Characterisation of the infection of *E. coli* by a Shiga-toxin encoding bacteriophage.

Thesis submitted in accordance with the requirements of the University of Liverpool for the Degree in Philosophy by Darren L. Smith

December 2005

Acknowledgements

I would like to thank Alan, Heather and Jon for their support and supervision throughout my Ph.D. I would also like to thank the honours and research students who contributed to my research, Ben McDermott for his help making the reporter gene construct John Connelly for his help with the reporter gene assays and Fiona McFarlane and Louise Ashall who all contributed to my research. I would also like to thank Caes Veltkamp, for his help with the Scanning Electron Microscopy.

I would mainly like to thank and acknowledge the support of Julie, who altered my opinion regarding entering scientific research and has always been a friend through the good times and the bad.

I would also like to thank my parents, grandma and family who still ask what area of biochemistry I work in! They have always been proud and supported me through things I have attempted over the years.

Thank you to all the people in 8.12 / Lab H for their support, friendship and binge drinking during this project including the old school Chloe, Steve, Chris, Rob, Paul, Coz, Claire, Jayne, Kate, Gro and Foggy, and the new school James, Jen, Mike, Sharon, Amy, Laura, Rachel, Alex, Brian and last but not least Mal for his help and over the last year of my PhD.

Live as if you were to die tomorrow. Learn as if you were to live forever. -- Gandhi

Abstract

The emergence of verocytotoxigenic or Shiga-like toxin producing *Escherichia coli* (VTEC / STEC) as food borne pathogens has become a worldwide public health concern. The most infamous serotype is O157, and it is the most virulent reported strain of enterohaemorrhagic *E. coli* (EHEC) to date. This severe infection can lead to haemorrhagic colitis (HC), haemolytic uremic syndrome (HUS) and thrombocytopenic purpura and can result in fatality. These downstream sequelae of the infection are the result of the action of Stx-toxins which have been found to be carried on lambdoid-like bacteriophages that have the ability to horizontally transfer toxin genes between *E. coli* strains. Bacteriophages are heterogeneous in nature and have been identified as almost modular genetic entities that evolve by recombinational events within the host. This leads to gene transfer and gene region swapping events over generations altering both phage and host genetic constitution resulting in a very heterogeneous population of lambdoid phages carrying Stx genes.

$\Phi 24_B$, a previously described Stx-encoding bacteriophage, was shown to unequivocally infect *E. coli* via an essential outer membrane protein (Vpr). Vpr is essential as it was not possible to interrupt function, and has been identified as being a key component in the biogenesis of the outer membrane. This protein is well conserved throughout the *Enterobacteriaceae* and has orthologs in a number of bacterial families. Vpr has been identified as being controlled by the bacterial stress response regulon (σ^E). A reporter gene system demonstrated increased transcription levels of Vpr when the culture was grown at 42 $^{\circ}$ C and under anaerobiosis, although levels of transcription were attenuated under growth in bile salts. Increasing Vpr expression also increases localisation at the cell surface which has been quantified by standardising a bacteriophage adsorption

assay. To completely categorise the protein-protein interaction between the bacteriophage and host antibodies were used to inhibit adsorption of $\Phi 24_B$ to Vpr (poly-clonal antibody made to recombinant Vpr). The ability of $\Phi 24_B$ to adsorb to its host can be acquired, as the *vpr* gene of *E.* coli was used to complement a resistant host (*Erwinia carotovora* sbsp. *atroseptica*). Surface localisation of Vpr has been demonstrated in the K-12 *E. coli* strain MC1061 using the anti-Vpr antibody by scanning laser confocal microscopy.

The tail of $\Phi 24_B$ has been shown to have an identical tail spike to Stx-phage 933W, the best studied Stx-phage; and they infect their host using the same bacterial receptor. A number of wild-type phages derived from environmentally isolated STEC strains were determined to have tail spikes similar to 933W using a PCR screening approach.

Induction of a $\Phi 24_B$ lysogen using norfloxacin has a dramatic effect on the morphology of the lysogen during the induction process, inhibiting the ability of cells to divide and yielding lysogen cells of up to 20 μm in length. Scanning electron microscopy revealed this morphological change and also the increasing presence of bacteriophage in the sample.

Lambdoid-like phages are thought to conform to the λ immunity model so that a prophage blocks subsequent infection by an identical bacteriophage. This has been disproved as part of this study as isogenic bacteriophages labelled with separate antibiotic resistance markers were shown to infect a single host genome.

i Table of Contents

Chapter 1: Introduction

Chapter 4: Identification and characterisation of Vpr

ii List of Figures

Chapter 1: Introduction

Chapter 2: Materials and Methods

Chapter 3 Visualisation of Induction of $\Phi 24_B::_\Delta Cat$ from an *E. coli* (MC1061) Host Using Scanning Electron Microscopy

Chaper 4 Identification and characterisation of Vpr

Chapter 5: Other bacteriophage encoded influences over infection of its host range

iii List of Tables

Chapter 1: Introduction

Chapter 2: Materials and Methods

Chapter 5 Bacteriophage encoded factors that influence infection and host range

Abbreviations

Φ	Bacteriophage
λ	Lambda
aa	amino acid
Amp	Ampicillin
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CAT	Chloramphenicol acetyl transferase
CDC	Centers for Disease Control
CNS	Central Nervous System
conc.	Concentration
CSPD	3-(4-methoxyspiro{1,2-dioxethane-3,2-(5'-chloro)tricycle[3.3.1.1^{3,7}]decan}-4-yl)
CTAB	Hexadecyltrimethylammonium Bromide
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
ds	Double Stranded
dUTP	Deoxy uracil triphosphate
EAEC	Enteroaggregative *Escherichia coli*
EDTA	Ethylene diamine tetra acetic acid
EHEC	Enterohaemorrhagic *Escherichia coli*
Ehx	Enterohaemolysin
EIEC	Enteroinvasive *Escherichia coli*

EPEC	Enteropathogenic *Escherichia coli*
Esp	*E. coli* secretory protein
FASTA	FAST-AII (nucleotide / amino acid sequence format)
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
Gb_3	Globotriaosyl ceramide
HC	Haemorrhagic colitis
HCT	Human colon adenocarcinoma cell line
Hep	Human epithelial cells
HIMEC	Human Intestinal microvascular Endothelial Cells
HUS	Haemolytic Uraemic Syndrome
IMS	Immuno-magnetic separation
Int	Integrase
IS	Insertion Sequence
Kan	Kanamycin
LA	Luria Bertani Agar
LB	Luria Bertani
LEE	Locus for enterocyte effacement
LPS	Lipopolysaccharide
MCS	Multiple cloning site
MIC	Minimum Inhibitory Concentration
MOI	Multiplicity of infection
mRNA	Messenger RNA
NCBI	National Centre for Biotechnology Information
NFLX	Norfloxacin
OM	Outer Membrane

OMP	Outer Membrane Protein
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
Pfx	Invitrogen Platinum DNA polymerase
Phage	Bacteriophage
PHLS	Public Health Laboratory Service
Rif	Rifampcin
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RTX	Repeats in toxin
sd	Sterile distilled
SDS	Sodium Dodecyl Sulphate
SE	Standard Error
SEM	Scanning Electron Microscope
SMAC	Sorbitol MaConkey
Sm	Streptomycin
ss	Single stranded
STEC	Shigatoxigenic *E. coli*
Stx	Shiga toxin
TAE	Tris-acetate EDTA
Taq	Recombinant DNA polymerase from *Thermus aquaticus*
TE	Tris EDTA
TEM	Transmission Electron Microscope
Tir	Translocated Intimin receptor

Tris	Tris(hydroxymethyl)methylamine
TTP	Thrombotic Thrombocytopenic purpura
UPEC	Uropathogenic *E. coli*
UV	Ultra-violet
Vpr	VT-phage receptor
VT	Verocytotoxin
VTEC	Verocytotoxigenic *E. coli*
WT	Wild type

Units

b	Bases
bp	Base pair
cfu	Colony forming units
g	Grams
h	Hours
kb	Kilobases
kDa	Kilodaltons
kV	Kilovolts
lfu	Lysogen Forming Units
M	Moles per Litre
mg	Milligrams
min	Minutes
ml	Millilitres
mM	Millimoles per litre
mm	Millimetre
MW	Molecular Weight
ng	Nanograms
nm	nanometres
°C	Degrees Centigrade
OD	Optical Density
rpm	Revolutions per minute

s	seconds
U	Units
µg	Micrograms
µM	Micromoles per litre
V	Voltage
v/v	Volume/volume
vol	Volume
w/v	Weight/Volume

Chapter 1: Introduction

1.1 Bacteriophages

Viruses that can infect bacteria are called bacteriophages or originally "bacteria eaters". Bacteriophages or phages were first identified by Twort (1915) and subsequently d'Herelle (1917), who described them as 'filterable, transmissible agents of bacterial lysis'. Elford and Andrews (1932) made the first estimation of viral particle size by a filtration method and determined that phage particles were between 25 and 100 nm in diameter. From his experiments published in 1934, Schlesinger described how phages were composed of mainly protein and nucleic acid. Lambda (λ) phage was first isolated and identified by Ester Lederberg (1951) and has become the model system for research on temperate phages *i.e.* bacterial viruses that can either enter the lytic or lysogenic life cycle. Hambly and Suttle's review (2005) postulates that there are approximately 10^{6} plaque forming units per ml (pfu ml^{-1}) on Earth and approximately 10^{27} phage particles in the sea.

Thus far, morphology has been the key factor in characterising and categorising phages. Taxonomically, viruses are categorised by their nucleic acid content (single or double stranded DNA or RNA), their morphology and bacterial host range. Fig 1.1 shows the morphological diversity of viruses that infect bacteria. Thus far, Shiga toxin (Stx)-encoding bacteriophages have fallen within one of three morphological groups: *Siphodoviridae* (H19B), *Podoviridae* (933W)(Plunkett *et al.*, 1999) and *Myoviridae* (P27) (Recktenwald *et al.*, 2002). These phage groups are commonly called tailed phages as all consist of both an icosahedral capsid and tail. The best described of these tailed phages are λ, Mu, P1, P2, T1, T5 and T7 (Casjens, 2005).

