiol.* 133: 3403-3413.
- Baughman, G., and Nomura, M. (1983) Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in *Escherichia coli*. *Cell* 34:979-988.
- Bäumler, A. J., and Hantke, K. (1992) Ferrioxamine uptake in *Yersinia enterocolitica*; characterisation of the receptor protein FoxA. *Mol. Microbiol.* 6:1309-1321.

- Bayer, M., Eferi, R., Zellnig, G., Teferle, K., Dijkstra, A., Koraimann, G., and Högenauer, G. (1991) Gene 19 of plasmid R1 is required for both efficient conjugative DNA transfer and bacteriophage R17 infection. *J. Bact.* **177**:4279-4288.
- Bayley, S. A. and Rees, D. A. (1986) Myosin light chain phosphorylation in fibroblast shape change, detachment and patching. *Eur. J. Cell. Biol.* **42**:10-16.
- Beckwith, J., Signer, E., and Epstein, W. (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**:77-90.
- Beebakhee, G., Louie, M., Deazavedo, J., and Brunton, J. (1992) Cloning and nucleotide-sequence of the *eae* gene homolog from enterohemorrhagic *Escherichia coli* serotype O157:H7. *FEMS Microbiol. Lett.* **91**:63-68.
- Beiber, D., Ramer, S. W., Cheng-Yen, W., Murray, W. J., Tobe, T., Fernandez, R., and Skoolnik, G. K. (1998) Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. *Science* **280**:2114-2118.
- Beinke, C., Laarmann, S., Wachter, C., Karch, H., Greune, L., and Schmidt, M. A. (1998) Diffusely adhering *Escherichia coli* strains induce attaching and effacing phenotypes and secrete homologs of Esp proteins. *Infect. Immun.* **66**:528-539.
- Benjamin, W. H., Turnbough, C. L., Posey, B. S., and Briles, D. (1985) The ability of *Salmonella typhimurium* to produce the siderophore enterochelin is not a virulence factor in mouse typhoid. *Infect. Immun.* **50**:392-397.
- Benz, I., and Schmidt, M. A. (1992) AIDA-1, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesised via a precursor molecule. *Mol. Microbiol.* **6**:1539-1546.
- Berger, B. R., and Christie, P. J. (1993) The *Agrobacterium tumefaciens virB4* gene product is an essential virulence protein requiring an intact nucleoside triphosphate-binding domain. *J. Bact.* **175**:1723-1734.
- Berridge, M. J. (1984) Inositol trisphosphate and diacylglycerol as second messengers. *Biochem. J.* **220**:345-360.
- Berridge, M. J. and Irvine, R. F. (1984) Inositol trisphosphate, a second messenger in cellular signal transduction. *Nature* **312**:414-321.
- Berlyn, M. K. B. (1998) Linkage map of *Escherichia coli* K-12, Edition 10: The traditional map. *Microbiol. Mol. Biol. Rev.* **62**:814-984.
- Bernet-Camard, M. F., Coconnier, M. H., Hudault, S., and Servin, A. L. (1996a) Differential expression of complement proteins and regulatory decay-accelerating factor in relation to differentiation of cultured human colon adenocarcinoma cell lines. *Gut* **38**:248-253.
- Bernet-Camard, M. F., Coconnier, M. H., Hudault, S., and Servin, A. L. (1996b) Pathogenicity of the diffusely-adhering strain *Escherichia coli* C1845: F1845 adhesin-decay accelerating factor interaction, brush border microvillus injury, and actin disassembly in cultured human intestinal epithelial cells. *Infect. Immun.* **64**:1818-1828.
- Bertani, G. (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bact.* **62**:293-300.
- Best, W. H. (1938) Epidemic diarrhea of the new-born. *JAMA* **110**:1155-1158.
- Bezkorovainy, A. (1987) Iron proteins. 27-67. In Bullen, J. J., and Griffiths, E. (Eds) *Iron and Infection: Molecular, Clinical and Physiological Aspects*. Wiley, Chichester.

- Bhatia, R. and Calvo, K. C. (1996) The sequencing expression, purification, and steady-state kinetic analysis of quinolinate phosphoribosyl transferase from *Escherichia coli*. *Arch. Biochem. Biophys.* **325**:270-278.
- Bilge, S. S., Clausen, C. R., Lau, W., and Moseley, S. L. (1989) Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *J. Bact.* **171**:4281-4289.
- Birnboim, H. C., and Doly, J. (1979) A rapid alkaline extraction procedure for the screening of recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.
- Black, R. E. (1990) Epidemiology of traveler's diarrhea and relative importance of various pathogens. *Rev. Infect. Dis.* **12**:S73-S79.
- Blake, P. A., Ramos, S., Macdonald, K. L., Rassi, V., Gomes, T. A. T., Ivey, C., Bean, N. H., and Trabulsi, L. R. (1993) Pathogen-specific risk factors and protective factors for acute diarrheal disease in urban Brazilian infants. *J. Infect. Dis.* **167**:627-632.
- Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**:1453-1474.
- Bliska, J. (1996) How pathogens exploit interactions mediated by SH3 domains. *Chem. Biol.* **3**:7-11.
- Blomfield, I. C., Vaughn, V., Rest, R. F., and Eisenstein, B. I. (1991) Allelic exchange in *Escherichia coli* using the *Bacillus subtilis* *sacB* gene and a temperature-sensitive pSC101 replicon. *Mol. Microbiol.* **5**:1447-1457.
- Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschape, H., and Hacker, J. (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. *Infect. Immun.* **62**:606-614.
- Blum, G., Falbo, B., Caprioli, A., and Hacker, J. (1995) Gene clusters encoding the cytotoxic necrotising factor type 1, Prs-fimbriae, and α -haemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microb. Lett.* **126**:189-196.
- Bochner, B. R. (1980) Positive selection for loss of tetracycline resistance. *J. Bact.* **143**:926-933.
- Bogdanove, A. J., Beer, S. V., Bonas, U., Boucher, C., Collmer, A., Coplin, D. L. *et al.* (1996) Unified nomenclature for broadly conserved *hrp* genes of phytopathogenic bacteria. *Mol. Microbiol.* **20**:681-683.
- Bower, J. R., Congenti, B. L., Cleary, T. G., Stone, R. T., Wanger, A., Murray, B. E., Mathewson, J. J., and Pickering, L. K. (1989) *Escherichia coli* O114: nonmotile as a pathogen in an outbreak of severe diarrhea associated with a day care center. *J. Infect. Dis.* **160**:243-247.
- Braun, V. (1995) Energy-coupled transport and signal transduction through the Gram-negative outer membrane via TonB-ExbB-ExbD-dependent receptor proteins. *FEMS Microbiol Revs.* **16**:295-307.
- Braun, V. (1997) Surface signaling: Novel transcription initiation mechanism starting from the cell surface. *Arch. Microbiol.* **167**:325-331.

- Braun, V. and Hantke, K. (1991) Genetics of bacterial iron transport. 107-130. In Winkelmann, G. (Ed.) *CRC Handbook of Microbiol Iron Chelates*. CRC Press Inc., Boca Raton, Fla.
- Braun, V., Hantke, K., and Koster, W. (1998) Bacterial iron transport: mechanisms, genetics and regulation. *Met. Ions Biol. Syst.* **35**:67-145.
- Bray, J. (1945) Isolation of antigenically homogenous strains of *Bact. coli neopolitanum* from summer diarrhoea of infants. *Journal of Pathology and Bacteriology* **57**:239-247.
- Bray, J., and Bevan, T. E. D. (1948) Slide agglutination of *Bacterium coli* var. *neopolitanum* in summer diarrhoea. *Journal of Pathology and Bacteriology* **60**:395-401.
- Brenner, D. J., Fanning, G. R., Miklos, G. V., and Steigerwalt, A. G. (1973) Polynucleotide sequence relatedness among *Shigella* species. *Int. J. Syst. Bacteriol.* **23**:1-7.
- Brickman, E., and Beckwith, J. (1975) Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and $\phi 80$ transducing phages. *J. Mol. Biol.* **96**:307-316.
- Brickman, T. J. and McIntosh, M. A. (1992) Overexpression and purification of ferric enterobactin esterase from *Escherichia coli*. Demonstration of enzymatic hydrolysis of enterobactin and its iron complex. *J. Biol. Chem.* **267**:12350-12355.
- Brinton, C. C. (1959) Non-flagellar appendages of bacteria. *Nature* **183**:782-786.
- Brock, J. H. (1989) The biology of iron. 35-54. In de Sousa, M., and Brock, J. H. (Eds) *Iron in Immunity, Cancer and Inflammation*. Wiley, Chichester.
- Brockman, R. W., and Heppel, L. A. (1968) On the localisation of alkaline phosphatase and cyclic phosphodiesterase in *Escherichia coli*. *Biochemistry* **7**:2554-2562.
- Brown, M. R. W., Anwar, H., and Lambert, P. A. (1984) Evidence that mucoid *Pseudomona aeruginosa* in the cystic fibrosis lung grows under iron restricted conditions. *FEMS Microbiol. Lett.* **21**:113-117.
- Byers, B. R. and Arceneaux, J. E. (1998) Microbial iron transport: iron acquisition by pathogenic microorganisms. *Met. Ions Biol. Syst.* **35**:37-66.
- Camara, L. M., Carbonare, S. B., Silva, M. L. M., and Carneirosampaio, M. M. S. (1994) Inhibition of enteropathogenic *Escherichia coli* (EPEC) adhesion to HeLa cells by human colostrum - detection of specific sIgA related to EPEC outer membrane proteins. *Int. Arch. Allergy Immunol.* **103**:307-310.
- Calderwood, S. B., and Mekalanos, J. J. (1988) Confirmation of the Fur operator site by insertion of a synthetic oligonucleotide into an operon fusion plasmid. *J. Bact.* **170**:1015-1017.
- Canil, C., Rosenshine, I., Ruschkowski, S., Donnenberg, M. S., Kaper, J. B., and Finlay, B. B. (1993) Enteropathogenic *Escherichia coli* decreases the transepithelial electrical resistance of polarized epithelial monolayers. *Infect. Immun.* **61**:2755-2762.
- Cantey, J. R., and Blake, R.K. (1977) Diarrhea due to *Escherichia coli* in the rabbit: a novel mechanism. *J. Infect. Dis.* **135**:454-462.
- Caprioli, A., Falbo, V., Ruggeri, F. M., Baldassarri, L., Bisicchia, R., Ippolito, G., Romoli, E., and Donelli, G. (1987) Cytotoxic necrotizing factor production by hemolytic strains of *Escherichia coli* causing extraintestinal infections. *J. Clin. Microbiol.* **25**:146-149.

- Carbonetti, N. H., Boonchai, S., Parry, S. H., Väisänen-Rhen, V., Korhonen, T. K., and Williams, P. H. (1986) Aerobactin-mediated iron uptake by *Escherichia coli* isolates from human extraintestinal infections. *Infect. Immun.* **51**:966-968.
- Casbadan, M. J., and Cohen, S. N. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.* **138**:179-207.
- Cathie, I. A. B., and MacFarlane, J. C. W. (1951) Incidence of *Bact. coli* O group 111 in sporadic infantile gastroenteritis. *Br. Med. J.* **2**:1002-1004.
- Chamberlain, J. P. (1979) Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Anal. Biochem.* **98**:132-135.
- Chen, W., and Kuo, T. (1993) A simple and rapid method for the preparation of Gram-negative bacterial genomic DNA. *Nucleic. Acids Res.* **21**:2260.
- Choa, K. L., and Dreyfus, L. A. (1997) Interaction of *Escherichia coli* heat-stable enterotoxin B with cultured human intestinal epithelial cells. *Infect. Immun.* **65**:3209-3217.
- Chou, P. Y., and Fasman, G. D. (1978) Empiric predictions of protein conformation. *Annu. Rev. Biochem.* **47**:251-276.
- Christiansen, M. (1917) Bakterien der Typhus-Coli-Gruppe im Darm von gesunden Spankälbern und bei deren Darminfektionen. *Centralblatt für Bakteriologie, Paraitenkunde und Infektionskrankheiten* **79**:196-248.
- Clausen, C. R., and Christie, D. L. (1982) Chronic diarrhea in infants caused by adherent *Escherichia coli*. *J. Pediatr.* **100**:358-361.
- Clowes, R. C. and Hayes, W. (1968) *Experiments in microbial genetics*. John Wiley and Sons Inc., New York.
- Collington, G. K., Booth, I. W., and Knutton, S. (1998) Rapid modulation of electrolyte transport in Caco-2 cell monolayers by enteropathogenic *Escherichia coli* (EPEC) infection. *Gut* **42**:200-207.
- Copass, M., Grandi, G., and Rappuoli, R. (1997) Introduction of unmarked mutations in the *Helicobacter pylori vacA* gene with a sucrose sensitivity marker. *Infect. Immun.* **65**:1949-52.
- Cornelis, G. R., and Wolf-Watz, H. (1997) The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. *Mol. Microbiol.* **23**:861-867.
- Coulton, J. W., Mason, P., Cameron, D. R., Carmel, G., Jean, R., and Rode, H. N. (1986) Protein fusions of β -galactosidase to the ferrichrome-iron receptor of *Escherichia coli* K-12. *J. Bact.* **165**:181-192.
- Cox, C. D. (1982) Effect of pyochelin on the virulence of *Pseudomonas aeruginosa*. *Infect. Immun.* **36**:17-23.
- Crane, J. K., and Oh, J. S. (1997) Activation of host cell protein kinase C by enteropathogenic *Escherichia coli*. *Infect. Immun.* **65**:3277-3285.
- Crane, J. K., Wehner, M. S., Bolen, E. J., Sando, J. J., Linden, J., Guerrant, R. L., and Sears, C. L. (1992) Regulation of intestinal guanylate cyclase by the heat-stable enterotoxin of *Escherichia coli* (STa) and protein kinase C. *Infect. Immun.* **60**:5004-5012.
- Cravioto, A., Gross, R. J., Scotland, S. M., and Rowe, B. (1979) An adhesive factor found in strains of *Escherichia coli* belonging to the traditional enteropathogenic serotypes. *Curr. Microbiol.* **3**:95-99.