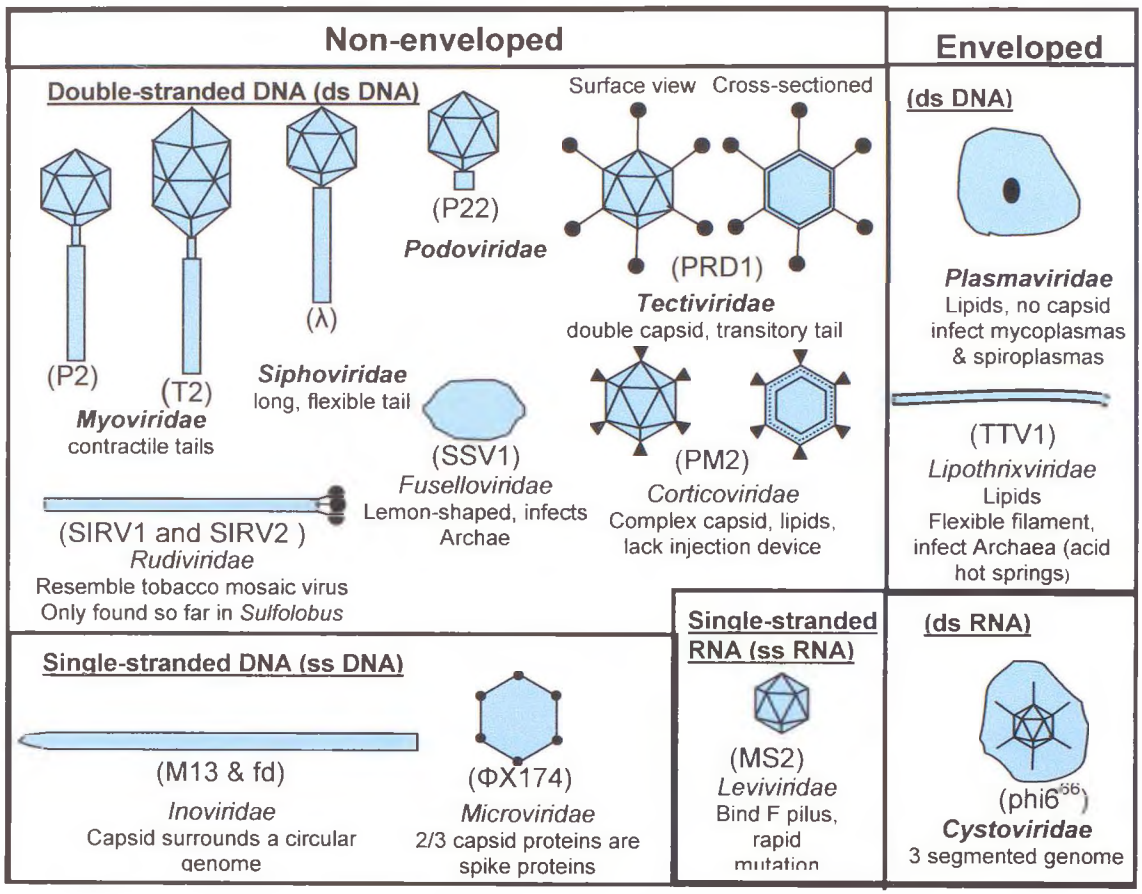


Fig 1.1 **Viruses that infect Bacteria.** From H. E. A. Allison (unpublished)

It has become apparent that characterisation of bacteriophages using phenotypic markers such as capsid and tail morphologies is a good approach to initially classify bacteriophages. Over the last 20 years, bacteriophages and bacterial sequencing projects have started to unravel the diverse genetic nature of bacteriophages and therefore the use of an exclusively morphological approach for classification can be flawed.

1.2 The Temperate Bacteriophage Life Cycle

Temperate bacteriophages can enter either the lysogenic or lytic pathway following infection of a host cell. Fig 1.2 details the lambda bacteriophage life cycle which is depicted as the model system for temperate lambdoid-like bacteriophages. The phage recognises its surface ligand on the bacteria and adsorbs. The linearised phage DNA is injected into the host bacterial cell through the tail of the lambdoid phage. A schematic of the linearised λ genome is presented in Fig 1.3. The phage genome re-circularises at the point of linearisation (*cos* site) and it is at this point that the decision between the lysogenic and lytic life cycles occurs. Lytic life cycle: Bacterial host RNA polymerase is subverted, binds four promoters *p*L, *p*RE, *p*O and *p*R (Herskowitz and Hagen, 1980) and begins to transcribe the early regulatory genes. Transcription from *p*L encodes N (anti-termination factor for genes involved in lysogeny) and the transcript from *p*RE encodes *cro* (Herskowitz and Hagen, 1980, Ptashne, 1992). Cro binds operator region O_R to prevent premature synthesis of cI by sequentially binding domains within the O_R (Herskowitz and Hagen, 1980). Production of O and P stimulates circularised phage DNA replication at point *ori*, and levels of Q anti-termination protein reach a level that has the ability to modify

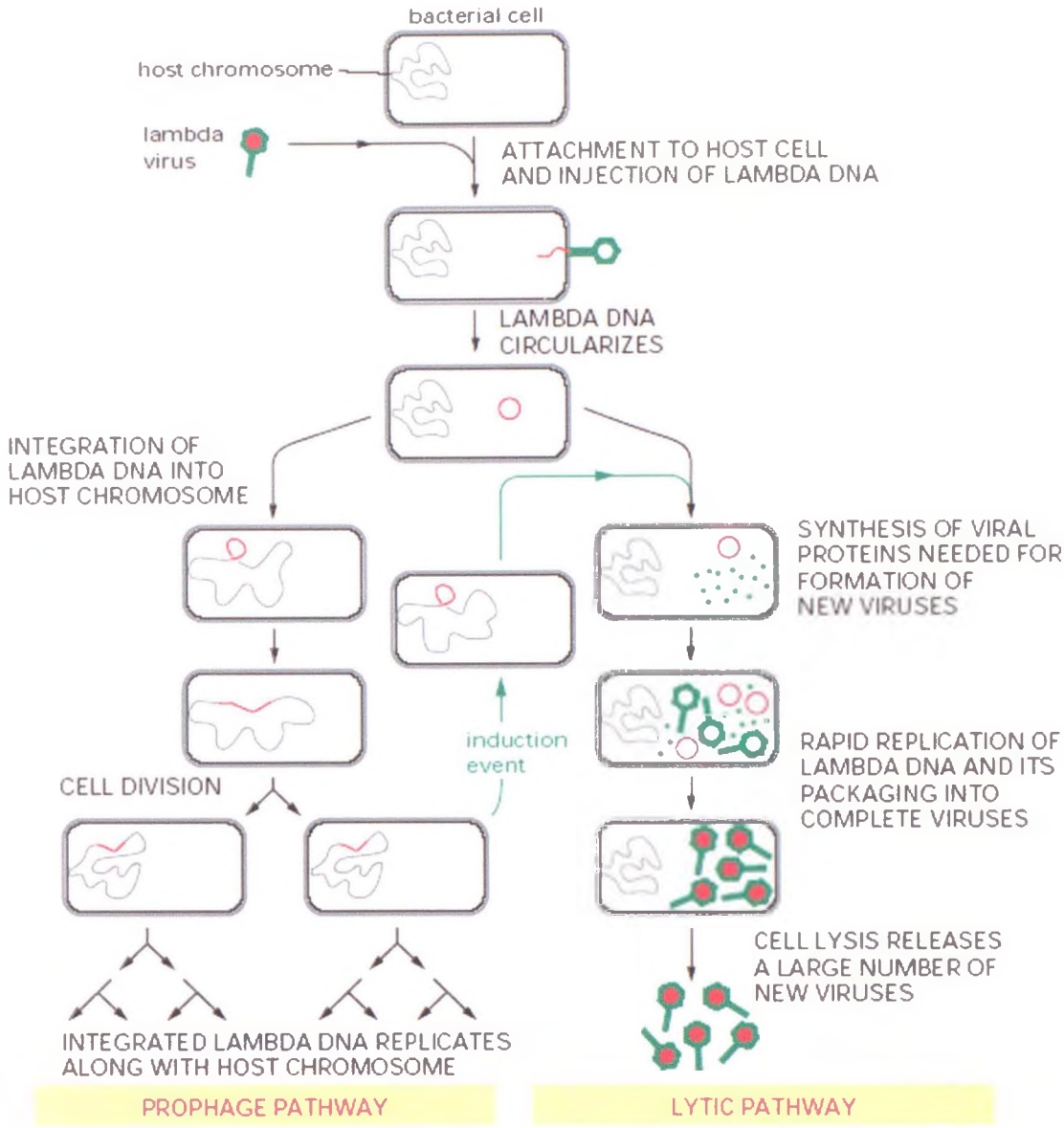


Fig 1. 2 Life cycle of Lambdoid-Like Viruses. Modified from Alberts *et al.*, 2003

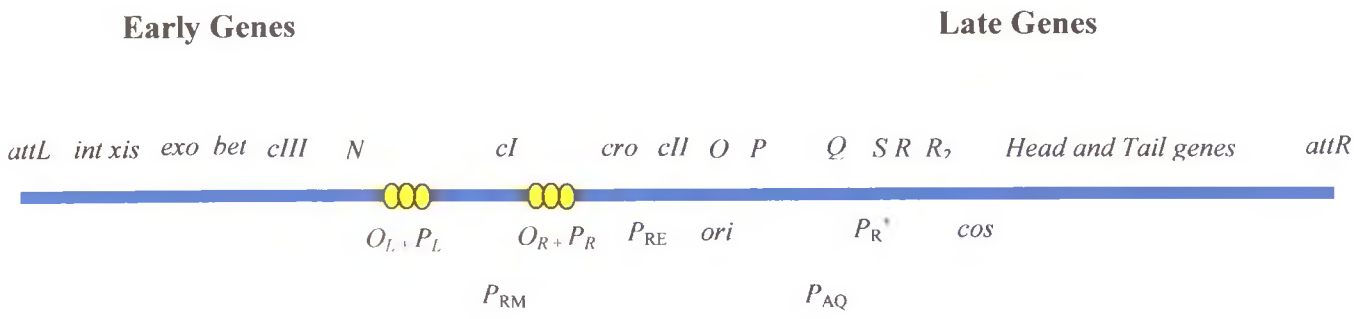


Figure 1.3 Genome orientation of the linearised λ genome. The genome has been linearised at the *att* (attachment sites) sites; this is the orientation in which the bacteriophage genome would be inserted into the bacterial genome. When the phage DNA is being packaged into the head, the circularised dsDNA is nicked at the end site (*cos*). The gene annotations are: *int* – integrase gene, *xis* – excisionase gene, *exo* and *bet* – involved in phage recombination; *cIII* – involved in the stabilisation of cII; *N* – involved in the regulation of the early genes; *cI* – λ repressor, inhibits lytic life cycle; *cro* – repressor, regulates the lytic life cycle; *cII* – regulator of λ repressor and integrase synthesis; *O* and *P* – phage DNA replication proteins; *Q* – involved in regulation of late genes; *S, R, R_2* – lysis proteins; head genes – involved in the production of the phage capsid; tail genes – cascade of eleven genes involved in the construction of the tail fibre and tail spike. Early promoters and operator regions: $O_L + P_L$ – operator region left and promoter left; $O_R + P_R$ – operator region right and promoter right; P_{RE} and P_{RM} – early promoters. Late gene promoters; P_{AQ}, P_R'.

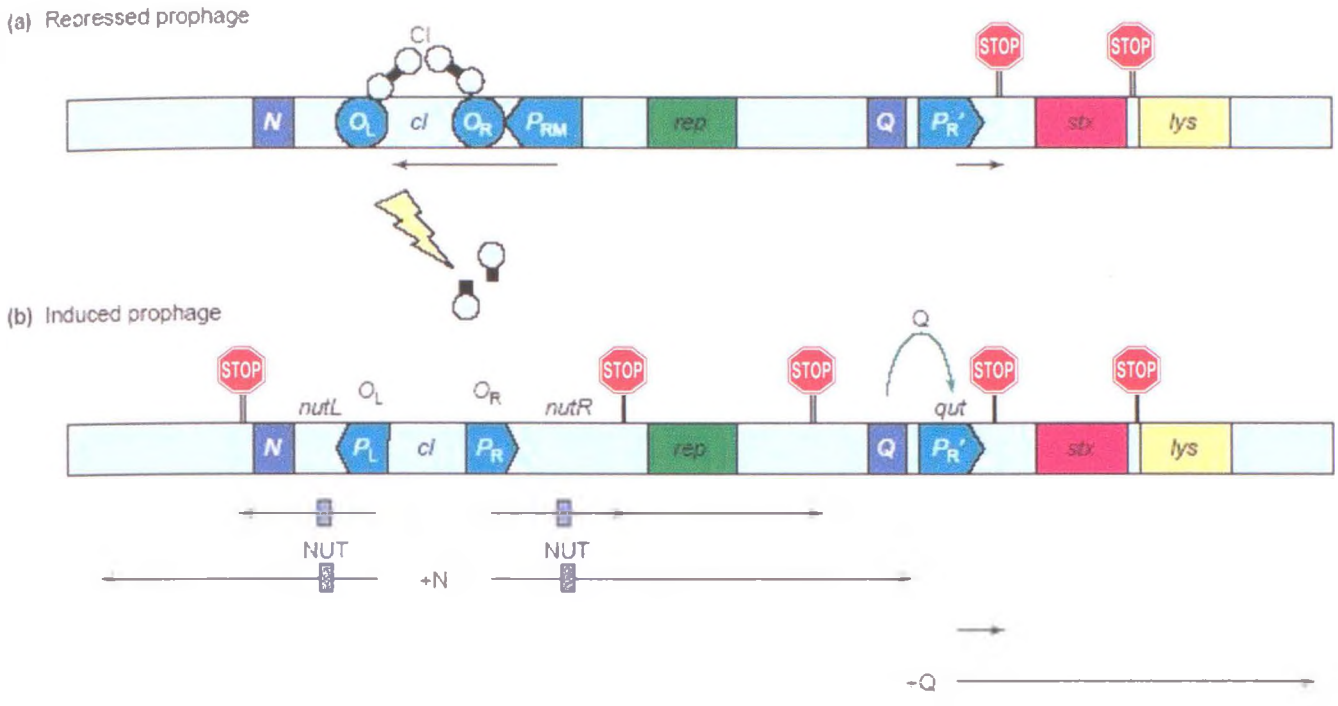
the host RNA polymerase at the Q utilisation site, the implication of which is that the downstream termination sequence is ignored and the late genes (including the capsid, tail etc) are transcribed (Herskowitz and Hagen, 1980). For lysogeny to occur upon infection, *p*L and *p*RE regulate the transcription of flanking regions including cII and cIII, necessary for the production of cI from *p*RE and integrase from *p*int. Integrase is needed for the integration of the bacteriophage into the bacterial genome at *att*B. On infection λ ensures that cI is not immediately produced. cI is transcribed from *p*RM, but this requires cI for positive regulation of itself. The presence of Cro after initial infection blocks cI transcription from *p*RM; cI is thus transcribed from *p*RE, which requires the action of cII and cIII. It is for these reasons that the decision between the lytic life cycle and lysogeny after infection is governed by cII and Cro (Michalowski *et al.*, 2005). Once cII has forced lysogenic conversion, cI takes over to regulate lysgoeny (Michalowski *et al.*, 2005).

In lysogenic conversion, cII activates cI from *p*RE giving a large burst of cI which inhibits *p*L and *p*R. cII is very unstable *in vivo* and is a substrate of the bacterial protease FtsH (Shortland *et al.*, 2000). cIII is also a substrate for FtsH (Shortland *et al.*, 1997) and probably protects cII by offering an alternative substrate for FtsH and therefore lowering the probability of cleavage. Some bacteriophages induce the lytic life cycle at higher frequency than others and this may in part be due to the relative levels of cII / cIII (Little, 2005). Therefore, if cIII levels are low, which can sometimes occur in a low multiplicity of infection (MOI), cII is more regularly cleaved. This cleavage means that no cI is produced and thus N, O, P and Q are expressed until Cro has reached a level that it saturates O_L and O_R, levels of Q increase until anti-termination occurs and expression of the late genes is initiated. If cIII is high expressed at high levels subsequent to infection, cII stimulates production

of cI and integrase. cI binding to O_L inhibits O, P and Q (Herskowitz and Hagen, 1980). Operator region binding of cI stimulates self transcription from *p*RM. cI in this situation continues to be made in the absence of cII and cIII (Herskowitz and Hagen, 1980).

Once lysogenised, bacteriophage can subsequently be induced into the lytic life cycle (Fig 1.3). Induction of the lytic life cycle in a lysogen has been linked to the *E. coli SOS* response, where RecA is able to negate factors involved in the repression of phage lysis (Neely & Friedman, 1998). In all lambdoid phages detailed thus far the cI protein is cleaved by RecA-stimulated auto-proteolysis occurring at an Ala-Gly dipeptide sequence. Lambdoid-like phages that encode shiga toxins are called Stx-phages, of which 933W is the most well characterised. Stx-phage 933W differs from any lambdoid phage observed as auto-cleavage occurs at a Leu-Gly dipeptide (Koudelka *et al.*, 2004). Livny and Freidman (2004) were able to show that Stx-phage lysogens (infected with H19B) are more likely to undergo induction to the lytic life cycle than λ and proposed that other phage-encoded genes are responsible for heightening the sensitivity of the bacterial response to induce lysis.

Some Stx-phages show homology to lambda in the nature of their genome orientation and regulation. The well described short tailed Stx-phage, 933W (Plunkett *et al.*, 1999), has been shown to have identical orientation of the complete lysis/lysogeny regulatory region to λ, apart from the third binding site of the left operator region (O_L3), which is located within the coding region of the cI gene (Tyler *et al.*, 2004; Koudelka *et al.*, 2004). Even though this means that there are only 2 active binding sites in O_L3, it is still sufficient to effectively repress *p*L transcription (Tyler *et al.*, 2004; Koudelka *et al.*, 2004). An Stx-phage, isolated from a clinical isolate of *E. coli* O157:H7 and named $\Phi 24_B$, is used in the experiments next



from Waldor and Freidman (2005)

Fig 1.4 Early Regulation of Stx-phage (a) Repressed prophage: transcription that initiates the regulatory cascade (shown in **(b)**) is blocked by CI repressor binding at the operators O_L and O_R. The lightning strike indicates stimulation of the bacterial SOS response to DNA damage. Activated RecA facilitates autocleavage of repressor. Transcription from P_{RM} directs synthesis of repressor in the lysogen. In the absence of Q, transcription initiating at P_R' terminates at the immediate downstream terminator. Cleavage of repressor results in release of repression. **(b) Induced prophage**: the regulatory cascade begins with transcription initiating at PL and PR that terminates after synthesis of a short message. This early transcription allows expression of N, which acts at the NUT sites in the RNA to modify RNA polymerase to a form that transcends terminators. Q, which is then expressed, acting at the qut site in the DNA, modifies transcription initiating at P_R' to a termination resistant form that can transcribe downstream genes that include *stx* and *lys* as well as most of the genes required for production of viable phages. Shown below the schematic, designated by arrows, are transcription patterns in the absence and presence of N and Q. Stop signs denotes sites or regions that contain transcription terminators.

described in this thesis, includes only a 1 bp difference from 933W in the region between O_L and *cro* (Sharon Gossage oral communication). Two isogenic forms of bacteriophage ($\Phi 24_B$) marked with kanamycin and chloramphenicol resistance genes were repeatedly shown to infect a single host, integrating into separate positions in the bacterial chromosome (Allison *et al.*, 2002), meaning that $\Phi 24_B$ does not conform to the λ immunity model where a second identical bacteriophage cannot lysogenise a single host.

The genetic switching used by lambdoid-like bacteriophages to regulate lysogeny or the productive lytic response is complex. New bacteriophages are being isolated and characterised where subtle differences are being identified in the way that this decision is regulated compared to λ model. This has already been described with 933W (Tyler *et al.*, 2004, Koudelka *et al.*, 2004). 933W is the first Stx-phage to have its immunity region described and characterised in some detail.