- Cravioto, A., Reeyes, R. E., and Ortega, R. (1988) Prospective study of diarrhoeal disease in a cohort of rural Mexican children: incidence and isolated pathogens during the first two years of life. *Epidemiol. Rev.* **101**:123.
- Cravioto, A., Reeyes, R. E., Trujillo, F., Uribe, F., Navarro, A., de la Roca, J. M., Hernandez, J. M., Perez, G., and Vazquez, V. (1990) Risk of diarrhea during the first year of life associated with initial and subsequent colonization by specific enteropathogens. *Am. J. Epidemiol.* **131**:886-904.
- Cravioto, A., Tello, A., Villafan, H., Ruiz, J., Delvedovo, S., and Neeser, J. R. (1991) Inhibition of localized adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells by immunoglobulin and oligosaccharide fractions of human colostrum and breast milk. *J. Infect. Dis.* **163**:1247-1255.
- Creighton, C. (1975) *A history of epidemics in Britain. Vol 2.* Barnes and Noble, New York.
- Crosa, J. H. (1997) Signal transduction and transcriptional and postranscriptional control of iron-regulated genes in bacteria. *Microbiol. Mol. Biol. Rev.* **61**:319-336.
- Crowley, D. S., Ryan, M. J., and Wall, P. G. (1997) Gastroenteritis in children under 5 years of age in England and Wales. *Communicable Disease Report.* **7**:R82-R86.
- Cunliffe, H. E., Merriman, T. R., and Lamont, I. L. (1995) Cloning and characterization of *pvdS*, a gene required for pyoverdine synthesis in *Pseudomonas aeruginosa*: PvdS is probably an alternative sigma factor. *J. Bacteriol.* **177**:2744-2750.
- Czeczulin, J. R., Balepur, S., Hicks, S., Phillips, A., Hall, R., Kothary, M. H., Navarro-Garcia, F., and Nataro, J. P. (1997) Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect. Immun.* **65**:4135-4145.
- Dagbert, M. and Ehrlich, S. D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene.* **6**:23-28.
- Davis, L. S., Patel, S. S., Atkinson, J. P., and Lipsky, P. E. (1988) Decay-accelerating factor functions as a signal transducing molecule for human T cells. *J. Immunol.* **141**:2246-2252.
- Davison, W. C. (1925) The duodenal contents of infants in health and during and following diarrhea. *Am. J. Dis. Child* **29**:743-756.
- Day, N. P., Scotland, S. M., and Rowe, B. (1981) Comparison of an HEp-2 tissue culture test with the Serény test for detection of enteroinvasiveness in *Shigella* spp. and *Escherichia coli*. *J. Clin. Microbiol.* **17**:681-684.
- Dean, C. R., Neshat, S., and Poole, K. (1996) PfeR, an enterobactin-responsive activator of ferric enterobactin receptor gene expression in *Pseudomonas aeruginosa*. *J. Bact.* **178**:5361-5369.
- Dean, C. R., and Poole, K. (1993) Expression of the ferric enterobactin receptor (*pfeA*) of *Pseudomonas aeruginosa* - involvement of a 2-component regulatory system. *Mol. Microbiol.* **8**:1095-1103.
- Dean, D. (1981) A plasmid cloning vector for the direct selection of strains carrying recombinant plasmids. *Gene* **15**:99-102.
- Deibel, C., Krämer, S., Chakraborty, T., and Ebel, F. (1998) EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90 kDa protein. *Mol. Microbiol.* **28**:463-474.

- de Lorenzo, V., Giovannini, F., Herrero, M., and Neilands, J. B. (1988a) Metal ion regulation of gene expression. Fur repressor-operator interaction at the promoter region of the aerobactin system of pColV-K30. *J. Mol. Biol.* **203**:875-884.
- de Lorenzo, V., Giovannini, F., Herrero, M., and Neilands, J. B. (1988b) Fur (ferric uptake regulation) protein and CAP (catabolite-activator protein) modulate transcription of fur gene in *Escherichia coli*. *Eur. J. Biochem.* **173**:537-546.
- de Lorenzo, V., and Timmis, K. N. (1994) Analysis and construction of stable phenotypes in Gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol.* **235**:386-405.
- de Lorenzo, V., Wee, S., Herrero, M., and Neilands, J. B. (1987) Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulator (*fur*) repressor. *J. Bact.* **169**:2624-2630.
- Demchick, P., and Koch, A. L. (1996) The permeability of the cell wall fabric of *Escherichia coli* and *Bacillus subtilis*. *J. Bact.* **178**:768-773.
- Dersch, P., Schmidt, K., and Bremer, E. (1993) Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol. Microbiol.* **8**:875-889.
- de Rycke, J., Gonzalez, E. A., Blanco, J., Oswald, E., Blanco, M., and Boivin, R. (1990) Evidence for two types of cytotoxic necrotizing factor in human and animal clinical isolates of *Escherichia coli*. *J. Clin. Microbiol.* **28**:694-699.
- de Sauvage, F. J., Horuk, R., Bennet, G., Quan, C., Burnier, J. P., and Goeddel, D. V. (1992) Characterization of the recombinant human receptor for *Escherichia coli* heat-stable enterotoxin. *J. Biol. Chem.* **267**:6479-6482.
- Diderichsen, B. (1980) *flu*, a metastable gene controlling surface properties of *Escherichia coli*. *J. Bacteriol.* **141**:858-867.
- Dijkstra, A. J., Hermann, F., and Keck, W. (1995) Cloning and controlled expression of the gene encoding the 35 kDa soluble lytic transglycosylase from *Escherichia coli*. *FEBS Microbiol. Letts.* **366**:115-118.
- Dijkstra, A. J., and Keck, W. (1996a) Identification of new members of the lytic transglycosylase family in *Haemophilus influenzae* and *Escherichia coli*. *Microbial Drug Resistance.* **2**:141-145.
- Dijkstra, A. J., and Keck, W. (1996b) Peptidoglycan as a barrier to transenvelope transport. *J. Bact.* **178**:5555-5562.
- Dijkstra, B. W., and Thunnissen, A. M. W. H. (1994) Holy proteins .2. the soluble lytic transglycosylase. *Current Opinion Struct. Biol.* **4**:810-813.
- Dong, H., Nilsson, L., and Kurland, C. G. (1995) Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J. Bact.* **177**:1497-1504.
- Donnenberg, M. S. (1995) Enteropathogenic *Escherichia coli*. 709-726 In Blaser, M. J., Smith, P. D., Ravdin, J. I., Greenberg, H. B., and Guerrant, R. L. (Ed) *Infections of the gastrointestinal tract*. Raven Press, New York, N. Y.
- Donnenberg, M. S., Calderwood, S. B., Donohue-Rolfe, A., Keusch, G. T., and Kaper, J. B. (1990a) Construction and analysis of Tn ϕ A mutants of enteropathogenic *Escherichia coli* unable to invade HEp-2 cells. *Infect. Immun.* **58**:1565-1571.

- Donnenberg, M. S., Donohue-Rolfe, A., and Keusch, G. T. (1989) Epithelial cell invasion - an overlooked property of enteropathogenic *Escherichia coli* (EPEC) associated with the EPEC adherence factor. *J. Infect. Dis.* **160**:452-459.
- Donnenberg, M. S., Donohue-Rolfe, A., and Keusch, G. T. (1990b) A comparison of HEp-2 cell invasion by enteropathogenic and enteroinvasive *Escherichia coli*. *FEMS Microbiol. Lett.* **69**:83-86.
- Donnenberg, M. S., Girón, J. A., Nataro, J. P., and Kaper, J. B. (1992) A plasmid encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. *Mol. Microbiol.* **6**:3427-3437.
- Donnenberg, M. S., and Kaper, J. B. (1991) Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* **59**:4310-4317.
- Donnenberg, M. S., and Kaper, J. B. (1992) Enteropathogenic *Escherichia coli*. *Infect. Immun.* **60**:3953-3961.
- Donnenberg, M. S., Kaper, J. B., and Finlay, B. B. (1997a) Interactions between enteropathogenic *Escherichia coli* and host epithelial cells *Trends Microbiol.* **5**:109-114.
- Donnenberg, M. S., Tacket, C. O., James, S. P., Losonsky, G., Nataro, J. P., Wasserman, S. S., Kaper, J. B., and Levine, M. M. (1993a) Role of the *eaeA* gene in experimental enteropathogenic *Escherichia coli* infection. *J. Clin. Invest.* **92**:1412-1417.
- Donnenberg, M. S., Yu, J., and Kaper, J. B. (1993b) A second chromosomal gene necessary for intimate attachment of enteropathogenic *Escherichia coli* to epithelial cells. *J. Bact.* **175**:4670-4680.
- Donnenberg, M. S., Zhang, H. Z., and Stone, K. D. (1997b) Biogenesis of the bundle-forming pilus of enteropathogenic *Escherichia coli*: reconstitution of fimbriae in recombinant *E. coli* and role of DsbA in pilin stability - a review. *Gene* **192**:33-38.
- Dorman, C. (1992) The VirF protein from *Shigella flexneri* is a member of the AraC transcription factor superfamily and is highly homologous to Rns a positive regulator of virulence genes in enterotoxigenic *E. coli* *Mol. Microbiol.* **6**:1575-1575.
- Dorman, C. (1994) *Genetics of bacterial virulence*. Blackwell Scientific Publications, Oxford.
- Dorman, C. J. (1995) 1995 Flemming Lecture. DNA topology and the global control of bacterial gene expression: implications for the regulation of virulence gene expression. *Microbiology* **141**:1271-1280.
- Dower, W. J., Miller, J. F., and Ragsdale, C. W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **16**:6127-6145.
- Drachmann, R. H. (1974) Acute infectious enteritis. *Pediatr. Clin. North Am.* **21**:711-737.
- Draper, D. E. (1996) Translational initiation. 902-908. In Neidhardt, F. C., Curtis III, R., Ingraham, J. L., Lin, E. E. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schacchter, M., and Umberger, H. E. (Ed). *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D. C.
- Dreyfus, L. A., Urban, R. G., Whipp, S. C., Slaughter, C., Tachias, K., and Kupersztoch. (1992) Purification of STb enterotoxin of *Escherichia coli* and the role of selected amino acids in its secretion, stability and toxicity. *Mol. Microbiol.* **6**:2397-2406.

- Drolet, R., Fairbrother, J. M., Harel, J., Helie, P. (1994) Attaching and effacing and enterotoxigenic *Escherichia coli* associated with enteric colibacillosis in the dog. *Can. J. Vet. Res.* **58**:87-92.
- Dulaney, A. D. and Michelson, I. D. (1935) A study of *B. coli mutabile* from an outbreak of diarrhea in the new-born. *Am. J. Public Health* **25**:1241-1251.
- Dums, F., Dow, J. M., and Daniels, M. J. (1991) Structural characterisations of protein secretion genes of the bacterial phytopathogen *Xanthomonas campestris* pathovar *campestris*: relatedness to secretion systems of other bacteria. *Mol. Gen. Genet.* **229**:357-364.
- DuPont, H. L., and Ericsson, C. D. (1993) Prevention and treatment of traveler's diarrhea. *N. Engl. J. Med.* **328**:1821-1827.
- DuPont, H. L., Formal, S. B., Hornick, R. B., Snyder, M. J., Libonati, J. P., Sheahan, D. G., LaBrec, E. H., and Kalas, J. P. (1971) Pathogenesis of *Escherichia coli* diarrhea. *N. Engl. J. Med.* **285**:1-9.
- Dytoc, M., Fedorko, L., and Sherman, P. M. (1994) Signal transduction in human epithelial cells infected with attaching and effacing *Escherichia coli* in vitro. *Gastroenterology* **106**:1150-1161.
- Earhart, C. F. (1996) Uptake and metabolism of iron and molybdenum. 1075-1090. In Neidhardt, F. C., Curtis III, R., Ingraham, J. L., Lin, E. E. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schacchter, M., and Umberger, H. E. (Ed). *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D. C.
- Ebel, F., Deibel, C., Kresse, A. U., Guzman, C. A., and Chakraborty, T. (1996) Temperature dependent and medium dependent secretion of proteins by Shiga toxin producing *Escherichia coli*. *Infect. Immun.* **64**:4472-4479.
- Edelman, R., and Levine, M. M. (1983) Summary of a workshop on enteropathogenic *Escherichia coli*. *J. Infect. Dis.* **147**:1108-1118.
- Eick-Helmerich, K., and Braun, V. (1989) Import of biopolymers into *Escherichia coli*: nucleotide sequences of the *exbB* and *exbD* genes are homologous to those of the *tolQ* and *tolR* genes, respectively. *J. Bact.* **171**:5117-5126.
- Eisenstein, B. (1981) Phase variation of type 1 fimbriae in *Escherichia coli* is under transcriptional control. *Science* **214**:337-339.
- Eisenstein, B. (1988) Type 1 fimbriae of *Escherichia coli*: genetic regulation, morphogenesis and role in pathogenesis. *Rev. Infect. Dis.* **10**:S341-S344.
- Ehlert, K., Holtje, J. V., and Templin, M. F. (1995) Cloning and expression of a murein hydrolase lipoprotein from *Escherichia coli*. *Mol. Microbiol.* **16**:761-768.
- Elliot, S. J., and Kaper, J. B. (1997) Role of type 1 fimbriae in EPEC infections. *Microb. Pathogen.* **23**:113-118.
- Elliot, S. J., Mellies, J. L., Wainwright, L. A. and Kaper, J. B. (1998a) Analysis of ORFs from the LEE pathogenicity island of enteropathogenic *E. coli* (EPEC). *Abstracts of the 98th General Meeting of the American Society for Microbiology.* B-154.
- Elliott, S. J., Wainwright, L. A., McDaniel, T. K., Jarvis, K. G., Deng, Y. K., Lai, L. C., McNamara, B. P., Donnenberg, M. S., and Kaper, J. B. (1998b) The complete

- sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol. Microbiol.* **28**:1-4.
- Emerson, S. U., Tokuyasu, K., and Simon, M. I. (1970) Bacterial flagella: polarity of elongation. *Science* **276**:190-192.
- Engel, H., Kazemier, B., and Keck, W. (1991) Murein-metabolising enzymes from *Escherichia coli*: sequence analysis and controlled over-expression of the *slt* gene which encodes the soluble lytic transglycosylase. *J. Bact.* **173**:6773-6782.
- Enz, S., Braun, V., and Crosa, J. H. (1995) Transcription of the region encoding the ferric dicitrate transport system in *Escherichia coli* - similarity between promoters for *fecA* and for extracytoplasmic function sigma factors. *Gene* **163**:13-18.
- Escherich, T. (1885) Die Darmbakterien des Neugeborenen und Säuglings. *Fortschr. Med.* **3**:515-522, 547-554.
- Escherich, T. (1886) *Die Darmbakterien des Säuglings und ihre Beziehungen zur Physiologie der Verdauung*. F. Enke, Stuttgart.
- Eslava, C., Villaseca, J., Morales, R., Navarro, A., and Ceavioto, A. (1993) Identification of a protein with toxigenic activity produced by enteroaggregative *Escherichia coli*. B-105. In Abstracts of the 93rd General Meeting of the American Society for Microbiology 1993. American Society for Microbiology. Washington D. C.
- Fasano, A., Fiorentini, C., Donelli, G., Uzzau, S., Kaper, J., Mangaretten, K., Ding, X., Gaundalini, S., Comstock, L., and Goldblum, S. (1995) Zonula occludens toxin modulates tight junctions through protein kinase C-dependent actin reorganisation in vitro. *J. Clin. Invest.* **96**:710-720.
- Fee, B. E., and Dempsey, W. B. (1988) Nucleotide sequence of gene X of antibiotic resistance plasmid R100. *Nucleic Acids Res.* **16**:4726.
- Felsenstein, J. (1989) PHYLIP - phylogeny inference package (version 3.2). *Cladistics.* **5**:164-166.
- Ferguson, W. W. and June, R. C. (1952) Experiments on feeding adult volunteers with *Escherichia coli* 111, B4, a coliform organism associated with infant diarrhea. *American Journal of Hygiene.* **55**:155-169.
- Finkel, S. E., and Johnson, R. C. (1993) The Fis protein: it's not just for DNA inversion anymore. *Mol. Microbiol.* **7**:1023.
- Finlay, B. B., DeVinney, R., Kenny, B., and Stein, M. (1998) EPEC delivers the goods: Response. *Trends Microbiol.* **6**:172-173.
- Finlay, B. B., and Falkow, S. (1997) Common themes in microbial pathogenicity revisited. *Microbiol. Mol. Biol. Rev.* **61**:136-169.
- Finlay, B. B., Rosenshine, I., Donnenberg, M. S., and Kaper, J. B. (1992) Cytoskeletal composition of attaching and effacing lesions associated with enteropathogenic *Escherichia coli* adherence to HeLa cells. *Infect. Immun.* **60**:2541-2543.
- Fischer, J., Maddox, C., Moxley, R., Kinden, D., and Miller, M. (1994) Pathogenicity of a bovine attaching effacing *Escherichia coli* isolate lacking Shiga-like toxins. *Am. J. Vet. Res.* **55**:991-999.
- Fitch, W. M., and Margoliash, E. (1967) Construction of phylogenetic trees. *Science.* **155**:279-284.

- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. -F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L. -I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*. **269**:449-604.
- Flitter, R. W., Rowley, D. A., and Halliwell, B. (1983) Superoxide dependent formation of hydroxide radicals in the presence of iron salts. *FEBS Lett.* **158**:310-312.
- Fogh, J., Wright, W. C., and Loveless, J. D. (1977) Absence of HeLa cell contamination in 169 cell lines derived from human tumours. *J. Natl. Cancer Inst.* **58**:209-214.
- Folkhard, W., Marvin, D. A., Watts, T. H., and Paranchych, W. (1984) Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. *Mol. Biol.* **149**:79-93.
- Fondacaro, J., and Henderson, L. (1985) Evidence for protein kinase C as a regulator of intestinal electrolyte transport. *Am. J. Physiol.* **249**:G422-G426.
- Foubister, V., Rosenshine, I., Donnenberg, M. S., and Finlay, B. B. (1994a) The *eaeB* gene of enteropathogenic *Escherichia coli* is necessary for signal transduction in epithelial cells. *Infect. Immun.* **62**:3038-3040.
- Foubister, V., Rosenshine, I., and Finlay, B. B. (1994b) A diarrheal pathogen, enteropathogenic *Escherichia coli* (EPEC), triggers a flux of inositol phosphates in infected epithelial cells. *J. Exp. Med.* **179**:993-998.
- Francis, C. L., Jerse, A. E., Kaper, J. B., and Falkow, S. (1991) Characterization of interactions of enteropathogenic *Escherichia coli* O127-H6 with mammalian cells *in vitro*. *J. Infect. Dis.* **164**:693-703.
- Frankel, G., Candy, D. C. A., Everest, P., and Dougan, G. (1994) Characterization of the C-terminal domains of intimin like proteins of enteropathogenic and enterohemorrhagic *Escherichia coli*, *Citrobacter freundii*, and *Hafnia alvei*. *Infect. Immun.* **62**:1835-1842.
- Frankel, G., Candy, D. C. A., Fabiani, E., Adu-Bobie, J., Gil, S., Novakova, M., Phillips, A. D., and Dougan, G. (1995) Molecular characterization of a carboxy terminal eukaryotic cell binding domain of intimin from enteropathogenic *Escherichia coli*. *Infect. Immun.* **63**:4323-4328.
- Frankel, G., Lider, O., Hershkoviz, R., Mould, A. P., Kachalsky, S. G., Candy, D. C. A., Cahalon, L., Humphries, M. J., and Dougan, G. (1996a) The cell binding domain of intimin from enteropathogenic *Escherichia coli* binds to β_1 integrins. *J. Biol. Chem.* **271**:20359-20364.
- Frankel, G., Phillips, A. D., Novakova, M., Field, H., Candy, D. C. A., Schauer, D. B., Douce, G., and Dougan, G. (1996b) Intimin from enteropathogenic *Escherichia coli* restores murine virulence to a *Citrobacter rodentium eaeA* mutant - induction of an immunoglobulin A response to intimin and EspB. *Infect. Immun.* **64**:5315-5325.
- Freestone, P., Trinei, M., Clarke, S. C., Nyström, T., and Norris, V. (1998) Tyrosine phosphorylation in *Escherichia coli*. *J. Mol. Biol.* **279**:1045-51.

- Frost, G. E., and Rosenberg, H. (1973) The inducible citrate-dependent iron transport system in *Escherichia coli* K-12. *Biochim. Biophys. Acta.* **330**:90-101.
- Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., and Kado, C. I. (1985) Positive selection procedure for the entrapment of insertion sequence elements in Gram-negative bacteria. *J. Bact.* **164**:918-921.
- Gay, P., Le Coq, D., Steinmetz, M., Ferrari, E., and Hoch, J. A. (1983) Cloning the structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. *J. Bact.* **153**:1424-1431.
- Geissler, S., and Drummond, M. (1993) A counterselectable pACYC184-based *lacZ* alpha-complementing plasmid vector with novel multiple cloning sites; construction of chromosomal deletions in *Klebsiella pneumoniae*. *Gene* **136**:253-255.
- Gibson, F., and Magrath, D. J. (1969) The isolation and characterization of a hydroxamic acid (aerobactin) formed by *Aerobacter aerogenes* 62-1. *Biochim. Biophys. Acta.* **192**:175-184.
- Giles, C., Sangster, G., and Smith, J. (1949) Epidemic gastroenteritis of infants in Aberdeen during 1947. *Arch. Dis. Child* **24**:45-53.
- Girón, J. A., Ho, A. S. Y., and Schoolnik, G. K. (1991) An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* **254**:710-713.
- Girón, J. A., Ho, A. S. Y., and Schoolnik, G. K. (1993) Characterization of fimbriae produced by enteropathogenic *Escherichia coli*. *J. Bact.* **175**:7391-7403.
- Goldberg, M. B., Barzu, O., Parsot, C., and Sansonetti, P. J. (1993) Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. *J. Bacteriol.* **175**:2189-2196.
- Goldberg, M. B., and Sansonetti, P. J. (1993) *Shigella* subversion of the cellular cytoskeleton: a strategy for epithelial colonization. *Infect. Immun.* **61**:4941-4946.
- Goldschmidt, R. (1933) Untersuchungen zur Ätiologie der Durchfallserkrankungen der Säuglinge. *Jahrbuch für Kinderheilkunde und physische Erziehung* **139**:318-358.
- Gomes, T. A. T., Rassi, V., Macdonald, K. L., Ramos, S. R. T. S., Trabulsi, L. R., Vieira, M. A. M., Guth, B. E. C., Candeias, J. A. N., Ivey, C., Toledo, M. R. F., and Blake, P. A. (1991) Enteropathogens associated with acute diarrheal disease in urban infants in Sao Paulo, Brazil. *J. Infect. Dis.* **164**:331-337.
- Gomes, T. A. T., Blake, A., and Trabulsi, L. R. (1989) Prevalence of *Escherichia coli* strains with localized, diffuse, and aggregative adherence to HeLa cells in infants with diarrhea and matched controls. *J. Clin. Microbiol.* **27**:266-269.
- Gomez, O., Donnenberg, M. S., and Kaper, J. B. (1993) Plasmid encoded regulator A (PerA) is a transcriptional activator of *eaeA* and *eaeB* of enteropathogenic *Escherichia coli*. B-310. In Abstracts of the 93rd General Meeting of the American Society for Microbiology 1993. American Society for Microbiology. Washington D. C.
- Gomez-Duarte, O. G., and Kaper, J. B. (1995) A plasmid encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. *Infect. Immun.* **63**:1767-1776.
- González-Gil, G., Bringmann, P., and Kahmann, R. (1996) FIS is a regulator of metabolism in *Escherichia coli*. *Mol. Microbiol.* **22**: 21-29.

- González-Pastor, J. E., San Millán, J. L., Castilla, M. A., and Moreno, F. (1995) Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7. *J. Bact.* **177**:7131-7140.
- Gorbach, S. L., Banwell, J. G., Chatterjee, B. D., Jacobs, B., and Sack, R. B. (1971) Acute undifferentiated diarrhoea in the tropics. I Alterations in intestinal microflora. *J. Clin. Invest.* **50**:881-889.
- Graus, H., Hodl, A., Wallner, P., and Högenauer, G. (1990) The sequence of the leading region of the resistance plasmid R1. *Nucleic Acids Res.* **18**:1046.
- Graus-Goldner, A., Graus, H., Schlacher, T., and Hogenauer, G. (1990) The sequences bordering *oriT* in the enterotoxin plasmid P307: comparison with the sequences of plasmids F and R1. *Plasmid* **24**:119-131.
- Greenberg, M., and Wronker, B. M. (1938) An outbreak of epidemic diarrhea in the newborn. *JAMA* **110**:563-566.
- Greenwood, K. T., and Luke, R. K. J. (1978) Enzymatic hydrolysis of enterobactin and its iron complex in *Escherichia coli* K-12. *Biochim. Biophys. Acta.* **525**:209-218.
- Griffiths, E., and Humphreys, J. (1980) Isolation of enterochelin from the peritoneal washings of guinea pigs lethally infected with *Escherichia coli*. *Infect. Immun.* **28**:286-289.
- Griffiths, E., Stevenson, P., Thorpe, R., and Chart, H. (1985) Naturally occurring antibodies in human sera that react with the iron-regulated outer membrane proteins of *Escherichia coli*. *Infect Immun.* **47**:808-813.
- Griffiths, E., Stevenson, P., and Joyce, P. (1983) Pathogenic *Escherichia coli* express new outer membrane proteins when growing in vivo. *FEMS Microbiol. Lett.* **16**:95-99.
- Griggs, D. W., and Konisky, J. (1989) Mechanism for iron-regulated transcription of the *Escherichia coli* *cir* gene: metal-dependent binding of the Fur protein to the promoter. *J. Bact.* **171**:1048-1054.
- Gross, R. (1993) Signal transduction and virulence regulation in human and animal pathogens. *FEMS Microbiol. Rev.* **10**:301-326.
- Gross, R. J., Scotland, S. M., and Rowe, B. (1976) Enterotoxin testing of *Escherichia coli* causing epidemic infantile enteritis in the U. K. *Lancet* **1**:629-630.
- Gunzberg, S. T., Chang, B. J., Elliot, S. J., Burke, V., and Gracey, M. (1993) Diffuse and enteroaggregative patterns of adherence of enteric *Escherichia coli* isolated from aboriginal children from the Kimberley region of western Australia. *J. Infect. Dis.* **167**:755-758.
- Gurwith, M. J., Wiseman, D. A., and Chow, P. (1977) Clinical and laboratory assessment of the pathogenicity of serotyped enteropathogenic *Escherichia coli*. *J. Infect. Dis.* **135**:736-743.
- Guyer, D. M., Kao, J.-S., and Mobley, H. L. T. (1998) Characterisation and analysis of pathogenicity island sequences in *Escherichia coli*. *Abstracts of the 98th General Meeting of the American Society for Microbiology.* B-167.
- Hacker, J., Blum-Oehler, G., Mühldorfer, I., and Tschäpe, H. (1997) Pathogenicity islands of virulent bacteria: structure function and impact on microbial evolution. *Mol. Microbiol.* **23**:1089-1097.

- Haigh, R., Baldwin, T., Knutton, S., and Williams, P. H. (1995) Carbon dioxide regulated secretion of the EaeB protein of enteropathogenic *Escherichia coli*. *FEMS Microbiol. Lett.* **129**:63-67.
- Hale, T. L., Oaks, E. V., and Formal, S. B. (1985) Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infect. Immun.* **50**:620-629.
- Hales, B. A., Hart, C. A., Batt, R. M., and Saunders, J. R. (1992) The large plasmids found in enterohemorrhagic and enteropathogenic *Escherichia coli* constitute a related series of transfer defective Inc F-IIa replicons. *Plasmid.* **28**:183-193.
- Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557-580.
- Hancock, R. W., Hantke, K., and Braun, V. (1976) Iron transport in *Escherichia coli* K-12: involvement of the colicin B receptor and of a citrate-inducible protein. *J. Bact.* **127**:1370-1375.
- Hantke, K. (1981) Regulation of ferric iron transport in *Escherichia coli* K-12: Isolation of a constitutive mutant. *Mol. Gen. Genet.* **182**:288-292.
- Hantke, K. (1983) Identification of an iron uptake system specific for coprogen and rhodotorulic acid in *Escherichia coli* K12. *Mol. Gen. Genet.* **191**:301-306.
- Hantke, K. (1990) Dihydroxybenzoylserine - a siderophore for *Escherichia coli*. *FEMS Microbiol. Letts.* **67**:5-8.
- Harle, C., Kim, I., Angerer, A., and Braun, V. (1995) Signal transfer through 3 compartments - transcription initiation of the *Escherichia coli* ferric citrate transport system from the cell surface. *EMBO J.* **14**:1430-1438.
- Harlow, E., and Lane, D. (1988) *Antibodies: A laboratory manual*. Cold Spring Harbour Laboratory, Cold Spring Harbour, NY, USA.
- Hase, C. C., Bauer, M. E. and Finkelstein, R. A. (1994) Genetic characterization of mannose-sensitive hemagglutinin (MSHA)-negative mutants of *Vibrio cholerae* derived by Tn5 mutagenesis. *Gene.* **150**:17-25.
- Hecht, G., Pestic, L., Nikcevic, G., Koutsouris, A., Tripuraneni, J., Lorimer, D. D., Nowak, G., Guerriero, V., Elson, E. L., and de Lanerolle, P. (1996) Expression of the catalytic domain of myosin light chain kinase increases paracellular permeability. *Am. J. Physiol.* **271**:C1678-C1684.
- Hedberg, C. W., Savarino, S. J., Besser, J. M., Paulus, C. J., Thelen, V. M., Myers, L. J., Cameron, D. N., Barrett, T. J., Kaper, J. B., Osterholm, M. T., Boyer, W., Kairis, F., Gabriel, L., Soler, J., Gyswyt, L., Bray, S., Carlson, R., Hooker, C., Fasano, A., Jarvis, K., McDaniel, T., and Tornieporth, N. (1997) An outbreak of foodborne illness caused by *Escherichia coli* O39:NM, an agent not fitting into the existing scheme for classifying diarrheogenic *E. coli*. *J. Infect. Dis.* **176**:1625-1628.
- Heinrichs, D. E., and Poole, K. (1993) Cloning and sequence-analysis of a gene (*pchR*) encoding an AraC family activator of pyochelin and ferripyochelin receptor synthesis in *Pseudomonas aeruginosa*. *J. Bact.* **175**:5882-5889.
- Heinrichs, D. E., and Poole, K. (1996) PchR, a regulator of ferripyochelin receptor gene (*fptA*) expression in *Pseudomonas aeruginosa*, functions both as an activator and as a repressor. *J. Bact.* **178**:2586-2592.