1.3 Integration of viable prophage

Lysis can occur when the MOI of the infection is low. There is also a bias towards lysogenic infection when the MOI is high enabling an increased frequency of lysogeny (Herskowitz and Hagen, 1980). What advantage does integration of a viable prophage benefit the bacteriophage and the host?

We are only beginning to understand the function of a wide range of phage-encoded proteins. Livny and Freidman (2004) described how H19B (Watarai *et al.*, 1998) lysogens show increased probability of induction compared to λ (section 1.2). This may be due to the increased sensitivity of the genetic switch between lysogeny and the lytic life cycle when compared to λ (Livny and Freidman, 2005). The benefit to the phage of this sensitive response and the early induction may be an increase in

the probability of propagation by out-competing other phage for the next host. This early induction results in increased pathogenicity of the bacteria, as the lytic life cycle stimulates the production of Stx-toxins.

Bacteriophage genes that have been identified as providing an advantage to the lysogen and are well described include the λ-encoded Lom protein, which increases bacterial resistance to macrophages (Barondess *et al.*, 1990). Lambda also encodes an outer membrane protein, Bor; which has been shown to increase the resistance of *E. coli* to animal serum (Barondess *et al.*, 1995). Genome sequencing projects have identified high levels of cryptic phages in all members of the *Enterobacteriaceae* sequenced thus far. Most of these bacteriophage regions are inactivated and are thus cryptic prophages unable to make viable virions. When a phage becomes unable to excise and propagate due to genetic deletion or mutation, this may provide an increased burden on the host to replicate and thus maintain this extra prophage DNA (Canchaya *et al.*, 2004). On the other hand, cryptic genes that increase viability or cell fitness will be subject to positive selection (Canchaya *et al.*, 2004). A good example was recently provided by Dziva *et al* (2004) who identified proteins of unknown function, encoded on cryptic prophages CP933X and CP933M of *E. coli* O157:H7 strain EDL933, that increase gut adhesion of the bacteria in a calf intestinal ileal loop model. Mann *et al.* (2003) described a marine T4-like phage (S-PM2) that carries D1 and D2 proteins associated with photosytem II, whose host is a photosynthetic cyanobacterium *Synechococcus*. These phage-encoded D1 and D2 proteins could help the bacteria to maintain photosynthetic function, possibly providing the energy required for the phage lytic life cycle to proceed (Bailey *et al.*, 2004). T4-phage (RB49) encodes a protein that has the ability to replace the host *E. coli* chaperonin protein GroES, which is involved in protein folding (Ang *et al.*,

2001; Keppel *et al.*, 2002). Another example of the host benefiting from remnant bacteriophage genes occurs in *Vibrio cholerae*. The Tcp pilus, which aids colonisation of the gut epithelia shows distinct homology to a bacteriophage-encoded gene (Karaolis *et al.*, 1999; Lee, 1999). Other bacterial functionality genes that have been postulated to be previously phage borne include the needle-like pilus proteins of the type III secretory systems that are used for gut colonisation by enteropathogenic *E. coli* and *Salmonella spp.* (Ehrbar *et al.*, 2005).

1.4 Bacteriophage genomics.

Over the last five years, bacterial genome sequencing has revealed a large amount of phage DNA studding bacterial host chromosomes. Blattner *et al.* (1997) published the sequence of the *E. coli* K-12 genome (~ 4.6 Mbp). Subsequently, in 2001, two sequences of *E. coli* O157:H7 strains were published: one (Sakai) from a large outbreak in Japan (Hayashi *et al.*, 2001), the other from a US outbreak (Perna *et al.*, 2001). Table 1.1 identifies the differences in genome size, G + C content, predicted ORFs and number of prophage or cryptic prophage related regions in the 4 sequenced *E. coli* strains. There is a higher number of phage-related regions in both of the Enterohaemorrhagic *E. coli* (EHEC) compared to *E. coli* K-12 (Table 1.1) and Sakai O157:H7 strain with approximately twice the number of cryptic phages present with respect to *E. coli* K-12. The difference in genome sizes between *E.* coli K-12 and both of the EHEC strains is ~ 1 Mbp, which is made up of both putative phage genes and insertion (IS) elements. From the genomes in Table 1.1, it has been identified that there is ~ 4 Mbp of colinear core chromosome that is separated by strain-specific regions and is responsible for the inter-strain genome size differences

Table 1.1 **Genomic sequence data comparison of sequenced O157:H7 EHEC strains (EDL933 and Sakai) compared to *E. coli* K-12 (MG1655) and a sequenced Uropathogenic *E. coli* (UPEC) strain CFT073.**

Strain	Pathotype	Size (bp)	G + C content (%)	No. of predicted ORFs	No. of prophage related regions	Ref
MG1655	K-12	4,639,221	50.8	4,294	10	Blattner *et al.*, 1997
CFT073	UPEC	5,231,428	50.5	5,553	5	Welch *et al.*, 2002
EDL933	EHEC O157:H7	5,528,445	50.5	5,361	16	Perna *et al.*, 2001
Sakai	EHEC O157:H7	5.594,477	50.5	5,361	24	Hayashi *et al.*, 2001

Modified from Dobrindt (2005).

Table 1.2 List of integrated bacteriophage regions in K-12 *E. coli* (Blattner *et al., 1997)* and both sequenced *E. coli* O157:H7 strains EDL933 (Perna *et al.*, 2001) and Sakai (Hayashi *et al.*, 2001).

***E. coli* strain and cryptic prophage**	**Length (bp)**	**Chromosomal insertion site**	**Characteristics**
E. coli K-12			
CP4-6	34,308	*thrW*	CP4 element
DLP12	21,302	*argU*	Lambdoid-like
e14	15,204	between *icdA* and *mcrA*	-
Rac	23,060	*ydaO*	Lambdoid-like
Qin	16,381	between b1543 and b1582	Lambdoid-like
CP4-44	12,873	Between b1955 and *yeeX*	CP4 element contains *flu*
CPZ-55	6,782	Between *eutB* and *eutA*	-
CP4-57	22,030	*ssrA*	*CP4* element
KpLE1	10,216	*argW*	Same integrase as Sp16
KpLE2	40,071	*leuX*	Partial integrase of SpLEE6
E. coli O157:H7			
EDL933	10,586	*thrW*	Lambdoid-like
CP933H	12,895	*thrW*	P4-like integrated in tandem with CP-933H
CP-933I			
CP-933K	38,588	between *ybhC* and *ybhB*	Lambdoid-like
CP-933M	45,244	*serT*	Lambdoid-like
BP-933W	61,663	*wrbA*	Viable Lambdoid-like, carries *stx2* genes
CP-933N	47,315	*potB*	Lambdoid-like, 2 tRNA loci
CP-933C	15,227	between *ycfD* and *phoQ*	Similar to Sp7 (Sakai)
CP-933X	54,000	between *icdA* and *minE*	Lambdoid-like
CP-933O	80,826	*yciD*	Lambdoid-like, 2 tRNA loci
CP-933R	49,797	*ydaO*	Lambdoid-like, 2 tRNA loci
CP-933P	57,984	b1582	Lambdoid-like, 3 tRNA loci
CP-933T	21,120	*leuZ*	P2 like phage
CP-933U	45,177	*serU*	Lambdoid-like, 3 tRNA loci
CP-933V	48,916	*yehV*	Lambdoid-like, *stx1* genes

CP-933Y	21,681	*ssrA*	Lambdoid-like
E. coli O157:H7 (Sakai)			
Sp1	10,586	*thrW*	Lambdoid-like
Sp2	12,887	*thrW*	P4-like integrated in tandem-Sp1
Sp3	38,586	between *ybhC* and *ybhB*	Lambdoid-like
Sp4	49,650	*serT*	Lambdoid-like, 3 tRNA loci
Sp5	62,708	*wrbA*	Lambdoid-like, stx genes, 3 tRNA loci
Sp6	48,423	*potB*	Lambdoid-like
Sp7	15,463	between *ycfD* and *phoQ*	-
Sp8	46,897	between *icdA* and *minE*	Lambdoid-like
Sp9	58,175	*yciD*	Lambdoid-like, 3 tRNA loci
Sp10	51,112	*ydaO*	Lambdoid-like, 3 tRNA loci
Sp12	45,778	between b1543 and b1582	Lambdoid-like, 3 tRNA loci
Sp12	46,142	between b1543 and b1582	Lambdoid-like, integrated in tandem with Sp11, 3 tRNA loci
Sp13	21,120	*leuZ*	P2-like
Sp14	44,029	*serU*	Lambdoid-like, 3 tRNA loci
Sp15	47,879	*yehV*	Lambdoid-like *stx1*
Sp16	8,551	*argW*	P22-like
Sp17	24,199	*ssrA*	Lambdoid-like
Sp18	38,759	between ECs4941 and ECs4999	Mu-like
SpLE1	86,249	*serX*	Cp4-like element
SpLE2	13,459	between b1955 and *yeeX*	Corresponds to CP4-44 of K-12
SpLE3	23,454	*pheV*	-
SpLE4	43,450	*selC*	CP4-like element, contains LEE
SpLE5	10,235	*leuX*	Integrated in tandem with SpLE5
SpLE6	34,148	*leuX*	

and genetic divergence with *E. coli* (Dobrindt, 2005). Approximately 40 % of the differences between the K-12 and O157:H7 strains are the located on the cryptic plus 1 viable in EDL933, 933W (Perna *et al.*, 2001) bacteriophage regions (Dobrindt, 2005). Table 1.2 lists all of the bacteriophage-related genes in *E. coli* K-12 and the 2 sequenced *E. coli* O157:H7 strains EDL 933 (Perna *et al.* 2001) and Sakai (Hayashi *et al.*, 2001). These data include the size of the putative phage regions and the points of integration in the bacterial core chromosome. The data also include phage-specific characteristics *e.g.* lambdoid-like, Mu –like, the number of tRNA genes they carry and whether the region is inserted in tandem with another bacteriophage region.

Cryptic bacteriophages are probably the result of mutation either through deletion in a functional gene necessary for the lytic life cycle or by mis-integration / excision of the phage genome on entry into or exit from the cell. These cryptic genes originally from viable bacteriophages, will be selected for in subsequent generations if they offer a selective advantage to the host. Alternatively, if the genes do not confer a disadvantage there is a possibility that these cryptic prophage genes will simply persist.

Bacteriophages, thus far, have been shown to have a distinct mosaic nature in the organisation and orientation of their genomes (Casjens, 2005). An example of this mosaicism can be observed in bacteriophages P22 and N15; both have some genes that are very similar to λ, but if their overall genome sequences were compared little homology would be registered (Hendrix *et* al. 2005; Casjens, 2003). The term 'novel sequence joints' has been invoked by Casjens (2005) to describe the points in the bacteriophage genome where similarities between two bacteriophages immediately stops, thus giving qualitative points of mocaicism. These novel

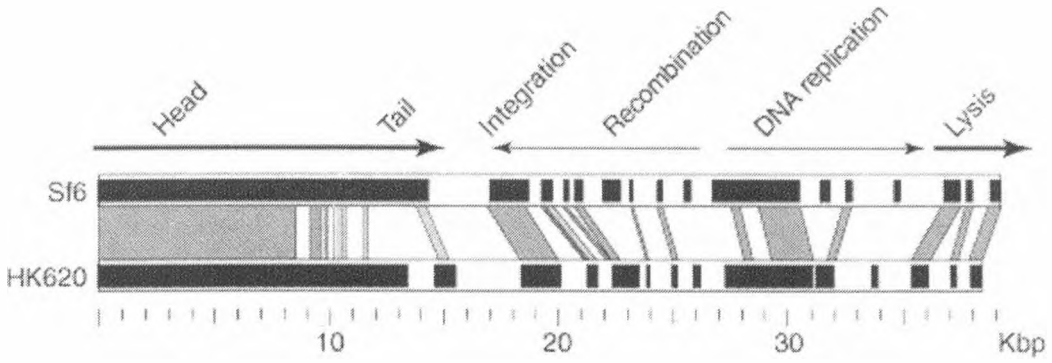


Fig 1.5 Qualitative mosaic relationship between the whole genomes of phage Sf6 and HK620, with their circular genomes opened at the 5′-end of their small terminase genes. Above, the arrows denote the major transcripts (thin arrows represent early operons; thick arrows represent the late operon), and representative functions of both genomes are provided. Black horizontal bars indicate regions of sequence similarity, and grey areas between the genomes indicate regions of >95% (dark), 90–95% (medium) and 80–90% (light) nucleotide sequence identity. Taken from Casjens, 2005.

sequence joints can be located in a gene, at the end of a complete gene or separating gene clusters. *Shigella flexneri* bacteriophage Sf6 has seven 50 – 150 bp motifs where gene swapping between bacteriophage could theoretically occur, although such gene shuffling has not actually been demonstrated (Casjens *et al.,* 2004). Fig 1.4 where the genomes of phages HK620 and Sf6 are compared and high sequence identity is increased in certain regions, but overall main functional genes are conserved between phage types *i.e.* the head genes are similar, divergence is first seen towards the tail of the phage *etc.* (Casjens, 2005).

Sequence identity between bacteriophages does not always complete the genetic story of bacteriophage genomics, as similarity can be compared at not only the gene level, but at the level of protein folding: these have been classified by Casjens (2005) as 'non-recognisable homologues' (NRH). An example is the capsid gene of P22 and HK97, they are divergent at the nucleotide sequence level, but have similarities in the protein folding regions (Jiang *et al.*, 2003; Helgstrand *et al.*, 2003).

Levels of recombination between both cryptic and viable phages could be extensive as Herold *et al.*, (2004) have shown that on norfloxacin induction of bacteriophage 933W from O157:H7 strain EDL933, identified up-regulation of 85 genes, of which 55 were related to Stx-phage 933W and CP-933V, was detected by micro-array analysis. Twenty-two of the remaining 33 up-regulated genes were related to other cryptic bacteriophages; the remaining 11 genes were associated with recombination and bacterial stress functions (Herold *et al.*, 2004). The *stxA* $_2$ gene was shown to be up-regulated 158-fold on induction with norfloxacin (Herold *et al.*, 2004). Thus Herold *et al.* (2004) data identifies the level of shiga-toxin up-regulation on induction which would concur with increased virulence. It also identifies how recombinational events could occur with the increased regulation of

all these phage genes leading to the evolution of the bacteriophage. There is also a possibility, as structural components of all these different bacteriophages are being made on induction, that chimeric phage incorporating structural proteins from some of the cryptic phages present join to form a viable virion that has the ability to infect a host previously inaccessible due to tail spike differences *etc*. This virion would contain the genome of the original viable bacteriophage.

An example of gene acquisition from other phages has been described in a number of bacteriophages. Wild type λ has been shown to have acquired side tail fibres associated with T4 (Hashemolhosseini *et al*., 1996). There is a distinct gene motif in P22 and a T7-like phage, SP6, in the polysaccharide-binding domains of these phages and demonstrates that of the >500 amino acid domain there is 58 % protein sequence identity (Dobbins *et al*., 2004; Scholl *et al*., 2002). Apart from these tail regions, both bacteriophages do not encode any other similar proteins (Scholl *et al*., 2002). This is best explained by these phages having shared the same gene pool at some point in their evolution, possibly in a single host cell (Brussow *et al*., 2004).

Stx-phage genomes are larger than these of λ (*e.g*. ~ 61,670 bp compared to ~ 48,502 bp respectively) and there are a many genes of unknown function (~ 74 %). These genes could contribute to the virulence of the phage and even the lysogenised bacteria themselves. The fully sequenced bacteriophages that have been heavily studied are still not fully annotated with respect to gene function. These genes may have less significance to the phage to central processes such as replication, but may play other roles in their survival and evolution. There are also gene products that may only affect the biology of the host cell *in vivo e.g. E. coli* in the gastrointestinal tract.

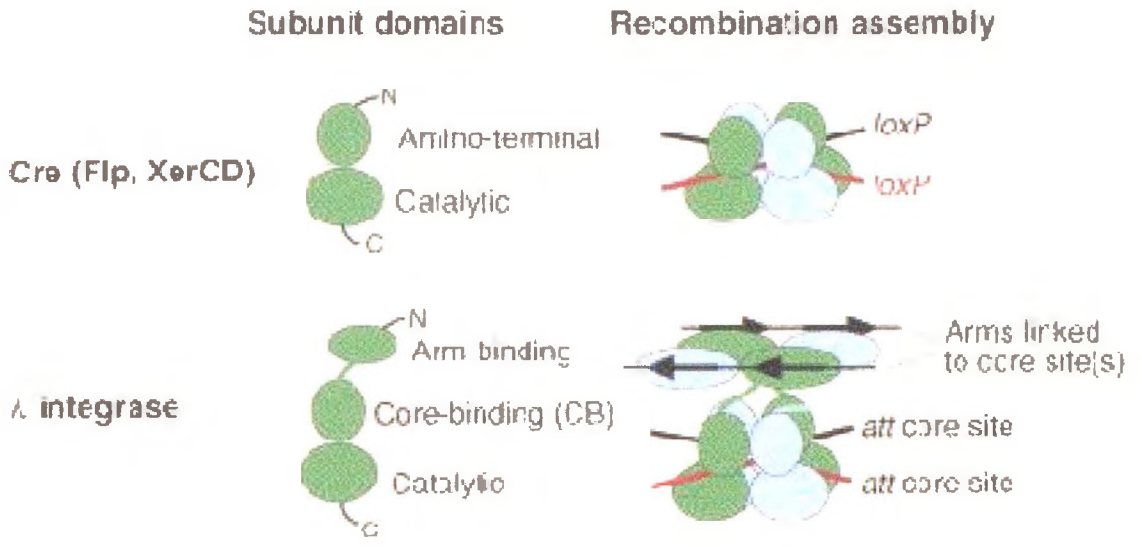


Fig 1.6 Comparison of λ integrase to a simpler version of recombinase found in the Cre system. The Cre complex when compared to the λ integrase shows the addition of the core binding site bound tetramer of the arm binding region. This diagram shows a schematic of the complex representing the crystal structures of the integrase gene identified by Biswas *et al.* (2005) (From Van Duyne, 2005).

1.5 Bacteriophage integration into the bacterial host

As discussed in section 1.2 bacteriophages, under the correct conditions, can lysogenise their bacterial hosts. Tyrosine recombinases are present in yeast (mu2 plasmid) and bacterial integrons, and are responsible for catalysing the integration and excision of mobile genetic elements (including bacteriophage) into their associated target site. The integration of λ has been modelled on the Cre and Flp recombinase sytems (Chen *et* al., 2000; Van Duyne, 2001) and has been invaluable in unravelling the complex integration of lambda into the *E. coli* genome (Van Duyne *et al.*, 2005). The λ integrase, starting from the N terminus, includes an arm-binding site, a core-binding site, and towards the C- terminus, a catalytic site depicted in Fig 1.4. This figure also shows schematic differences in the structures between the of the Cre / Flp recombinase complex compared to the λ integrase complex (Van Duyne, 2005). Through crystallography of integration and excision intermediates, it was possible to further categorise the integration/excision complex of λ (Biswas *et al.*, 2005). A tetramer (integrase complex per strand) of the core binding and catalytic site is described and is similar to the structure produced in the Cre and Flp system (Chen *et al.*, 2000; Azaro and Landy, 2002). The main difference is that the λ integrase has arm binding sites that bind regulatory DNA sequences flanking the regions where the strand exchange takes place (*att* sites) (Fig1.5) (Landy, 1989; Azaro and Landy, 2002). Integration Host Factor (IHF) plays the key role in bending the DNA from its recognition site so that flanking regions to the site of exchange are held at the arm-binding sites (Rice *et al.*, 1996).