- Heinrichs, D. E., Young, L., and Poole, K. (1991) Pyochelin-mediated iron transport in *Pseudomonas aeruginosa* - involvement of a high molecular mass outer membrane protein. *Infect. Immun.* **59**:3680-3684.
- Heller, K., Mann, B. J., and Kadner, R. J. (1985) Cloning and expression of the gene for the vitamin B-12 receptor protein in the outer membrane of *Escherichia coli*. *J. Bact.* **161**:896-903.
- Henderson, I. R., Meehan, M., and Owen, P. (1997) A novel regulatory mechanism for a novel phase-variable outer membrane protein of *Escherichia coli*. 349-355. In Paul, P. S., Francis, D. H. and Benfield, D. A. (Eds) *Mechanisms in the Pathogenesis of Enteric Diseases*. Plenum Press, New York.
- Henderson, I., and Owen, P. (1996) Cloning and sequencing of the gene encoding the phase-variable outer membrane protein from *Escherichia coli*. *Abstr. Irish Branch Meet. Soc. Gen. Microbiol.* **13**.
- Henikoff, S. (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* **155**:156-165.
- Herweg, J. C., Middelkamp, N., and Thornton, H. K. (1956) *Escherichia coli* diarrhea: the relationship of certain serotypes of *Escherichia coli* to sporadic and epidemic cases of infantile diarrhea. *J. Pediatr.* **49**:629-641.
- Hicks, S., Candy, D. C. A., and Phillips, A. D. (1996) Adhesion of enteroaggregative *Escherichia coli* to pediatric intestinal mucosa in vitro. *Infect. Immun.* **64**:4751-4760.
- Hicks, S., Frankel, G., Kaper, J. B., Dougan, G., and Phillips, A. D. (1998) Role of intimin and bundle-forming pili in enteropathogenic *Escherichia coli* adhesion to pediatric intestinal tissue in vitro. *Infect. Immun.* **66**:1570-1578.
- Hider, R. C. (1984) Siderophore mediated absorption of iron. *Structure and Bonding (Berlin)*. **58**:25-87.
- Higgins, C. F. (1992) ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**:67-113.
- Higgins, C. F., Hyde, S. C., Mimmack, M. M., Gileadi, U., Gill, D. R., and Gallagher, M. P. (1990) Binding protein-dependent transport systems. *J. Bioenerg. Biomembr.* **22**:571-592.
- Higgins, C. F., Dorman, C. J., Stirling, D. A., Waddell, L., Booth, I. R., May, G., and Bremer, E. (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. *Cell* **52**:569-584.
- Higgins, D. C., and Sharp, P. M. (1989) Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS*. **5**:151-153.
- Hobbs, M., Collie, E. S. R., Free, P. D., Livingston, S. P., and Mattick, J. S. (1993) PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **7**:669-683.
- Hobbs, M., and Mattick, J. S. (1993) Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes. *Mol. Microbiol.* **10**:233-243.

- Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M., and Schuster, P. (1994) Fast Folding and Comparison of RNA Secondary Structures *Monatsh.Chem.* **125**:67-188.
- Hoffman, C. S., and Wright, A. (1985) Fusions of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. *Proc. Natl. Acad. Sci. USA.* **82**:5107-5111.
- Höltje, J. V. (1996a) Molecular interplay of murein synthases and murein hydrolases in *Escherichia coli*. *Microbial Drug Resistance-Mechanisms Epidemiology and Disease* **2**:99-103.
- Höltje, J. V. (1996b) A hypothetical holoenzyme involved in the replication of the murein sacculus of *Escherichia coli*. *Microbiology UK* **142**:1911-1918.
- Höltje, J. V. (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **62**:181-203.
- Höltje, J.-V., Mirelman, D., Sharon, N., and Schwarz, U. (1975) Novel type of murein transglycosylase in *Escherichia coli*. *J. Bact.* **124**:1067-1076.
- Höltje, J.-V., and Tuomanen, E. I. The murein hydrolases of *Escherichia coli*: properties, functions and impact on the course of infections in vivo. *J. Gen. Microbiol.* **137**:441-454.
- Hueck, C. J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**:379-435.
- Hussein, S., Hantke, K., and Braun, V. (1981) Citrate-dependent iron transport system in *Escherichia coli* K-12. *Eur. J. Biochem.* **117**:431-437.
- Isberg, R. R., Voorhis, D. L., and Falkow, S. (1987) Identification of invasins: a protein that allows enteric bacteria to penetrate cultured mammalian cells. *Cell* **50**:769-778.
- Ismaili, A., McWhirter, E., Handelsman, M. Y., Brunton, J. L., and Sherman, P. M. (1998) Divergent signal transduction responses to infection with attaching and effacing *Escherichia coli*. *Infect. Immun.* **66**:1688-1696
- Ismaili, A., Philpott, D. J., Dytoc, M. T., and Sherman, P. M. (1995) Signal transduction responses following adhesion of verocytotoxin producing *Escherichia coli*. *Infect. Immun.* **63**:3316-3326.
- Jacobson, S. H., Tullus, K., Wretling, B., and Brauner, A. (1988) Aerobactin-mediated uptake of iron by strains of *Escherichia coli* causing acute pyelonephritis and bacteraemia. *J. Infect.* **16**:147-152.
- Jarvis, K. G., Giron, J. A., Jerse, A. E., McDaniel, T. K., Donnenberg, M. S., and Kaper, J. B. (1995) Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc. Natl. Acad. Sci. USA.* **92**:7996-8000.
- Jarvis, K. G., and Kaper, J. B. (1996) Secretion of extracellular proteins by enterohemorrhagic *Escherichia coli* via a putative type III secretion system. *Infect. Immun.* **64**:4826-4829.
- Jensen, C. O. (1892-3) Ueber die Kälberruhr und deren Aetiologie. *Monatsschrift für praktische Tierheilkunde* **4**:97-124.
- Jerse, A. E., and Kaper, J. B. (1991) The *eae* gene of enteropathogenic *Escherichia coli* encodes a 94 kilodalton membrane protein, the expression of which is influenced by the EAF plasmid. *Infect. Immun.* **59**:4302-4309.

- Jerse, A. E., Yu, J., Tall, B. D., and Kaper, J. B. (1990) A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proc. Natl. Acad. Sci. USA* 87:7839-7843.
- Jiang, B., and Howard, S. P. (1992) The *Aeromonas hydrophila* *exeE* gene required for both protein secretion and normal outer membrane biogenesis, is a member of the general secretion pathway. *Mol. Microbiol.* 6:1351-1361.
- Johnson, R. P., Clarke, R. C., Wilson, J. B., Read, S. C., Rahn, K., Renwick, S. A., Sandhu, K. A., Alves, D., Karmali, M. A., Lior, H., McEwen, S. A., Spika, J. S., and Gyles, C. L. (1996) Growing concerns and recent outbreaks involving non-O157:H7 serotypes of verotoxigenic *Escherichia coli*. *J. Food Prot.* 59:1112-1122.
- Jones, R. L., Peterson, C. M., Grady, R. W., Kumbaraci, T., Cerami, A., and Graziano, J. H. (1977) Effects of iron chelators and iron overload on *Salmonella* infection. *Nature* 267:63-65.
- June, R. C., Ferguson, W. W., and Worfel, M. T. (1953) Experiments in feeding adult volunteers with *Escherichia coli* 55, B_s, a coliform organism associated with infant diarrhea. *American Journal of Hygiene* 57:222-236.
- Kammer, G. M., Walter, E. I., and Medof, M. E. (1988) Association of cytoskeletal re-organisation with capping of the complement decay-accelerating factor on T lymphocytes. *J. Immunol.* 141:2924-2928.
- Kamoun, S., Tola, E., Kamdar, H., and Kado, C. I. (1992) Rapid generation of directed and unmarked deletions in *Xanthomonas*. *Mol. Microbiol.* 6:809-816.
- Kaniga, K., Delor, I., and Cornelis, G. R. (1991) A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene* 109:137-141.
- Kao, J.-S., Stucker, D. M., Warren, J. W., and Mobley, H. L. T. (1997) Pathogenicity island sequences of pyelonephritogenic *Escherichia coli* CFT073 are associated with virulent uropathogenic strains. *Infect. Immun.* 65:2812-2820.
- Kaper, J. B. (1994) Molecular pathogenesis of enteropathogenic *Escherichia coli*. 173-195. In Miller, V. L., Kaper, J. B., Portnoy, D. A., and Isberg, R. R. (Eds) *Molecular Genetics of Bacterial Pathogenesis*. American Society for Microbiology, Washington.
- Kaper, J. B. (1998) EPEC delivers the goods. *Trends Microbiol.* 6:169-173.
- Karmali, M. A. (1989) Infection by verocytotoxin-producing *Escherichia coli*. *Clin. Microbiol. Rev.* 2:15-38.
- Karmali, M. A., Petric, M., Lim, C., Fleming, P. C., and Steele, B. T. (1983a) *Escherichia coli* cytotoxin, haemolytic-uraemic syndrome and haemorrhagic colitis. *Lancet* ii:1299-1300.
- Karmali, M. A., Steele, B. T., Petric, M., and Lim, C. (1983b) Sporadic cases of haemolytic uremic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. *Lancet* i:619-620.
- Kauffmann, F. (1947) The serology of the *coli* group. *J. Immunol.* 57:71-100.
- Kauffmann, F., and DuPont, A. (1950) *Escherichia* strains from infantile epidemic gastroenteritis. *Acta Pathologica et Microbiologica Scandinavica* 27:552-564.

- Kaufman, M. R., Shaw, C. E., Jones, I. D., and Taylor, R. K. (1993) Biogenesis and regulation of the *Vibrio cholerae* toxin-coregulated pilus: analogies to other virulence factor secretory systems. *Gene* 126:43-49.
- Kaufman, M. R., and Taylor, R. K. (1994) Identification of bacterial cell-surface virulence determinants with TnphoA. *Methods Enzymol.* 235:426-448.
- Kenny, B., Abe, A., Stein, M., and Finlay, B. B. (1997a) Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. *Infect. Immun.* 65:2606-2612.
- Kenny, B., DeVinney, R., Stein, M., Reinscheid, D. J., Frey, E. A., and Finlay, B. B. (1997b) Enteropathogenic E-coli (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 91:511-520.
- Kenny, B., and Finlay, B. B. (1995) Protein secretion by enteropathogenic *Escherichia coli* is essential for transducing signals to epithelial cells. *Proc. Natl. Acad. Sci. USA* 92:7991-7995.
- Kenny, B., and Finlay, B. B. (1997) Intimin-dependent binding of enteropathogenic *Escherichia coli* to host cells triggers novel signaling events, including tyrosine phosphorylation of phospholipase C- γ 1. *Infect. Immun.* 65:2528-2536.
- Kenny, B., Lai, L. C., Finlay, B. B., and Donnenberg, M. S. (1996) EspA, a protein secreted by enteropathogenic *Escherichia coli*, is required to induce signals in epithelial cells. *Mol. Microbiol.* 20:313-323.
- Keon, R. G., Fu, R., and Voordouw, G. (1997) Deletion of two downstream genes alters expression of the *hmc* operon of *Desulfovibrio vulgaris* subsp. *vulgaris* Hildenborough. *Arch. Microbiol.* 167:376-383.
- Kim, I., Stiefel, A., Plantor, S., Angerer, A., and Braun, V. (1997) Transcription induction of the ferric citrate transport genes via the N-terminus of the FecA outer membrane protein, the Ton system and the electrochemical potential of the cytoplasmic membrane. *Mol. Microbiol.* 23:333-344.
- Kim, S. R., and Komano, T. (1997) The plasmid R64 thin pilus identified as a type IV pilus. *J. Bact.* 179:3594-3603.
- Kingsley, R. A. (1997) Iron uptake and metabolism by *Salmonella enterica*. Thesis, University of Leicester.
- Kirby, A. C., Hall, E. G., and Coackley, W. (1950) Neonatal diarrhoea and vomiting: outbreaks in the same maternity unit. *Lancet* 2:201-207.
- Klapproth, J. M., Donnenberg, M. S., Abraham, J. M., Mobley, H. L. T., and James, S. P. (1995) Products of enteropathogenic *Escherichia coli* inhibit lymphocyte activation and lymphokine production. *Infect. Immun.* 63:2248-2254.
- Knutton, S., AduBobie, J., Bain, C., Phillips, A. D., Dougan, G., and Frankel, G. (1997) Down regulation of intimin expression during attaching and effacing enteropathogenic *Escherichia coli* adhesion. *Infect. Immun.* 65:1644-1652.
- Knutton, S., Baldini, M. M., Kaper, J. B., and McNeish, A. S. (1987a) Role of plasmid encoded adherence factors in adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells. *Infect. Immun.* 55:78-85.
- Knutton, S., Baldwin, T., Williams, P. H., and McNeish, A. S. (1988) New diagnostic test for enteropathogenic *Escherichia coli*. *Lancet* 1:1337.

- Knutton, S., Baldwin, T. J., Haigh, R. D., Williams, P. H., Manjarrez-Hernandez, A., and Aitken, A. (1994) Intracellular changes in "attaching and effacing" adhesion. 215-222. *In* Karmali, M. A., and Goglio, A. G. (Eds) *Recent advances in verocytotoxin-producing Escherichia coli infections*. Elsevier Science B. V., Amsterdam, Netherlands.
- Knutton, S., Baldwin, T., Williams, P. H., and McNeish, A. S. (1989a) Actin accumulation at sites of bacterial adhesion to tissue culture cells - basis of a new diagnostic test for enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect. Immun.* **57**:1290-1298.
- Knutton, S., Collington, G. K., Baldwin, T. J., Haigh, R. D., and Williams, P. H. (1996) Cellular responses to EPEC infection. *Rev. Microbiol. Sao Paulo* **27**(Suppl. 1):89-94
- Knutton, S., Lloyd, D. R., and McNeish, A. S. (1987b) Adhesion of enteropathogenic *Escherichia coli* to human intestinal enterocytes and cultured human intestinal mucosa. *Infect. Immun.* **55**:69-77.
- Knutton, S., Phillips, A., Shaw, R., Price, E., and Watson, P. (1989b) Evaluation of a specific test for attaching and effacing enteropathogenic *Escherichia coli*. *Pediatr. Res.* **26**:273.
- Knutton, S., Phillips, A. D., Smith, H. R., Gross, R. J., Shaw, R., Watson, P., and Price, E. (1991) Screening for enteropathogenic *Escherichia coli* in infants with diarrhea by the fluorescent actin staining test. *Infect. Immun.* **59**:365-371.
- Knutton, S., Rosenshine, I., Pallen, M. J., Nisan, I., Neves, B. C., Bain, C., Wolff, C., Dougan, G., and Frankel, G. (1998) A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *EMBO J.* **17**:2166-2176.
- Koch, A. L., and Woeste, S. (1994) Elasticity of the sacculus of *Escherichia coli*. *J. Bact.* **174**:4811-4819.
- Koebnik, R. (1993) Structural organisation of TonB-dependent receptors. *Trends Microbiol.* **1**:201.
- Kolter, R., Inuzuka, M., and Helinski, D. R. (1978) Transcomplementation-dependent replication of a low molecular weight origin fragment from plasmid R6K. *Cell.* **15**:1199-1208.
- Konowalchuk, J., Speirs, J. L., and Stavric, S. (1977) Vero response to a cytotoxin of *Escherichia coli*. *Infect. Immun.* **18**:775-779.
- Kopecko, D. J., Baron, L. S., and buysse, J. (1985) Genetic determinants of virulence in Shigella and dysenteric strains of *Escherichia coli* : their involvement in the pathogenesis of dysentery. *Curr. Top. Microbiol. Immunol.* **118**:71-95.
- Koster, W. (1991) Iron(III) hydroxamate transport across the cytoplasmic membrane of *Escherichia coli*. *BioMetals* **4**:23-32.
- Koster, M., Vandevossenberg, J., Leong, J., and Weisbeek, P. J. (1993) Identification and characterization of the *pupB* gene encoding an inducible ferric-pseudobactin receptor of *Pseudomonas putida* wcs358. *Mol. Microbiol.* **8**:591-601.
- Koster, M., Vanklompenburg, W., Bitter, W., Leong, J., and Weisbeek, P. (1994) Role for the outer membrane ferric siderophore receptor *pupB* in signal transduction across the bacterial cell envelope. *EMBO J.* **13**:2805-2813.