Since the finer points of λ integration / excision are now being discovered, these observations can only be compared to other lambdoid-like bacteriophages

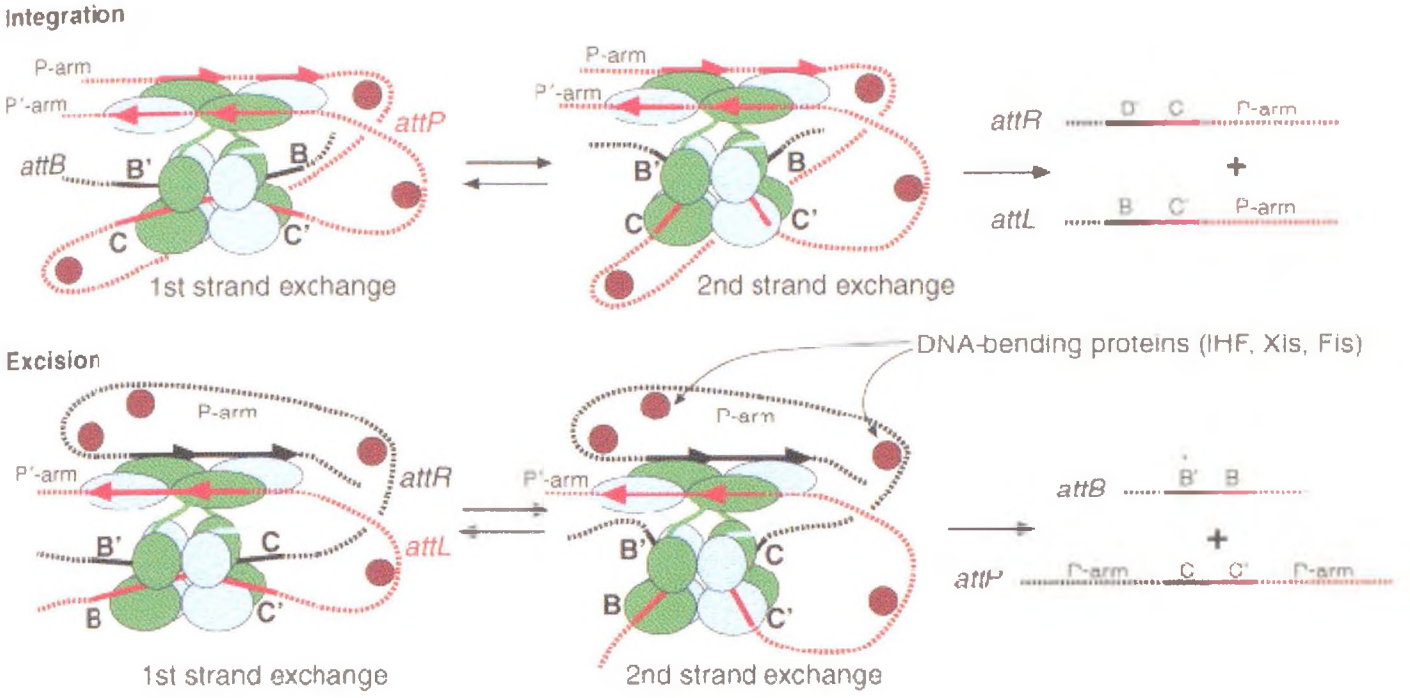


Fig 1.7 Integration and excision by the λ integrase. The diagram shows a schematic of the integration / excision pathways taken by the λ integrase from the crystallography of integration / excision intermediates by Biswas *et al.*, (2005). During integration the *attB* (bacterial integration point) and *attP* (phage site of cleavage and integration) sites are brought together within the core complex and the 1st exchange of strands occurs to form a Holliday junction intermediate. The 2nd strand exchange forms the recombinant *attL* (left end of the integrated phage chromosome) and *attR*. Excision is thought to work in the opposite orientation. Integration Host Factor (IHF), excisionase (Xis) and Fis are involved in bending the DNA to bind the arm-binding sites for the subsequent exchange of strands to occur. (From Van Duyne, 2005).

that are currently being studied. Differences may be seen with respect to the bacteriophage arm binding or other further nuances that separate phage types. Balding *et al.* (2005) showed the divergent nature of the integrase genes associated with lambdoid-like phages (including Stx-phages) outside of the 3 major integrase sites previously mentioned. Genetic divergence and distance between these sites may alter integration potential or define differences in integration. It has also been shown that bacteriophages can hijack bacterial tyrosine recombinases to integrate into their bacterial host. Virulent filamentous bacteriophage CTX, which carries the cholera toxin, uses the host tyrosine recombinase XerC/D in *Vibrio cholerae* to integrate into the bacterial chromosome as the phage lacks its own integrase gene (Mcleod and Waldor, 2004). Eventually, it may be possible to use the functional sites of the integrase to determine integration sites on the bacterial chromosome.

The ability of phages to integrate into their bacterial host along with all other recombinases are of scientific interest as these 'integration mediators' have the ability to promote integration into the host genome and therefore provide a means of introducing DNA to another host (transduction). There is also interest in phage recombinases for introducing DNA into mammalian genomes: it has been found that mammalian DNA can undergo integration of DNA using the P1 phage Cre system via recognition by utilising sequence information other than its native *loxP* site in the mammalian genome (Thyagarajan *et al.*, 2000). This has led to further research on the use of more efficient recombinase elements for mammalian chromosome integration including bacteriophages phiC31 (Thyagarajan *et al.*, 2001), R4 (Olivares *et al.*, 2001) and TP901-1 (Stoll *et al.*, 2002).

1.6 Phage Tail Fibres and infection

Bacteriophage tail fibres are responsible for adsorption of the bacteriophage to the bacterial cell via a specific epitope on the cell surface. *Siphoviridae* and *Podoviridae* phages are tailed, but differ in their tail morphology. *Siphoviridae* and *Podoviridae* tails are non-contractile compared to those of *Myoviridae* Fig 1.1. Tail shape is important, and length has been associated with efficiency of adsorption of the phage to its bacterial host ligand; the longer the tail, the greater the adsorption potential (Schwartz, 1976). Phage tail fibres can be either single gene products, as in some *Podoviridae* phages such as Stx-phage 933W, or produced by a cascade of genes *e.g.* long tailed phage such as λ (Xu *et al.*, 2004). SEM images of lambda, 933W and Mu are presented in Fig. 1.8. The λ tail is encoded by 11 genes *gpZ, gpU*, gpV, *gpG*, *gpT*, *gpH*, *gpM*, *gpL*, *gpK*, *gpI*, *gp,J*. The λ long tail fibre is approximately 140 nm in length and is constructed from 32 hexameric protein disks encoded by *gpV* (Casjens and Hendrix, 1974). The length of the tail fibre is determined by *gpH* which encodes a tail tape measure protein; this is a large ORF (~2 Kbp) that yields a α helical protein (Xu *et al.*, 2004). gpH is stored in the injection tube of the phage and is ejected into the naïve cell before the phage DNA is transferred (Roessner and Ihler, 1984). Tail gene order is highly conserved in the long tailed phages *i.e.* the major tail gene is found upstream of the tape measure gene (Xu *et al.*, 2004). The head and tail proteins of λ are constructed in different amounts due to the efficiency of translation (Sampson *et al.*, 1988); and a translational frame shift has been associated with controlling the levels of tail fibre proteins produced (Xu *et al.*, 2004). A frame shift (-1) in *gpG* and *gpT* gives a gpGT protein because of the read-through across both ORFs, and this has been

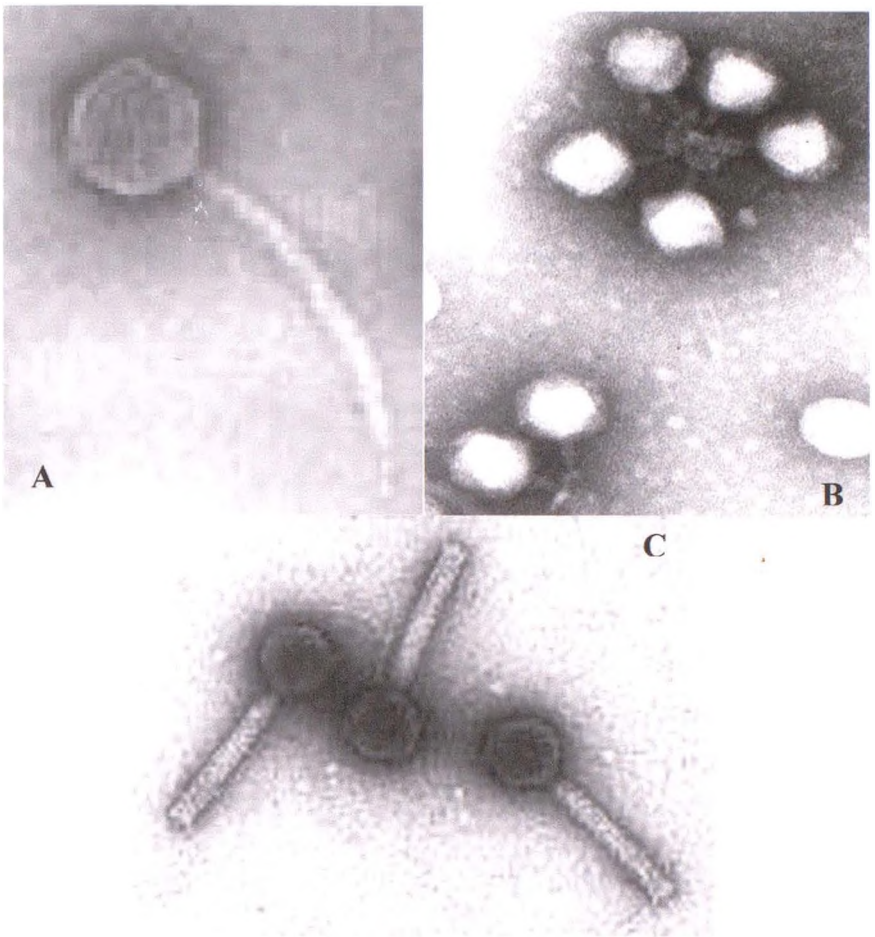


Fig 1.8 TEM images of wild type bacteriophage. **A + C** wild type λ (*Siphodoviridae*) and Mu (*Myoviridae*) respectively (from:www.biochem.wisc.edu/ inman/empics/virus.htm). **B** Image of Stx-phage 933W (*Podoviridae*) (from Plunkett *et al.*, 1999)

hypothesised as a possible chaperone protein for the tail fibre assembly (Xu *et al.*, 2004). This kind of frame shift regulation has also been described in the capsid genes of phage T4 where a frame shift (-1) has been identified in the scaffolding proteins (Casjens and Hendrix, 1988).

The key stage of phage adsorption is the association between, initially side tails of the bacteriophage to their respective receptors, and then the subsequent adsorption of tail spike to the bacterial receptor specific for infection. Phages obviously benefit from using a protein that is usually exposed to the environment during host cell growth and propagation. The λ *gpJ* gene has been associated with adsorption of the phage to the host cell outer membrane protein (OMP) as an anti-gpJ polyclonal antibody inhibited lambda infection (Wang *et al.*, 1998). The distal end of gpJ protein (last 248 amino-acid residues) is responsible for lambda adsorption to its bacterial ligand (Wang *et al.*, 2000). LamB, a maltose-binding protein, is the surface epitope in *E. coli* for lambda recognition. This was unequivocally established when an anti-LamB polyclonal antibody inhibited infection of the λ phage (Wang *et al.*, 2000). Environmental factors have been shown to inhibit lambda adsorption, such as bile salts. Under these conditions, the bacterial Ag43 gene (*E. coli* auto-transporter) would be repressed and it has been shown that this gene also has an effect on the bacteriophage adsorbing to its target protein (Gabig *et al.*, 2002). The *oxyR* gene (involved in the bacterial response under oxidative conditions) has a similar effect (Gabig *et al.*, 2002). Although the interaction between the λ tail spike and LamB is well characterised, little is known about other lambdoid-like phages and their bacterial surface targets. Dupont *et al.* (2004) identified the OMP responsible for *Lactococcus lactis* phage infection by using a

random insertion mutagenesis approach in the host bacteria to identify a trans-membrane protein involved in polysaccharide transport.

1.7 Capsid

The virus capsid has a symmetry which results from the binding of symmetrical protein shapes that, on completion, form a closed unit (Casjens, 1985). This is constructed from major capsid proteins that are arranged around a protein scaffold (Casper and Klug, 1962). The minor capsid gene encodes the head-tail connector joining the tail fibre and the capsid together, connecting the 5-fold symmetry of the capsid portal vertex with the six fold symmetry of the tail fibre (Casjens, 1985). Lurz *et al.* (2001) identified the head tail connector of phage SPP1 by EM by visualising phage heads that had not been packaged.

The procapsid, also known as the prehead or prohead, is the empty capsid that still contains the preliminary scaffolding, which in λ is digested by proteolytic cleavage and thus disappears from the protein head. In the case of phage such as P22, the scaffolding is cleaved and re-cycled for use in the further creation of capsid (King and Casjens, 1974). Lambda (*Siphoviridae)* has an extra capsid protein gpD that binds to the capsid surface several steps after the shell is complete (Sternbergt and Weisber, 1977). For the capsid to change from an almost spherical procapsid to its mature form requires a maturation step or expansion that is usually triggered by DNA packaging (Lee *et al.*, 2004) (see Fig 1.7). The capsid assembly in some bacteriophages as previously mentioned *e.g.* HKO22 has been well categorised, this is not the case for all lambdoid-like phages.

1.8 Phage DNA Packaging and ejection

Bacteriophage DNA is packaged into the viral head when the phage is being constructed. The capsid construction has been discussed above as a stimulus for packaging of the phage DNA into phage capsid, thus making a mature head. Terminases are responsible for cutting the circular phage genome at a specific site; in lambda this is the cohesive ends site (*cos*) and can be located upstream of the capsid genes (Fig 1.2). Terminases are involved in the packaging of the linearised phage DNA into the capsid (Black *et al.* 1989). The phage DNA enters the capsid via a portal protein, which is initially 'not passive' for DNA entry (Black, 1989). In T4 DNA packaging, a conformational change in the portal protein has been associated with the stimulus for the packaging event (Hsaio *et al.*, 1977). Phage DNA replication can lead to the production of concatamers that can be resolved by the terminases (Black, 1989). After the initial packaging cut of the DNA, it is packaged in a constant direction along the concatamer (Black, 1989). The terminal cutting in the packaging head by the terminase is called the headful cut; in λ this a sequence-specific cut at the *cos* site leaving a sticky ended 12 bp 5' overhang on the mature phage DNA (Feiss *et al.*, 1983; Hohn, 1983; Miwa *et al.*, 1983). Bacteriophage DNA, when being packaged into the head is thought to store sufficient energy from the packaging event that can be used to inject the DNA into the bacterial host. Evilevitch *et al.* (2003) showed that by increasing the external osmotic pressure by several tens of atmospheres outside of the capsid inhibited λ DNA injection. This raises the question how the phage DNA torsion effect is produced? A number of potential models have been published to address how phage DNA is held in the packaged virion. Electron microscopy of the packaged phage heads of T7 (Cerritelli *et* al., 1997), T4 (Olson *et al.*, 2001) and P22 (Zhang *et al.*, 2000) has shown that the

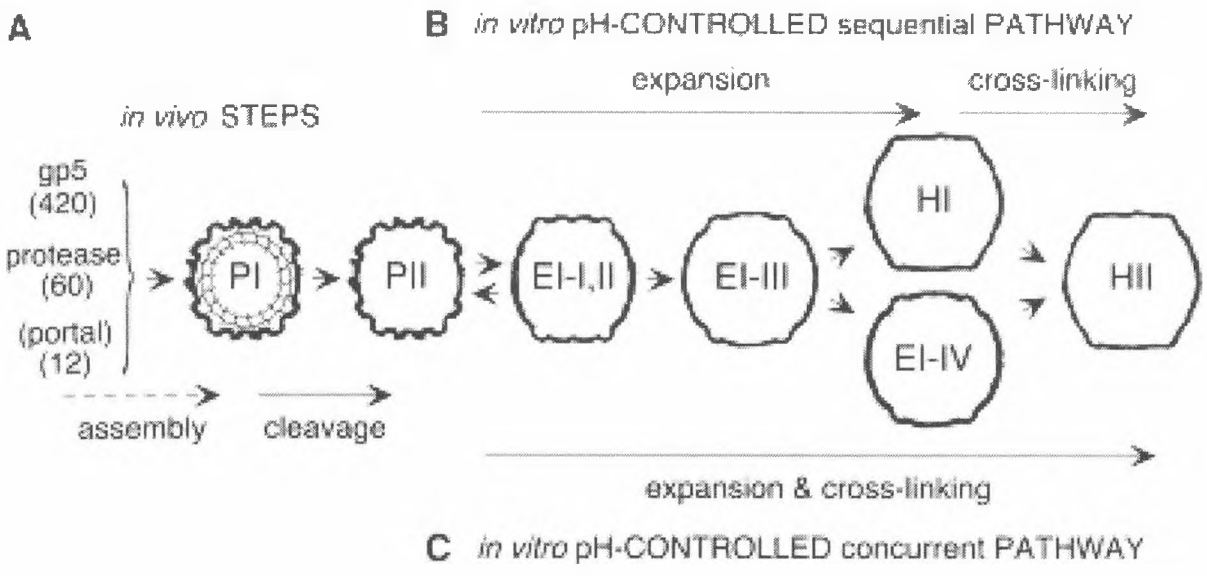


Fig 1.9 Diagram of bacteriophage HK97 capsid maturation / expansion (A) Assembly and processing. HK97 proheads assemble when capsid protein gp5 and protease gp4 are coexpressed in *E. coli*. Prohead I, a transient intermediate, consists of a shell formed from 60 hexamers and 12 pentamers of gp5 subunits with ~60 copies of gp4 inside. When present, a dodecamer of portal protein gp3 replaces one pentamer. The gp4 protease digests the N-terminal 102 residues of each gp5 and itself into fragments that exit from the shell to create Prohead II. Prohead II can be induced to expand *in vitro* at pH 4 (via expansion intermediates EI-I, EI-II, etc.) and further mature upon neutralization.(B) The *in-vitro* pH-controlled sequential pathway. In this older model expansion can be induced at low pH, but crosslinking does not begin until the entire structure has undergone the transition to Head I, identical to Head II, but without crosslinks. (C) The *in vitro* pH-controlled concurrent pathway. In the revised model of expansion, crosslinking begins early in the expansion process (from Lee *et al.*, 2004).

phage DNA takes on a spool-like conformation consistent with the coaxial spool model (Richards *et al.*, 1973; Earnshaw and Harrison, 1977). Spakowitz and Wang (2005) described an extra 12 ^{o} rotation for every 2 base-pairs that aids packaging of the phage head in Φ29. The model that they designed, including the extra twist by the packaging motor, is in agreement with the previous electron microscopy studies. The extra twist may also increase the stored pressure required for DNA ejection. It has been described that the genomes of λ (Virrankoski-Castrodeza *et al.*, 1982) and P22 (Casjens, 1989) show super-coiling; for this to occur the head would have to be packed in a spooling manner (Spakowitz and Wang, 2005).

1.9 The Pathogenicity of *Escherichia coli*

Escherichia coli are an important member of the intestinal microflora of humans and animals and typically become established in the gut of human infants within a few hours after birth, colonising the mucosal layer of the colon. *E. coli* as a commensal bacterium rarely causes symptoms of infection although there have been cases of peritonitis in the immuno-compromised patient when the gastrointestinal wall had been breached (Kaper *et. al.,* 2004). Certain *E. coli* types have, through evolution, obtained virulence factors that convert commensal strains to pathogenicity traits which form 6 pathotypes that clinically affect humans. These are enterohaemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC) and uropathogenic *E. coli* (UPEC) (Kaper, 2004). Infection by one of these pathotypes may yield one of three clinical syndromes; enteric / diarrhoeal disease, sepsis or urinary tract infection (UTI) (Kaper, 2004). All of these pathotypes mentioned, excluding EIEC, have been classed as ExPEC as they cause

extraintestinal infection. EIEC is an intracellular pathogen and can replicate both in epithelial cells and macrophages (Kaper, 2004). All pathotypes follow the sequential mode of infection that other mucosal pathogens follow in that they must first colonise, then evade the host defences, propagate and cause damage to the host (Kaper, 2004).