- Koya, G., Kosakai, N., and Fukasawa, Y. (1954a) Supplementary studies on the multiplication of *Escherichia coli* O-111 B4 in the intestinal tract of adult volunteers and its relation to manifestation of coli enteritis. *Jpn. J. Med. Sci. Biol.* 7:655-661.
- Koya, G., Kosakai, N., Kono, M., Mori, M., and Fukasawa, Y. (1954b) Observations on the multiplication of *Escherichia coli* O-111 B4 in the intestinal tract of adult volunteers in feeding experiments: the intubation study with Miller-Abbott's double lumen tube. *Jpn. J. Med. Sci. Biol.* 7:197-203.
- Kraft, A. R., Templin, M. F., and Höltje, J-V. (1998) Endo-glycosylase a novel representative of bacterial lytic transglycosylases. *J. Bact.* 180: 3441-3447.
- Krone, W. J. A., Steghuis, F., Koningstein, G., van Doorn, C., Roosedaal, B., de Graaf, F. K., and Oudege, B. (1985) Characterisation of the pColV-K30 encoded cloacin DF13/aerobactin outer membrane receptor protein of *Escherichia coli*; isolation and purification of the protein and analysis of its nucleotide structure and primary structure. *FEMS Microbiol. Lett.* 26:153-161.
- Labischinski, H., Goodell, E. W., Goodell, A., and Hochberg, M. L. (1979) Direct proof of a "more-than-single-layered" peptidoglycan architecture of *Escherichia coli* W7: a neutron small-angle scattering study. *J. Bact.* 173:751-756.
- Lacroix, J., Delage, G., Gosselin, F., and Chicoine, L. (1984) Severe protracted diarrhea due to multiresistant adherent *Escherichia coli*. *Am. J. Dis. Child* 138:693-696.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.
- Lai, L. C., Wainwright, L. A., Stone, K. D., and Donnenberg, M. S. (1997) A third secreted protein that is encoded by the enteropathogenic *Escherichia coli* pathogenicity island is required for transduction of signals and for attaching and effacing activities in host cells. *Infect. Immun.* 65:2211-2217.
- Lam, C., Turnowsky, F., Schwarzing, E. coli and Neruda, W. (1984) Bacteria recovered without subculture from infected human urines expressed iron regulated outer membrane proteins. *FEMS Microbiol. Lett.* 24:255-259.
- Law, D., Wilkie, K. M., Freeman, R., and Gould, F. K. (1992) The iron uptake mechanisms of enteropathogenic *Escherichia coli*: use of haem and haemoglobin during growth in an iron-limited environment. *J. Med. Microbiol.* 37:15-21.
- Lawlor, K. M., Daskaleros, P. A., Robinson, R. E., and Payne, S. M. (1987) Virulence of iron transport mutants of *Shigella flexneri* and utilisation of host compounds. *Infect. Immun.* 55:594-599.
- Lee, C. A. (1997) Type III secretion systems: machines to deliver bacterial proteins into eukaryotic cells? *Trends Microbiol.* 5:148-156.
- Lencer, W. I., Constable, C., Moe, S., Jobling, M. G., Webb, H. M., Ruston, S., Madara, J. L., Hirst, T. R., and Holmes, R. K. (1995) Targeting of cholera toxin and *Escherichia coli* heat-labile toxin in polarized epithelial cells: role of COOH-terminal KDEL. *J. Cell Biol.* 131:951-962.
- Lennox, E. S. (1955) Transduction of linked genetic characters of the host by bacteriophage P1. *Virology.* 1:190-206.
- Leong, J. M., Fournier, R. S., and Isberg, R. R. (1990) Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasin protein. *EMBO J.* 9:1979-1989.

- Lesage, A. A. (1897) Contribution à l'étude des entérites infantiles - séro-diagnostic des races de *Bacterium coli*. *C. R. Soc. Biol. (Paris)* 49:900-901.
- Letain, T. E., and Postle, K. (1997) TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Mol. Microbiol.* 24:271-283.
- Levine, M. M. (1985) *Escherichia coli* infections. *N. Engl. J. Med.* 313:445-447.
- Levine, M. M. (1987) *Escherichia coli* that cause diarrhea - enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J. Infect. Dis.* 155:377-389.
- Levine, M. M., Bergquist, E. J., Nalin, D. R., Waterman, D. H., Hornik, R. B., Young, C. R., Sotman, S., and Rowe, B. (1978) *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* 1:1119-1122.
- Levine, M. M., and Edelman, R. (1984) Enteropathogenic *Escherichia coli* of classical serotypes associated with infant diarrhea: epidemiology and pathogenesis. *Epidemiol. Rev.* 6:31-51.
- Levine, M. M., Nataro, J. P., Karch, H., Baldini, M. M., Kaper, J. B., Black, R. E., Clements, M. L., and O'Brien, A. D. (1985) The diarrheal response of humans to some classic serotypes of enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. *J. Infect. Dis.* 152:550-559.
- Lindeberg, M., and Collmer, A. (1992) Analysis of eight *out* genes in a cluster required for pectic enzyme secretion by *Erwinia chrysanthemi*: sequence comparison with secretion genes from other Gram-negative bacteria. *J. Bact.* 174:7385-7397.
- Linggood, M. A., Roberts, M., Ford, S., Parry, S. H., and Williams, P. H. (1987) Incidence of the aerobactin iron uptake system among *Escherichia coli* isolates from infections of farm animals. *J. Gen. Microbiol.* 133:835-842.
- Link, A. J., Phillips, D., and Church, G. M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterisation. *J. Bact.* 179:6228-6237.
- Liu, H., Haga, K., Yasumoto, K., Ohashi, Y., Yoshikawa, H., and Takahashi, H. (1997) Sequence and analysis of a 31 kb segment of the *Bacillus subtilis* chromosome in the area of the *rnnH* and *rnnG* operons. *Microbiology.* 143:2763-2767.
- Liu, L., Whalen, W., Das, A., and Berg, C. M. (1984) Rapid sequencing of cloned DNA using a transposon for bidirectional priming: sequence of *Escherichia coli* K-12 *autA* gene. *Nucl. Acid. Res.* 15:9461-9469.
- Loh, S., Cram, D., and Surray, R. (1989) Nucleotide sequence of the leading region adjacent to the origin of transfer on plasmid F and its conservation amongst conjugative plasmids. *Mol. Gen. Genet.* 219:177-186.
- Long-Krug, S. A., Weikel, C. S., Tiemens, K. T., Hewlett, E. L., and Levine, M. M. (1984) Does enteropathogenic *Escherichia coli* produce heat-labile enterotoxin, heat-stable enterotoxin a or b, or cholera toxin A subunits? *Infect. Immun.* 46:612-614.
- Lundrigan, M., and Earhart, C. F. (1981) Reduction in three iron-regulated outer membrane proteins and protein A by the *Escherichia coli perA* mutation. *J. Bact.* 146:804-807.

- Lundrigan, M., and Kadner, R. J. (1986) Nucleotide sequence of the gene for the ferrienterochelin receptor FepA in *Escherichia coli*. Homology among outer membrane receptors that interact with TonB. *J. Biol. Chem.* **261**:10797-10801.
- Macnab, R. M. (1992) Genetics and biogenesis of bacterial flagellae. *Annu. Rev. Genet.* **26**:131-158.
- Madara, J. L., Patapoff, T. W., Gillece-Gastro, B., Colgan, S. P., Parkos, C. A., Delp, C., and Mrsny, R. J. (1993) 5'-adenosine monophosphate is the neutrophil-derived paracrine factor that elicits chloride secretion from T84 intestinal epithelial cell monolayers. *J. Clin. Invest.* **91**:2320-2325.
- Maloy, S. R., and Nunn, W. D. (1981) Selection for loss of tetracycline resistance in *Escherichia coli*. *J. Bact.* **145**:1110-1112.
- Manjarrez-Hernandez, H. A., Amess, B., Sellers, L., Baldwin, T. J., Knutton, S., Williams, P. H., and Aitken, A. (1991) Purification of a 20 kDa phosphoprotein from epithelial cells and identification as a myosin light chain - phosphorylation induced by enteropathogenic *Escherichia coli* and phorbol ester. *FEBS Lett.* **292**:121-127.
- Manjarrez-Hernandez, H. A., Baldwin, T. J., Aitken, A., Knutton, S., and Williams, P. H. (1992) Intestinal epithelial cell protein phosphorylation in enteropathogenic *Escherichia coli* diarrhea. *Lancet* **339**:521-523.
- Manjarrez-Hernandez, H. A., Baldwin, T. J., Williams, P. H., Haigh, R., Knutton, S., and Aitken, A. (1996) Phosphorylation of myosin light chain at distinct sites and its association with the cytoskeleton during enteropathogenic *Escherichia coli* infection. *Infect. Immun.* **64**:2368-2370.
- Manning, P. A. (1997) The *tcp* gene cluster of *Vibrio cholerae*. *Gene* **192**:63-70.
- Manning, P. A., and Meyer, T. F. (1997) Type 4 pili: Biogenesis, Adhesins, Protein Export and DNA Import - Schloss Ringberg, Rottach-Egern, Germany, 26-29 November 1995 - Preface. *Gene* **192**:1.
- Manoil, C., and Beckwith, J. (1985) *TnphoA*: A transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA.* **82**:8129-8133.
- Mattila, L. (1994) Clinical features and duration of traveler's diarrhea in relation to its etiology. *Clin. Infect. Dis.* **19**:728-734.
- Matzanke, B. F. (1991) Structures, coordination chemistry and functions of microbial iron chelates. 15-64. In Winkelmann, G. (Ed) *CRC Handbook of Microbiol Iron Chelates*. CRC Press Inc., Boca Raton, Fla.
- Maupin-Furlow, J. A., Rosentel, J. K., Lee, J. H., Deppenmeier, U., Gunsalus, R. P., and Shanmugan, K. T. (1995) Genetic analysis of the *modABCD* (molybdate transport) operon of *Escherichia coli*. *J. Bact.* **177**:4851-4856.
- Maurelli, A. T. (1994) Virulence protein export systems in *Salmonella* and *Shigella*: a new family or lost relatives. *Trends Cell Biol.* **4**:240-242.
- Mayer, B. J., and Eck, M. J. (1995) SH3 domains. Minding your p's and q's. *Curr. Biol.* **5**:364-367.
- McDaniel, T. K., Jarvis, K. G., Donnenberg, M. S., and Kaper, J. B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Nat. Acad. Sci. USA* **92**:1664-1668.

- McDaniel, T. K., and Kaper, J. B. (1997) A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol. Microbiol.* **23**:399-407.
- McNamara, B. P., and Donnenberg, M. D. (1998) A novel proline-rich protein, EspF, is secreted from enteropathogenic *Escherichia coli* via the type III export pathway. *Abstracts of the 98th General Meeting of the American Society for Microbiology.* B-150.
- Melby, K., Slordahl, S., Guttenberg, T. J., and Nordbro, S. A. (1982) Sceptacaemia due to *Yersinia enterocolitica* after oral overdose of iron. *Br. Med. J.* **285**:467-468.
- Menard, R., Prevost, M. C., Gounon, P., Sansonetti, P. J., and Dehio, C. (1996) The secreted Ipa complex of *Shigella flexneri* promotes entry into mammalian cells. *Proc. Natl. Acad. Sci. USA* **93**:1254-1258.
- Menard, R., Sansonetti, P. J., and Parsot, C. (1993) Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J. Bacteriol.* **175**:5899-5906.
- Mezoff, A. G., Giannella, R. A., Eade, M. N., and Cohen, M. B. (1992) *Escherichia coli* enterotoxin (STA) binds to receptors, stimulates guanylate cyclase, and impairs absorption in rat colon. *Gastroenterology* **102**:816-822.
- Miller, J. H. (1992) *A short course in bacterial genetics: a laboratory manual and handbook for Escherichia coli and related bacteria.* Cold Spring Harbour Laboratory Press.
- Miller, V. L., and Mekalanos, J. J. (1988) A novel suicide vector and its use in construction of insertion mutants: Osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J. Bact.* **170**:2575-2583.
- Mills, M., and Payne, S. M. (1995) Genetics and regulation of heme iron transport in *Shigella dysenteriae* and detection of an analogous system in *Escherichia coli* O157-H7. *J. Bact.* **177**:3004-3009.
- Mills, M., and Payne, S. M. (1997) Analysis of *shuA*, the gene encoding the heme receptor of *Shigella dysenteriae* and analysis of invasion and intracellular multiplication of a *shuA* mutant. *Infect. Immun.* **65**:5358-5363.
- Miras, I., Hermant, D., Arricau, N., and Popoff, M. Y. (1995) Nucleotide sequence of *iagA* and *iagB* genes involved in invasion of HeLa cells by *Salmonella typhi* subsp. *enterica* ser. Typhi Res. *Microbiol.* **6**:17-20.
- Missiakas, D., and Satish, R. (1998) The extracytoplasmic function sigma factors: role and regulation. *Mol. Microbiol.* **28**:1059-1066.
- Miyazaki, H., Kato, H., Nakazawa, T., and Tsuda, M. (1995) A positive regulatory gene, *pvdS*, for expression of pyoverdinin biosynthetic genes in *Pseudomonas aeruginosa* pao. *Mol. Gen. Genet.* **248**:17-24.
- Mobley, H. L., Jarvis, K. G., Elwood, J. P., Whittle, D. I., Lockett, C. V., Russell, R. G., Johnson, D. E., Donnenberg, M. S., and Warren, J. W. (1993) Isogenic P-fimbrial deletion mutants of pyelonephritogenic *Escherichia coli*: the role of alpha Gal(1-4) beta Gal binding in virulence of a wild-type strain. *Mol. Microbiol.* **10**:143-155.
- Moeck, G. S., and Coulton, J. W. (1998) TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport. *Mol. Microbiol.* **28**:675-681.

- Moon, H. W., Whipp, S. C., Argenzio, R. A., Levine, M. M., and Giannella, R. A. (1983) Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect. Immun.* **41**:1340-1351.
- Montgomerie, J. Z., Bindereif, A., Neilands, J. B., Kalmanson, G. M., and Guze, L. B. (1984) Association of hydroxamate siderophore (aerobactin) with *Escherichia coli* isolated from patients with bacteraemia. *Infect. Immun.* **46**:835-838.
- Nakae, T., and Nikaido, H. (1975) Outer membrane as a diffusion barrier in *Salmonella typhimurium*. *J. Biol. Chem.* **250**:7359-7365.
- Nakai, K. and Kanehisa, M. (1991) Expert system for predicting protein localization sites in Gram-negative bacteria. *PROTEINS: Structure, Function, and Genetics.* **11**: 95-110.
- Nassif, X., and Sansonetti, P. J. (1986) Correlation of the virulence of *Klebsiella pneumoniae* K1 and K2 with the presence of a plasmid encoding aerobactin. *Infect. Immun.* **54**:603-608.
- Nassif, X., Mazert, M. C., Mounier, J., and Sansonetti, P. J. (1987) Evaluation with an *iuc::Tn10* mutant of the role of aerobactin production in the virulence of *Shigella flexneri*. *Infect. Immun.* **55**:1963-1969.
- Nataro, J. P., Baldini, M. M., Kaper, J. B., Black, R. E., Bravo, N., and Levine, M. M. (1985a) Detection of an adherence factor of enteropathogenic *Escherichia coli* with a DNA probe. *J. Infect. Dis.* **152**:560-565.
- Nataro, J. P., Deng, Y., Maneval, D. R., German, A. L., Martin, W. C., and Levine, M. M. (1992) Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. *Infect. Immun.* **60**:2297-2304.
- Nataro, J. P., and Kaper, J. B. (1998) Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**:142-201.
- Nataro, J. P., Maher, K. O., Mackie, P., and Kaper, J. B. (1987) Characterization of plasmids encoding the adherence factor of enteropathogenic *Escherichia coli*. *Infect. Immun.* **55**:2370-2377.
- Nataro, J. P., Scaletsky, I. C. A., Kaper, J. B., Levine, M. M., and Trabulsi, L. R. (1985b) Plasmid mediated factors conferring diffuse and localized adherence of enteropathogenic *Escherichia coli*. *Infect. Immun.* **48**:378-383.
- Nataro, J. P., Yikang, D., Cookson, S., Cravioto, A., Savarino, S. J., Guers, L. D., Levine, M. M., and Tacket, C. O. (1995) Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *J. Infect. Dis.* **171**:465-468.
- Nataro, J. P., Yikang, D., Giron, J. A., Savarino, S. J., Kothary, M. H., and Hall, R. (1993) Aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli* requires two unlinked plasmid regions. *Infect. Immun.* **61**:1126-1131.
- Nataro, J. P., Yikang, D., Yingkang, D., Walker, K. (1994) AggR a transcriptional activator of aggregative adherence fimbria I expression in in enteroaggregative *Escherichia coli*. *J. Bacteriol.* **176**:4691-4699.
- Nau, C. D., and Koninsky, J. (1989) Evolutionary relationship between the TonB-dependent outer membrane transport proteins: nucleotide and amino acid sequences of the *Escherichia coli* colicin I receptor gene. *J. Bact.* **171**:1041-1047.

- Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990) *Physiology of the bacterial cell: a molecular approach*. Sinauer Associates Inc., Sunderland, Massachusetts, USA.
- Neilands, J. B. (1981a) Microbiol iron compounds. *Ann. Rev. Biochem.* **50**:715-731.
- Neilands, J. B. (1981b) Iron absorption and transport in microorganisms. *Ann. Rev. Nutr.* **1**:27-45.
- Neter, E. (1959) Enteritis due to enteropathogenic *Escherichia coli*: present day status and unsolved problems. *J. Pediatr.* **55**:223-239.
- Neter, E., Korns, R. F., and Trussell, R. E. (1953) Association of *Escherichia coli* O111 with two hospital outbreaks of epidemic diarrhea of the new born infant in New York state during 1947. *Pediatrics* **12**:377-383.
- Neter, E., and Shumway, C. N. (1950) *E. coli* serotype D433: occurrence in intestinal and respiratory tracts, cultural characteristics, pathogenicity, sensitivity to antibiotics. *Proc. Soc. Exp. Biol. Med.* **75**:504-507.
- Neter, E., Westphal, O., Luderitz, O., Gino, R. M., and Gorzynski, E. A. (1955) Demonstration of antibodies against enteropathogenic *Escherichia coli* in sera of children of various ages. *Pediatrics* **16**:801-808.
- Nicholson-Weller, A., and Wang, C. (1994) Structure and function of decay-accelerating factor CD55. *J. Lab. Clin. Med.* **123**:485-491.
- O'Brien, A. D., and Holmes, R. K. (1996) Protein toxins of *Escherichia coli* and *Salmonella*. 2788-2802. In Neidhardt, F. C., Curtis III, R., Ingraham, J. L., Lin, E. E. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schacchter, M., and Umbarger, H. E. (Ed). *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, D. C.
- O'Brien, A. D., Lively, T. A., Chen, M. E., Rothman, S. W., and Formal, S. B. (1983) *Escherichia coli* O157:H7 strains associated with hemorrhagic colitis in the United States produce a *Shigella dysenteriae* I (Shiga) like cytotoxin. *Lancet* **i**:702.
- O'Brien, A. D., Newland, J. W., Miller, S. F., Holmes, R. K., Smith, H. W., and Formal, S. B. (1984) Shiga-like toxin-converting phages from *Escherichia coli* strains that cause haemorrhagic colitis or infantile diarrhoea. *Science* **226**:694-696.
- O'Brien, A. D., Thompson, M. R., Cantey, J. R., and Formal, S. B. (1977) Production of a *Shigella dysenteriae*-like toxin by pathogenic *Escherichia coli*. B-103. In Abstracts of the 77th General Meeting of the American Society for Microbiology 1977. American Society for Microbiology. Washington D. C.
- O'Brien, I. G., and Gibson, F. (1970) The structure of enterochelin and related 2,3-dihydroxy-N-benzoyl-serine conjugates from *Escherichia coli*. *Biochim. Biophys. Acta.* **215**:393-402.
- Ochs, M., Angerer, A., Enz, S., and Braun, V. (1996) Surface signalling in transcriptional regulation of the ferric citrate transport system of *Escherichia coli*: Mutational analysis of the alternative sigma factor FecI supports its essential role in *fecI* transport gene transcription. *Mol. Gen. Genet.* **250**:455-465.
- Ochs, M., Veitinger, S., Kim, I., Welz, D., Angerer, A., and Braun, V. (1995) Regulation of citrate-dependent iron transport of *Escherichia coli* -FecR is required for transcription activation by FecI. *Mol. Microbiol.* **15**:119-132.

- O'Connor, C. D., and Timmis, K. N. (1987) Highly repressible expression system for cloning genes that specify potentially toxic proteins. *J. Bact.* **169**:4457-4462.
- Ogierman M. A., Zabihi S., Mourtzio, L., Manning, P. A. (1993) Genetic organization and sequence of the promoter-distal region of the *tcp* gene cluster of *Vibrio cholerae*. *Gene*. **126**:51-60.
- Old, D. C., Corneil, I., Gibson, L. F., Thomson, A. D., and Duguid, J. P. (1968) Fimbriation, pellicle formation and the amount of growth of salmonellas in broth. *J. Gen. Microbiol.* **51**:1-16.
- Ørskov, F., Ørskov, I., Rees, T. A., and Sahab, K. (1960) Two new *E. coli* O-antigens: O141 and O142 and two new coli K-antigens: K85 and K86. *Acta Pathologica et Microbiologica Scandinavica* **48**:48-50.
- Ørskov, F., Whittam, T. S., Cravioto, A., and Ørskov, I. (1990) Clonal relationships among classical enteropathogenic *Escherichia coli* (EPEC) belonging to different O groups. *J. Infect. Dis.* **162**:76-81.
- Oshima, T., Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., *et al.* (1996) A 718-kb DNA sequence of the *Escherichia coli* K-12 genome corresponding to the 12.7-28.0 min region on the linkage map. *DNA Res.* **3**:137-155.
- Oswald, E., Pohl, P., Jacquemin, E., Lintermans, P., van Muylem, K., O'Brien, A. D., and Mainil, A. (1994) Specific DNA probes to detect *Escherichia coli* stains producing cytotoxic necrotising factor type 1 or type 2. *J. Med. Microbiol.* **40**:1757-1760.
- Owen, P., Caffrey, P., and Josefsson, L.-G. (1987) Identification and partial characterisation of a novel bipartite protein antigen associated with the outer membrane of *Escherichia coli*. *J. Bact.* **169**:3770-3777.
- Owen, P., Meehan, M., de Loughry-Douherty, H., and Henderson, I. (1996) Phase-variable outer membrane proteins in *Escherichia coli*. *FEMS Immuno. Med. Microbiol.* **16**:63-76.
- Pallen, M. J., Dougan, G., and Frankel, G. (1997) Coiled-coil domains in proteins secreted by type III secretion systems. *Mol. Microbiol.* **25**:423-425.
- Parsot, C., and Sansonetti, P. J. (1996) Invasion and the pathogenesis of *Shigella* infections. *Curr. Top. Microbiol.* **209**:25-42.
- Paulozzi, L. J., Johnson, K. E., Kamahale, L. M., Clausen, C. R., Riley, L. W., and Helgerson, S. D. (1986) Diarrhea associated with adherent enteropathogenic *Escherichia coli* in an infant and toddler center, Seattle, Washington. *Pediatrics* **77**:296-300.
- Payne, A. M. M., and Cook, G. T. (1950) A specific serological type of *Bact. coli* found in infant homes in absence of epidemic diarrhoea. *Br. Med. J.* **2**:192-195.
- Peiffer, I., Servin, A. L., and Bernet-Camard, M-F. (1998) Piracy of decay-accelerating factor (CD55) signal transduction by the diffusely adhering strain *Escherichia coli* C1845 promotes cytoskeletal F-actin rearrangements in cultured human intestinal INT407 cells. *Infect. Immun.* **66**:4036-4042.
- Pelicic, V., Reyrat, J. M., and Gicquel, B. (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol. Microbiol.* **20**:919-925.
- Pelicic, V., Reyrat, J. M., and Gicquel, B. (1996) Positive selection of allelic exchange mutants in *Mycobacterium bovis* BCG. *FEMS Microbiol. Lett.* **144**:161-166.

- Perna, N. T., Mayhew, G. F., Posfai, G., Elliott, S., Donnenberg, M. S., Kaper, J. B., and Blattner, F. R. (1998) Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. *Infect. Immun.* **66**:3810-3817.
- Pettersson, J., Nordfelth, R., Dubinina, E., Bergman, T., Gustafsson, M., Magnusson, K. E., and Wolf-Watz, H. (1996) Modulation of virulence factor expression by pathogen target cell contact. *Science* **273**:1231-1233.
- Philpott, D. J., McKay, D. M., Sherman, P. M., and Perdue, M. H. (1996) Infection of T84 cells with enteropathogenic *Escherichia coli* alters barrier and transport functions. *Am. J. Physiol.* **33**:G634-G645.
- Picciotto, M., Cohn, J., Bertuzzi, G., Greengard, P., and Nairn, A. (1992) Phosphorylation of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **267**:12742-12752.
- Pierce, J. R., and Earhart, C. F. (1986) *Escherichia coli* K-12 envelope proteins specifically required for ferrienterobactin uptake. *J. Bact.* **166**:930-936.
- Pinto, M., Robine-Leon, S., Appay, M. D., Keding, M., Triadou, N., Dussaulx, E., Lacroix, B., Simon-Assmann, P., Haffen, K., Fogh, J., and Zweibaum, A. (1983) Enterocyte-like development and polarisation of the human colon carcinoma cell line Caco-2 in culture. *Biol. Cell.* **47**:323-330.
- Polotsky, Y. E., Dragunskaya, E. M., Seliverstova, V. G., Avdeeva, T. A., Chakhutinskaya, M. G., Keytyi, I., Vertenyi, A., Ralovich, B., Emody, L., Malovics, I., safonova, N. V., Snigirevskaya, E. S., and Karyangina, E. I. (1977) Pathogenic effect of enterotoxigenic *Escherichia coli* and *Escherichia coli* causing infantile diarrhoea. *Acta. Microbiol. Acad. Sci. Hung.* **24**:221-236.
- Posset, O., d'Enfert, C., Reyss, I. and Pugsley, A. P. (1992) Pullulanase secretion in *Escherichia coli* K-12 requires a cytoplasmic protein and a putative polytopic cytoplasmic membrane protein. *Mol. Microbiol.* **6**:95-105.
- Postle, K., and Good, R. F. (1983) DNA sequence of the *Escherichia coli tonB* gene. *Proc. Natl. Acad. Sci. USA* **80**:5235-5239.
- Pressler, U., Staudenmaier, H., Zimmerman, L., and Braun, V. (1988) Genetics of iron dicitrate transport system of *Escherichia coli*. *J. Bact.* **170**:2716-2724.
- Puente, J. L., Bieber, D., Ramer, S. W., Murray, W., and Schoolnik, G. K. (1996) The bundle-forming pili of enteropathogenic *Escherichia coli* - transcriptional regulation by environmental signals. *Mol. Microbiol.* **20**:87-100.
- Pugsley, A. P. (1993) The complete general secretory pathway in Gram-negative bacteria. *Microbiol. Revs.* **57**:50-108.
- Quirk, P. G., Guffanti, A. A., Clejan, S., Cheng, J., and Krulwich, T. A. (1994) Isolation of Tn917 insertional mutants of *Bacillus subtilis* that are resistant to the protonophore carbonyl cyanide *m*-chlorophenylhydrazone. *Biochim. Biophys. Acta.* **1186**:27-34.
- Rabinowitz, R. P., Lai, L. C., Jarvis, K., McDaniel, T. K., Kaper, J. B., Stone, K. D., and Donnenberg, M. S. (1996) Attaching and effacing of host cells by enteropathogenic *Escherichia coli* in the absence of detectable tyrosine kinase mediated signal transduction. *Microb. Pathog.* **21**:157-171.
- Rajakumar, K., Sasakawa, C., and Adler, B. (1997) Use of a novel approach, termed island probing, identifies the *Shigella flexneri she* pathogenicity island which encodes a

- homolog of the immunoglobulin A protease-like family of proteins. *Infect. Immun.* **65**:4606-4614.
- Rakin, A., Saken, E., Harmsen, D., and Heeseman, J. (1994) The pesticin receptor of *Yersinia enterocolitica*: a novel virulence factor with dual function. *Mol. Microbiol.* **13**:253-263.
- Ramer, W. S., Beiber, D., and Skoolnik, G. K. (1996) BfpB, an outer membrane lipoprotein required for the biogenesis of bundle-forming pili in enteropathogenic *Escherichia coli*. *J. Bact.* **178**:6555-6563.
- Ramos, J. L., Rojo, F., Zhou, L., and Timmis, K. N. (1990) A family of positive regulators related to the *Pseudomonas putida* TOL plasmid XylS and the *Escherichia coli* AraC activators. *Nucleic Acids Res.* **18**:2149-2152.
- Rasheed, J. K., Guzman-Versezco, L.-M., and Kupersztoch, Y. M. (1990) Two precursors of the heat-stable enterotoxin of *Escherichia coli*: evidence of extracellular processing. *Mol. Microbiol.* **4**:265-273.
- Rech, S., Deppenmeier, U., and Gunsalus, R. P. (1995) Regulation of molybdate transport operon, *modABCD*, of *Escherichia coli* in response to molybdate availability. *J. Bact.* **177**:1023-1029.
- Reyrat, J.-M., Pelicic, V., Giquel, B., and Rappuoli, R. (1998) Counters selectable markers: untapped tools for bacterial genetics and pathogenesis. *Infect. Immun.* **66**:4011-4017.
- Riley, L. W., Remis, R. S., Helgerson, S. D., McGee, H. B., Wells, J. G., Davis, B. R., Hebert, R. J., Olcott, E. S., Johnson, L. M., Hargrett, N. T., Blake, P. A., and Cohen, M. L. (1983) Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* **308**:681-685.
- Roberts, M., Wooldridge, K. G., Gavine, H., Kuswandi, S. I. and Williams, P. H. (1989) Inhibition of biological activities of the aerobactin receptor protein in rough strains of *Escherichia coli* by polyclonal antiserum raised against native protein. *J. Gen. Microbiol.* **135**:2387-2398.
- Robins-Browne, R. M. (1987) Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. *Reviews of Infectious Diseases* **9**:28-53.
- Robins-Browne, R. M., Levine, M. M., Rowe, B., and Gabriel, E. M. (1982) Failure to detect conventional enterotoxins in classical enteropathogenic (serotyped) *Escherichia coli* strains of proven pathogenicity. *Infect. Immun.* **38**:798-801.
- Robins-Browne, R. M., and Prpic, J. K. (1985) Effects of iron and desferrioxamine on infections with *Yersinia enterocolitica*. *Infect. Immun.* **47**:774-777.
- Robins-Browne, R. M., Still, C. S., Miliotis, M. D., Richardson, N. J., Koornhof, H. J., Frieman, I., Schoub, B. D., Lecatsas, G., and Hartman, E. (1980) Summer diarrhoea in African infants and children. *Arch. Dis. Child* **55**:923-928.
- Rogers, H. J. (1973) Iron-binding catechols and virulence in *Escherichia coli*. *Infect. Immun.* **7**:445-456.
- Rogers, H. J., Perkins, H. R., and Ward, J. B. (1980) *Microbiol cell walls and membranes*. Chapman and Hall, London.
- Rogers, H. J., Synge, C., Kimber, B., and Bayley, P. M. (1977) Production of enterochelin by *Escherichia coli* O111. *Biochim. Biophys. Acta.* **497**:548-557.
- Rosenshine, I., Donnenberg, M. S., Kaper, J. B., and Finlay, B. B. (1992) Signal transduction between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells: EPEC induces

- tyrosine phosphorylation of host cell proteins to initiate cytoskeletal rearrangement and bacterial uptake. *EMBO J* 11:3551-3560.
- Rosenshine, I., Ruschkowski, S., Stein, M., Reinscheid, D. J., Mills, S. D., and Finlay, B. B. (1996) A pathogenic bacterium triggers epithelial signals to form a functional bacterial receptor that mediates actin pseudopod formation. *EMBO J* 15:2613-24.
- Rothbaum, R., McAdams, A. J., Giannella, R., and Partin, J. C. (1982) A clinicopathologic study of enterocyte-adherent *Escherichia coli*: a cause of protracted diarrhea in infants. *Gastroenterology* 83:441-454.
- Rothbaum, R., Partin, J. C., Saalfield, K., and McAdams, A. J. (1983) An ultrastructural study of enteropathogenic *Escherichia coli* infection in human infants. *Ultrastruct. Pathol.* 4:291-304.
- Rowe, B. (1979) The role of *Escherichia coli* in gastroenteritis. *Clin. Gastroenterol.* 8:625-644.
- Rowe, B., Gross, R. J., Lindop, R., and Baird, R. B. (1974) A new *E. coli* O group O158 associated with an outbreak of infantile enteritis. *J. Clin. Pathol.* 27:832-833.
- Rozenboom, H. J., Dijkstra, B. W., Engel, H., and Keck, W. (1990) Crystallization of the soluble lytic transglycosylase from *Escherichia coli* K-12. *J. Mol. Biol.* 212:557-559.
- Rudd, K. E. (1998) Linkage map of *Escherichia coli* K-12, Edition 10: the physical map. *Microbiol. Mol. Biol. Rev.* 62:985-1019.
- Rush, B. (1789) An inquiry into the cause and cure of cholera infantum. *Medical inquiries and observations*. Prichard and Hall, Philadelphia.
- Russel, M. (1994) Phage assembly: a paradigm for bacterial virulence factor export? *Science* 265:612-614.
- Sack, R. B. (1980) Enterotoxigenic *Escherichia coli*: identification and characterisation. *J. Infect. Dis.* 142:279-286.
- Sakazakai, R., Tamura, K., and Saito, M. (1967) Enteropathogenic *Escherichia coli* associated with diarrhea in children and adults. *Jpn. J. Med. Sci. Biol.* 20:387-39.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: A laboratory manual. 2nd Edition*. Cold Spring Harbour Laboratory Press.
- Sandilands, J. E. (1909) The communication of diarrhoea from the sick to the healthy. *Proceedings of the Royal Society of Medicine* 3:95-121.
- Sansonetti, P. J. (1992) Molecular and cellular biology of *Shigella flexneri* invasiveness: from cell assay systems to shigellosis. *Curr. Top. Microbiol. Immunol.* 180:1-19.
- Savarino, S. J., Fasano, Robertson, D. C., and A., Levine, M. M. (1991) Enteroaggregative *Escherichia coli* elaborate a heat-stable enterotoxin demonstrable in an in vitro rabbit intestinal model. *J. Clin. Invest.* 87:1450-1455.
- Savarino, S. J., Fasano, A., Watson, J., Martin, B. M., Levine, M. M., Guandalini, S., and Guerry, P. (1993) Enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 represents another subfamily of *E. coli* heat-stable toxin. *Proc. Natl. Acad. Sci. USA* 90:3093-3097.
- Savarino, S. J., Fox, P., Yikang, D., and Nataro, J. P. (1994) Identification and characterization of a gene cluster mediating enteroaggregative *Escherichia coli* aggregative adherence fimbria I biogenesis. *J. Bacteriol.* 176:4949-1957.

- Savkovic, S. D., Koutsouris, A., and Hecht, G. (1996) Attachment of a noninvasive enteric pathogen, enteropathogenic *Escherichia coli*, to cultured human intestinal epithelial monolayers induces transmigration of neutrophils. *Infect. Immun.* **64**:4480-4487.
- Savkovic, S. D., Koutsouris, A., and Hecht, G. (1997) Activation of NF- κ B in intestinal epithelial cells by enteropathogenic *Escherichia coli*. *Am. J. Physiol.* **42**:C1160-C1167.
- Scaletsky, I. C. A., Silva, M. L. M., Toledo, M. R. F., Davis, B. R., and Trablusi, L. R. (1985) Correlation between adherence to HeLa cells and serogroups, serotypes and bioserotypes of *Escherichia coli*. *Infect. Immun.* **49**:528-532.
- Scaletsky, I. C. A., Silva, M. L. M., and Trablusi, L. R. (1984) Distinctive patterns of adherence of enteropathogenic *Escherichia coli* to hela cells. *Infect. Immun.* **45**:534-536.
- Schägger, H., and Von Jagow, G. (1987) Tricine-sodium dodecyl sulphate polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368-379.
- Schauer, D. B., and Falkow, S. (1993) Attaching and effacing locus of a *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infect. Immun.* **61**:2486-2492.
- Schnaitman, C. A. (1971) Solubilisation of the cytoplasmic membrane of *Escherichia coli* by Triton X-100. *J. Bact.* **108**:115-119.
- Schweizer, H. P. (1992) Allelic exchange in *Pseudomonas aeruginosa* using novel ColE1-type vectors and a family of cassettes containing a portable *oriT* and the counter-selectable *Bacillus subtilis sacB* marker. *Mol. Microbiol.* **6**:1195-1204.
- Schweizer, H. P., and Hoang, T. T. (1995) An improved system for gene replacement and *xylE* fusion analysis in *Pseudomonas aeruginosa*. *Gene* **158**:15-22.
- Sciortino, C. V., and Finkelstein, R. A. (1983) *Vibrio cholerae* expresses iron regulated outer membrane proteins in vivo. *Infect. Immun.* **42**:990-996.
- Scotland, S. M., Richmond, J. E., and Rowe, B. (1983a) Adhesion of enteropathogenic strains of *Escherichia coli* (EPEC) to HEP-2 cells is not dependent on the presence of fimbriae. *FEMS Microbiol. Lett.* **20**:191-195.
- Scotland, S. M., Smith, H. R., Willshaw, G. A., and Rowe, B. (1983b) Vero cytotoxin production in strain of *Escherichia coli* is determined by genes carried on a bacteriophage. *Lancet* **ii**:216.
- Sears, C. L., and Kaper, J. B. (1996) Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. *Microbiol. Rev.* **60**:167-215.
- Sedgwick, S. G., Robson, M., and Malik, F. (1988) Polymorphisms in the *umuDC* region of *Escherichia* species. *J. Bacteriol.* **170**:1610-1616.
- Sedgwick, S. G. and Morgan, B. A. (1994) *Meths. Molec. Genetics.* **3**:131-140.
- Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. H., and Whittam, T. S. (1986) Methods of multilocus electrophoresis for bacterial population genetics and systematics. *J. Appl. Environ. Microbiol.* **51**:873-884.
- Selbitschka, W., Niemann, S., and Puhler, A. (1993) Construction of gene replacement vectors for Gram-negative bacteria using a genetically modified *sacRB* gene as a positive selection marker. *Appl. Microbiol. Biotech.* **38**:615-618.

- Sellers, J. R. (1991) Regulation of cytoplasmic and smooth muscle myosin. *Curr. Opin. Cell Biol.* 3:98-104.
- Serény, B. (1957) Experimental keratoconjunctivitis shigellosa. *Acta. Microbiol. Acad. Sci. Hung.* 4:367-376.
- Shäffer, S., Hantke, K., and Braun, V. (1985) Nucleotide sequence of the iron-regulatory gene *fur*. *Mol. Gen. Genet.* 200:110-113.
- Shand, G. H., Anwar, H., Kaduruganuwa, J., Brown, M. R. W., Silverman, S. H., and Melling, J. (1985) In vivo evidence that bacteria in urinary tract infection grow under iron restricted conditions. *Infect Immun.* 48:35-39.
- Shanks, R. A., and Studzinski, L. P. (1952) *Bact. coli* in infantile diarrhoea. *Br. Med. J.* 2:119-123.
- Shaw, C. E., and Taylor, R. K. (1990) *Vibrio cholerae* O395 *tcpA* pilin gene sequence and comparison of predicted protein structural features to other type 4 pilins. *Infect. Immun.* 58:3042-3049.
- Shea, C. M. and McIntosh, M. A. (1991) Nucleotide sequence and genetic organization of the ferric enterobactin transport system: homology to other periplasmic binding protein-dependent systems in *Escherichia coli*. *Mol. Microbiol.* 5:1415-1428.
- Shenoy-Scaria, A. M., Kwong, J., Fijita, T., Olszowy, M. W., Shaw, A. S., and Lublin, D. M. (1992) Signal transduction through decay-accelerating factor. Interaction of glycosylphosphatidylinositol anchor and tyrosine-kinases p56^{lck} and p59^{fyn1}. *J. Immunol.* 149:3535-3541.
- Shibuya, K., Abe, T., and Fujita, T. (1992) Decay-accelerating factor functions as a signal transducing molecule for human monocytes. *J. Immunol.* 149:1758-1762.
- Simon, R., Priefer, U., and Puhler, A. (1983) A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Biotechnology* 1:784-791.
- Sixma, T. K., Kalk, K. H., van Zanten, A., Dauter, Z., Kingma, J., Witholt, B., and Hol, W. G. (1993) Refined structure of *Escherichia coli* heat-labile enterotoxin, a close relative of cholera toxin. *J. Mol. Biol.* 230:890-918.
- Skare, J. T., and Postle, K. (1991) Evidence for a TonB-dependent energy transduction complex in *Escherichia coli*. *Mol. Microbiol.* 5:2883-2890.
- Sleisenger, M., and Fordtran, J. (1993) *Gastrointestinal Disease: Pathophysiology, Diagnosis, Management*. W. B. Saunders, Philadelphia.
- Small, P. L. C., and Falkow, S. (1988) Identification of regions on a 230-kilobase plasmid from enteroinvasive *Escherichia coli* that are required for entry into Hep-2 cells. *Infect. Immun.* 56:225-229.
- Smellie, J. M. (1939) Infantile diarrhoea. *Lancet* 1:969-973.
- Smith, H. W., Green, P., and Parsell, Z. (1983) Vero cell toxins in *Escherichia coli* and related bacteria: transfer by phage and conjugation and toxic action in laboratory animals, chickens, and pigs. *J. Gen. Microbiol.* 129:3121-3137.
- Smith, J. (1955) The aetiology of epidemic infantile gastroenteritis. Edinburgh: Royal College of Physicians.

- Sohel, I., Puente, J. L., Murray, W. J., Vuopiovarkila, J., and Schoolnik, G. K. (1993) Cloning and characterization of the bundle forming pilin gene of enteropathogenic *Escherichia coli* and its distribution in *Salmonella* serotypes. *Mol. Microbiol.* **7**:563-575.
- Sohel, I., Puente, J. L., Ramer, S. W., Bieber, D., Wu, C. Y., and Schoolnik, G. K. (1996) Enteropathogenic *Escherichia coli* - identification of a gene cluster coding for bundle forming pilus morphogenesis. *J. Bact.* **178**:2613-2628.
- Sojka, W. J. (1965) *Escherichia coli* in domestic animals and poultry. Farnham Royal, Bucks, England. Commonwealth Agricultural Bureaux. 65-96.
- South, M. A. (1971) Enteropathogenic *Escherichia coli* disease: new developments and perspectives. *J. Pediatr.* **79**:1-11.
- Spitz, J., Yuhan, R., Koutsouris, A., Blatt, C., Alverdy, J., and Hecht, G. (1995) Enteropathogenic *Escherichia coli* adherence to intestinal epithelial monolayers diminishes barrier function. *Am. J. Physiol.* **268**:G374-G379.
- Staley, T. E., Jones, E. W., and Corley, L. D. (1969) Attachment and penetration of *Escherichia coli* into intestinal epithelium of the ileum of newborn pigs. *Am. J. Pathol.* **56**:371-392.
- Stallmeyer, M. J., Aizawa, S., Macnab, R. M., and DeRosier, D. J. (1989) Image reconstruction of the flagellar basal body of *Salmonella typhimurium*. *J. Mol. Biol.* **205**:519-528.
- Staudenmaier, H., van Hove, B., Yaraghi, Z., and Braun, V. (1989) Nucleotide sequences of the *fecBCDE* genes and locations of the proteins suggest a periplasmic-binding-protein-dependent transport mechanism for iron(III) dicitrate in *Escherichia coli*. *J. Bact.* **171**:2626-2633.
- Stein, M., Kenny, B., Stein, M. A., and Finlay, B. B. (1996a) Characterization of espc, a 110-kilodalton protein secreted by enteropathogenic *Escherichia coli* which is homologous to members of the immunoglobulin-a protease-like family of secreted proteins. *J. Bact.* **178**:6546-6554.
- Stein, M. A., Mathers, D. A., Yan, H., Baimbridge, K. G., and Finlay, B. B. (1996b) Enteropathogenic *Escherichia coli* markedly decreases the resting membrane potential of Caco-2 and HeLa human epithelial cells. *Infect. Immun.* **64**:4820-4825.
- Stevenson, J. S. (1952) Further observations on the occurrence of *Bact. coli* in adult faeces. *Br. Med. J.* **2**:123-124.
- Stibitz, S. (1994) Use of conditionally counterselectable suicide vectors for allelic exchange. *Methods Enzymol.* **235**:458-465.
- Stone, K. D., Zhang, H. Z., Carlson, L. K., and Sonnenberg, M. S. (1996) A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Mol. Microbiol.* **20**:325-337.
- Strauss, E. J. (1995) Bacterial pathogenesis. When a turn off is a turn on. *Curr. Biol.* **5**:706-709.
- Streatfield, S. J., Sandkvist, M., Sixma, T. K., Bagdasarian, M., Hol, W. G. J., and Hirst, T. R. (1992) Intramolecular interactions between the A and B subunits of heat-labile enterotoxin from *Escherichia coli* promote holotoxin assembly and stability in vivo. *Proc. Natl. Acad. Sci. USA* **89**:12140-12144.

- Strom, M. S., and Lory, S. (1993) Structure-function and biogenesis of the type IV pili. *Annu. Rev. Microbiol.* **47**:565-596.
- Strom, M. S., and Lory, S. (1992) Kinetics and sequence specificity of processing of prepilin by PilD, the type IV leader peptidase of *Pseudomonas aeruginosa*. *J. Bacteriol.* **174**:7345-7351.
- Struyve, M., Moons, M., and Tommassen, J. (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* **218**:141-148.
- Sussman, M. (1997) *Escherichia coli* and human disease. 3-48. In Sussman, M. (Ed) *Escherichia coli: Mechanisms of virulence*. Cambridge University Press, Cambridge.
- Swenson, D. L., Bukanov, N. O., Berg, D. E., and Welch, R. A. (1996) Two pathogenicity islands in uropathogenic *Escherichia coli* J96: cosmid cloning and sample sequencing. *Infect. Immun.* **64**:3736-3743.
- Tabor, S., and Richardson, C. C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074-1078.
- Tarr, P. I., and Neill, M. A. (1996) Perspective: The problem of non-O157:H7 Shiga toxin (verotoxigenic)-producing *Escherichia coli*. *J. Infect. Dis.*
- Tatum, E. L. (1945) X ray induced mutant strains of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **31**:215-219.
- Tatum, E. L. and Lederberg, E. M. (1947) Gene recombination in the bacteria *Escherichia coli*. *J. Bact.* **53**:673-684.
- Taylor, C. J., Hart, A., Batt, R. M., McDougall, C., and McLean, L. (1986) Ultrastructural and biochemical changes in human jejunal mucosa associated with enteropathogenic *Escherichia coli* (O111) infection. *J. Pediatr. Gastroenterol. Nutr.* **5**:70-73.
- Taylor, J., Powell, B. W., and Wright, J. (1949) Infantile diarrhoea and vomiting. A clinical and bacteriological investigation. *Br. Med. J.* **2**:117-125.
- Taylor, R. K., Manoil, C., and Mekalanos, J. J. (1989) Broad-host-range vectors for the delivery of *TnphoA*: use in the genetic analysis of secreted virulence determinants of *Vibrio cholerae*. *J. Bact.* **171**:1870-1878.
- Teneberg, S., Hirst, T. R., Ångström, J., and Karlson, K.-A. (1994) Comparison of the glycolipid-binding specificities of cholera toxin and porcine *Escherichia coli* heat-labile enterotoxin: identification of a receptor-active non-ganglioside glycolipid for the heat-labile toxin in infant rabbit small intestine. *Glycoconjugate J.* **11**:533-540.
- Thanabalu, T., Koronakis, E., Hughes, C., and Koronakis, V. (1998) Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner membrane translocase to an outer membrane exit pore. *EMBO* **17**:6487-6496.
- Thunnissen, A.-M. W. H., Dijkstra, A. J., Kalk, K. H., Rozenboom, H. J., Engel, H., Keck, W., and Dijkstra, B. W. (1994) Doughnut-shaped structure of a bacterial muramidase revealed by X-ray crystallography. *Nature*. **367**:750-753.
- Tinoco Jr., J., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., and Gralla, J. (1983) Improved estimation of secondary structure in ribonucleic acids. *Nature New Biol.* **133**:17-21.

- Tobe, T., Schoolnik, G. K., Sohel, I., Bustamante, V. H., and Puente, J. L. (1996) Cloning and characterization of *bfpTVW*, genes required for the transcriptional activation of *bfpA* in enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **21**:963-975.
- Toolan, H. W. (1954a) Transplantable human neoplasms maintained in cortisone-treated laboratory animals: H.S.#1; H.Ep.#2; H.Ep.#3; and H.Emb.Rh.#1. *Cancer Research* **14**:660-666.
- Toolan, H. W. (1954b) Growth of human embryonic tissues in cortisone-treated laboratory animals. *Proceedings of the Society of Experimental Biology and Medicine* **86**:607-612.
- Topley, W. W. C., and Wilson, G. S. (1936) *The principles of bacteriology and immunity*. Arnold, London. 1245-1254.
- Torres, A. G., and Payne, S. M. (1997) Haem iron transport system in enterohaemorrhagic *Escherichia coli* O157:H7. *Mol. Microbiol.* **23**:825-833.
- Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci. USA.* **76**:4350-4354.
- Tsolis, R. M., Bäumlér, A. J., Stojiljkovic, I., and Hefferon, F. (1995) Fur regulon of *Salmonella typhimurium* - identification of new iron regulated genes. *J. Bact.* **177**:4628-4637.
- Turbett, G. R., and Sellner, L. N. (1996) Digestion of PCR and RT-PCR products with restriction endonucleases without prior purification or precipitation. *Promega Notes* **60**:23-27.
- Tzipori, S., Montanaro, J., Robins-Browne, R. M., Vial, P., Gibson, R. and Levine, M. M. (1992) Studies with enteroaggregative *Escherichia coli* in the gnotobiotic piglet gastroenteritis model. *Infect. Immun.* **60**:5302-5306.
- Tzipori, S., Robins-Browne, R. M., Gonis, G., Hayes, J., Withers, M., and McCartney, E. (1985) Enteropathogenic *Escherichia coli* enteritis: evaluation of the gnotobiotic piglet as a model of human infection. *Gut* **26**:570-578.
- Ulshen, M. H., and Rollo, J. L. (1980) Pathogenesis of *Escherichia coli* gastroenteritis in man - another mechanism. *N. Engl. J. Med.* **302**:99-101
- Ursinus, A., and Holtje, J. V. (1994) Purification and properties of a membrane bound lytic transglycosylase from *Escherichia coli*. *J. Bact.* **176**:338-343.
- Vaandrager, A. B., Van der Wiel, E., Hom, M. L., Luthjens, L. H., and de Jonge, H. R. (1994) Heat-stable enterotoxin receptor/guanylate cyclase C is an oligomer consisting of functionally distinct subunits, which are non-covalently linked in the intestine. *J. Biol. Chem.* **269**:16409-16415.
- van de Putte, P., Plasterk, R., and Kuijpers, A. (1984) A *Mu gin* complementing function and an invertible DNA region in *Escherichia coli* K-12 are situated on the genetic element *e14*. *J. Bact.* **158**:517-522.
- van der Helm, D. (1998) The physical chemistry of bacterial outer-membrane siderophore receptor proteins. *Met. Ions Biol. Syst.* **35**:355-401.
- van Hove, B., Staudenmaier, H., and Braun, V. (1990) Novel 2-component transmembrane transcription control - regulation of iron-dicitrate transport in *Escherichia coli* K-12. *J. Bact.* **172**:6749-6758.

- Vasselon, T., Mounier, J., Hellio, R., and Sansonetti, P. J. (1992) Movement along actin filaments of the perijunctional area and de novo polymerization of cellular actin are required for *Shigella flexneri* colonization of epithelial Caco-2 cell monolayers. *Infect. Immun.* **60**:1031-1040.
- Venkatesan, M. M., Buysse, J. M., and Oaks, E. V. (1992) Surface presentation of *Shigella flexneri* invasion plasmid antigens require the products of the *spa* locus. *J. Bacteriol.* **174**:1990-2001.
- Vial, P. A., Robins-Browne, R. M., Loir, H., Prado, V., Kaper, J. B., Nataro, J. P., Maneval, D., Elsayed, A., and Levine, M. M. (1988) Characterisation of enteroadherent-aggregative *Escherichia coli*, a putative agent of diarrheal disease. *J. Infect. Dis.* **158**:70-79.
- Vipond, R., Bricknell, I. R., Durant, E., Bowden, T. J., Ellis, A. E., Smith, M., and MacIntyre, S. (1998) Defined deletion mutants demonstrate that the major secreted toxins are not essential for the virulence of *Aeromonas salmonicida*. *Infect Immun* **66**:1990-1998.
- von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids. Res.* **14**:4683-4690.
- Vuopio-Varkila, J., and Schoolnik, G. K. (1991) Localized adherence by enteropathogenic *Escherichia coli* is an inducible phenotype associated with the expression of new outer membrane proteins. *J. Exp. Med.* **174**:1167-1177.
- Wagegg, W., and Braun, V. (1981) ferric citrate transport in *Escherichia coli* requires the outer membrane receptor protein FecA. *J. Bact.* **145**:156-163.
- Wainwright, L. A., and Kaper, J. B. (1998) EspB and EspD require a specific chaperone for proper secretion from enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **27**:1247-1260.
- Wang, G., Angermüller, S., and Klipp, W. (1993) Characterisation of *Rhodobacter capsultus* genes encoding a molybdenum transport system and putative molybdenum-pterin-binding proteins. *J. Bact.* **175**:3031-3042.
- Wang, F. W., and Kushner, S. R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene.* **100**:195-199.
- Wang, Y., and Taylor, D. (1990) Chloramphenicol resistance in *Campylobacter coli*: nucleotide sequence, expression, and cloning vector construction. *Gene.* **94**:23-28.
- Warne, S. R., Varley, J. M., Boulnois, G. J., and Norton, M. G. (1990) Identification and characterization of a gene that controls colony morphology and auto-aggregation in *Escherichia coli* K12. *J. Gen. Microbiol.* **136**:455-62.
- Weinberg, E. D. (1978) Iron and infection. *Microbiol. Rev.* **42**:45-66.
- Weinberg, E. D. (1984) Iron withholding: a defense against infection and neoplasia. *Physiol. Rev.* **64**:65-102.
- Weinberg, E. D. (1989) Cellular regulation of iron assimilation. *Quart. Rev. Biol.* **64**:261-291.
- Weiss, B. (1976) Exonuclease II of *Escherichia coli* is exonuclease III. *J. Biol. Chem.* **251**:1896.
- Wentworth, F. H., Brock, D. W., Stulberg, C. S., and Page R. H. (1956) Clinical, bacteriological, and serological observations of two human volunteers following ingestion of *Escherichia coli* O127:B8. *Proc. Soc. Exp. Biol. Med.* **91**:586-588.

- Whitchurch, C. B., Hobbs, M., Livingston, S. P., Krishnapillai, V., and Mattick, J. S. (1991) Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. *Gene*. **101**:33-44.
- Whittam, T. S., Wolfe, M. L., Wachsmuth, I. K., Ørskov, F., Ørskov, I and Wilson, R. A. (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect. Immun.* **61**:1619-1629.
- Wieler, L. H., McDaniel, T. K., Whittam, T. S., and Kaper, J. B. (1997) Insertion site of the locus of enterocyte effacement in enteropathogenic and enterohemorrhagic *Escherichia coli* differs in relation to the clonal phylogeny of the strains. *FEMS Microbiol. Lett.* **156**:49-53.
- Williams, P. H. (1979) Novel iron uptake system specified by ColV plasmids: an important component in the virulence of invasive strains of *Escherichia coli*. *Infect. Immun.* **26**:925-932.
- Winkelmann, G. (1991) Specificity of iron transport in bacteria and fungi. 107-138. In Winkelmann, G. (Ed) *CRC Handbook of Microbiol Iron Chelates*. CRC Press Inc., Boca Raton, Fla.
- Wolff, C., Nisan, I., Hanski, E., Frankel, G., and Rosenshine, I. (1998) Protein translocation into host epithelial cells by infecting enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **28**:143-155.
- Wooldridge, K. G., Morrissey, J. A., and Williams, P. H. (1992) Transport of ferric-aerobactin into the periplasm of *Escherichia coli* K12: role of envelope-associated proteins and the effect of endogenous siderophores. *J. Gen. Microbiol.* **138**:597-603.
- Wooldridge, K. G., and Williams, P. H. (1993) Iron uptake mechanisms of pathogenic bacteria. *FEMS Microbiol. Rev.* **12**:325-348.
- World Health Organisation (1997) On-line World Health Report. URL www.who.org/whr/1996/exsume.htm
- World Health Organisation (1998) *The epidemiology and etiology of diarrhoea*. URL www.who.int/chd/pub/cdd/meded/1med.htm
- Wray, C., and Woodward, M. J. (1997) *Escherichia coli* infections in farm animals. 49-84. In Sussman, M. (Ed) *Escherichia coli: Mechanisms of virulence*. Cambridge University Press, Cambridge.
- Wren, B. W., Henderson, J., and Ketley, J. M. (1994) A PCR based strategy for the rapid construction of defined bacterial mutants. *BioTechniques* **16**:7-8.
- Wriedt, K., Angerer, A., and Braun, V. (1995) Transcriptional regulation from the cell surface - conformational changes in the transmembrane protein FecR lead to altered transcription of the ferric citrate transport genes in *Escherichia coli*. *J. Bact.* **177**:3320-3322.
- Write, A. C., Simpson, L. M., and Oliver, J. D. (1981) Role of iron in the pathogenesis of *Vibrio vulnificus* infections. *Infect. Immun.* **34**:503-507.
- Wu, S. S., and Kaiser, D. (1996) Markerless deletions of *pil* genes in *Myxococcus xanthus* generated by counterselection with the *Bacillus subtilis sacB* gene. *J. Bact.* **178**:5817-5821.

- Wu, S-X., and Peng, R-Q. (1992) Studies on an outbreak of neonatal diarrhea caused by EPEC O127:H6 with plasmid analysis restriction analysis and outer membrane protein determination. *Acta Paediatr. Scand.* **81**:217-221.
- Wyckoff, E. E., Duncan, D., Torres, A. G., Mills, M., Maase, K., and Payne, S. M. (1998) Structure of the *Shigella dysenteriae* haem transport locus and its phylogenetic distribution in enteric bacteria. *Mol. Microbiol.* **28**:1139-1152.
- Yahr, T. L., Goranson, J., and Frank, D. W. (1997) Exoenzyme S of *Pseudomonas aeruginosa* is secreted by a type III pathway. *Mol. Microbiol.* **22**:991-1003.
- Yamaguchi, K., Yu, F., and Inouye, I. (1988) A single amino acid determinant of the membrane localisation of lipoproteins in *E. coli*. *Cell.* **53**:423-432.
- Yancey, R. J., Breeding, S. A. L., and Klinkford, C. E. (1979) Enterochelin (enterobactin): virulence factor for *Salmonella typhimurium*. *Infect Immun.* **24**:174-180.
- Yannisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene.* **33**:103-119.
- Yoneyama, H., and Nakae, T. 1996 Protein C (OprC) of the outer membrane of *Pseudomonas aeruginosa* is a copper-regulated channel protein. *Microbiology.* **142**:2137-2144.
- Yu, J., and Kaper, J. B. (1992) Cloning and characterization of the *eae* gene of enterohemorrhagic *Escherichia coli* O157-H7. *Mol. Microbiol.* **6**:411-417.
- Yuhan, R., Koutsouris, A., Sankovic, S. D. and Hecht, G. (1997) Enteropathogenic *Escherichia coli*-induced myosin light chain phosphorylation alters intestinal epithelial permeability. *Gastroenterology.* **113**:1873-1882.
- Zhang, H. Z., and Donnenberg, M. S. (1996) DsbA is required for stability of the type IV pilin of enteropathogenic *Escherichia coli*. *Mol. Microbiol.* **21**:787-797.
- Zhang, H. Z., Lory, S., and Donnenberg, M. S. (1994) A plasmid-encoded prepilin peptidase gene from enteropathogenic *Escherichia coli*. *J. Bact.* **176**:6885-6891.
- Zhu, C., Harel, J., Jacques, M., Desautels, C., Donnenberg, M. S., Beaudry, M., and Fairbrother, J. M. (1994) Virulence properties and attaching-effacing activity of *Escherichia coli* O45 from swine postweaning diarrhea. *Infect. Immun.* **62**:4153-4159.
- Zimmermann, L., Hantke, K., and Braun, V. (1984) Exogenous induction of iron dicitrate transport system of *Escherichia coli* K-12. *J. Bact.* **159**:271-277.
- Zuber, F., Kotlarz, D., Rimsky, S., and Buc, H. (1984) Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol. Microbiol.* **12**:231-240.
- Zuker, M., and Stiegler, P. (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **9**:133-148.