1.10 Shiga toxigenic *Escherichia coli* (STEC) Epidemiology

1.10a Overview of STEC infection - the emergence of a global pathogen.

The emergence of verocytotoxigenic or Shiga-like toxin-producing *Escherichia coli* (VTEC / STEC) as food-borne pathogens has become a worldwide public health concern. The most infamous serotype is O157 and the most virulent strain of Enterohaemorrhagic *E. coli* (EHEC) described to date (Law, 2000). STEC infection can also be classed as an EHEC infection as symptoms can lead to haemorrhagic colitis (HC) manifested as bloody diarrhoea. Haemolytic ureamic syndrome (HUS) and thrombocytopenic purpura are potential fatal complications that can also ensue (Dundas *et al.*, 2001) and it is these symptoms that are responsible for setting STEC apart from other pathogenic *E. coli* as a very serious pathogen indeed.

The first recorded incidence of *E. coli* O157 was in 1982 when the American Center for Disease Control (CDC) identified a rare *E. coli* serotype O157:H7 that had been isolated from contaminated hamburgers (Riley *et al.*, 1983). One year later a surveillance system was implemented across England and Wales to monitor the incidence of VTEC infection (Willshaw *et al.*, 2001). In the UK VTEC strains isolated from human infections are forwarded to the Health Protection Agency,

laboratories for enteric pathogens (LEP) and subsequent strain typing is reported to the CDC (Thomas *et al.*, 1993; 1996).

In 1992, 500 cases of *E. coli* O157 infection were reported from a single outbreak, subsequently traced to undercooked hamburgers being sold in a US fast food outlet (Bell, 1994). As a result of this outbreak the US government ordered surveillance of this infection via the CDC and have since reported that STEC infection accounts for at least 73,000 cases and 61 deaths per year including an estimated 2,100 hospitalisations per year in the U.S, this amounting to medical costs of between 250 and 300 million dollars per annum (CDC website 2005: http://www.cdc.gov/ncidod/dbmd/diseaseinfo/escherichiacoli_g.htm). In England and Wales, the incidence of VTEC infection has increased over the last 15 years; *e.g.* reached 1087 cases in 1997 and 1084 in 1999. The increase in incidence of STEC infection globally may be due to increased screening and surveillance programs alongside advances in laboratory identification methods. In Canada, an exponential increase of STEC infection was reported in the 1980's, which reached a peak of 8.8 people per 100,000 of the population in 1989 (Sharp, 1990). The incidence of human infection between 1982 and 1993 showed that there were 2 to 3 outbreaks of VTEC infection per year in the U.S. In England and Wales in the 1995 - 1999 period, thirteen incidences of STEC were related to contaminated dairy products, the largest of which occurred in North Cumbria where 87 people became infected after drinking contaminated milk (Anon, 1999). A similar dairy product-related STEC outbreak was also described in May 1997 in which geriatric patients in a Scottish hospital became unwell. The incidence of infection occurred across four separate, isolated hospital departments and was finally traced to a local dairy that produced cream cakes that had been brought to a hospital social event (O'Brien *et al.*, 2001).

Between 1994 and 1999, 8 outbreaks comprising 51 cases, 25 of which were children, were related to visits to petting and other farms that are open to visitors (Milne *et al.,* 1999).

On the 22nd November 1996 the largest outbreak ever reported in the UK occurred in Lanarkshire, Scotland; 512 cases were reported leading to 22 deaths, 17 specifically from the outbreak (Cowden 2001). In the same year the largest occurance of infection was described in Japan when 7000 people became infected from contaminated bean sprouts and 9 children died (Watanabe *et al.*, 1996). The most recent outbreak reported in the UK was in September 2005 in South Wales, deriving from a local meat supplier, 1 child died. Outbreaks of the infection are sporadic, but the low infective dose means that there is usually a large number of people involved in each episode.

1.10b Shiga-like toxin (Stx) / Vero toxin (VT)

The Stx toxins are carried by temperate bacteriophages (Stx-phages), the Stx-phage encoded toxins are called Shiga toxin 1 (Stx1), also known as VT1 and Shiga toxin-2 (Stx2) also known as VT2. In the Stx-phage genome, *stx1* and *stx2* are normally located directly after the Q antiterminator site (Datz *et al.,* 1996) (Fig 1.8). Stx2 is one thousand times more toxic to renal endothelial cells than Stx1 (Louise and Obrig, 1995).

Stx are members of the AB holotoxin family in that they are constructed of a single A-subunit (32 kDa) surrounded by a pentameric ring of five identical B-subunits (Melton-Celsa *et al.*, 1998). The B-subunit is involved in binding the toxin to its specific target site and shows affinity for globotriaosyl ceramide (Gb_3) and

globotetraosyl ceramide (Gb_4) (O'Loughlin *et al.,* 2001). Stx show activity against different cells/organs depending on the distribution of the target Gb_3 / Gb_4 receptors on the Vero cell surface (African green monkey kidney cells) (O'Loughlin *et al.,* 2001). Richardson *et al* (1992) injected a rabbit with Stx1 toxin and found inflammation in the gastrointestinal tract and central nervous system. In contrast, when Stx2 was injected into a mouse it exhibited central nervous system and renal disease without involvement of the gut (Tesh *et al.*, 1991). Stx1 has the ability to up-regulate apoptosis of endothelial cells by the inhibition of MC1-1, an apoptotic protein (Erwent *et al.*, 2003).

Vero-toxins enter their target cells at the apical membrane by receptor-mediated endocytosis and are transported across the cell to exert an effect on protein synthesis (Sandvig *et al.,* 1996). The toxin subunit (Stx_2B) attaches to its glycoprotein receptor and is internalised via a clathrin-coated pit (endosome). At this point catalytic cleavage of the Stx2A sub unit occurs by furin in the endosome and at the *trans*-Golgi network, increasing its enzymatic activity (Sandvig *et al.,* 1996). The toxin inhibits protein synthesis by restricting the elongation of peptides (Obrig *et al.*, 1987). Acheson *et al.* (1998) describe how purified Stx1 and Stx2 were added to a renal CaCo2A cell layer with no effect; when STEC bacteria were added, the toxins were transported across the layer destroying it. This response indicates that toxicity may be a culmination of virulence factors that are involved in directing the pathogenicity of the toxins (Law, 2000).

Stx1 and Stx2 cause an inflammatory response when administered to human monocytes stimulating production of TNFα, IL-1β, IL-6 and IL-8 and very low production of IL-10; this cytokine is stimulated by the presence of lipo-polysaccharide (LPS) (Nakagawa *et al.*, 2003). High ratios of inflammatory

cytokines IL-6 / IL-10 or IL-8 / IL-10 are used clinically as indicators of risk for HUS development (Westerholt *et al.*, 2000). Menge (2001) hypothesises that Stx1 can modulate intestinal inflammation as it binds to the CD77 epitope inhibiting proliferation of host lymphocytes. It may also confer an advantage to the bacteria to help explain why STEC can survive for long periods of time in the gut (Menge *et al.*, 2004).

Transportation of Stx2 to target organs may occur in humans via binding to a human serum amyloid P component (HuSap). This has been shown to compete for binding of Stx2 with a humanised monoclonal antibody TMA-15 that was shown to lower the efficacy of Stx2 in a mouse model. HuSap is present in the blood at ~30 – 45 μg ml^{-1} (Kimura *et al.*, 2003). O'Brien *et al.* (1983) compared toxins produced by EHEC and *Shigella dysenteriae* type 1 and found that they showed distinct amino acid sequence similarity. Stx1 has been shown to be homologous to shiga toxin from *Shigella dysenteriae* with only 1 amino acid difference (Calderwood *et al.*, 1987). The variation between Stx1 and Stx2 is usually identified in the B-subunit which alters binding specificities and toxicity (Melton-Celsa *et al.*, 1998) There are now a number of identified variants of the more virulent Stx2 catalogued from Stx sub-variant $Stx2_a$ up to $Stx2_g$ (Leung *et al.*, 2003). This variation can modulate specificity for target receptors *e.g.* Stx2e has an increased affinity for porcine erythrocytes (Matise *et al.*, 2003). Variation between Stx1 and Stx2 is also significant, as Willshaw (1987) described how specific probes designed for *stx1* failed to bind to *stx2* over a 1.4 Kb region.

1.10c Regulation of *stx1* and *stx2* Transcription

As discussed previously *stx* genes are located in the late gene region of temperate Stx-phages, and from sequence data it is established that they lie in the same region of the phage genome, just downstream of the Q antiterminator site (Fig 1.8). This was confirmed by Unkmeier (2000) while screening 49 STEC strains for Stx variation and sequence identity amongst the toxin gene flanking regions. As both *stx1* and *stx2* are situated in the prophage or cryptic prophage late genes, they are transcribed when the lytic life cycle is induced. Further detail on stimulation of Stx production can be found in section 1.2 and Fig 1.3. Fig 1.8 shows an alternative to the well described regulation of Stx provided in Fig. 1.3 where transcriptional control of Stx1 can occur under low environmental levels of iron (Neely and Friedman, 1998). Gamage *et* al. (2003) have shown that the presence of other non-pathogenic gut commensals can lead to the upregulation of Stx production by 1000-fold; this is due to the propagation of the phage in the commensal *E. coli* population. Wagner *et al.* (1999) showed that levels of toxin production can vary upon induction of different lysogens when different bacterial hosts are infected with a single Stx-phage.

1.10d STEC pathogenicity factors

The pathogenicity of STEC is not a single factor; it is an accumulation of a range of pathogenicity determinants that combine to give the complete pathogenic form. For the STEC to deliver the toxins effectively there must be an intimate association of the bacteria to the gut mucosa. Adhesion and other pathogenicity factors will be discussed in greater detail in this section. When STEC infection is mentioned it is pertinent to directly think of the O157:H7 serotype, but increased detection and characterisation of outbreaks have determined non-O157:H7 serotypes that have been associated with STEC infection *e.g.* Bettelheim (2000) describes the

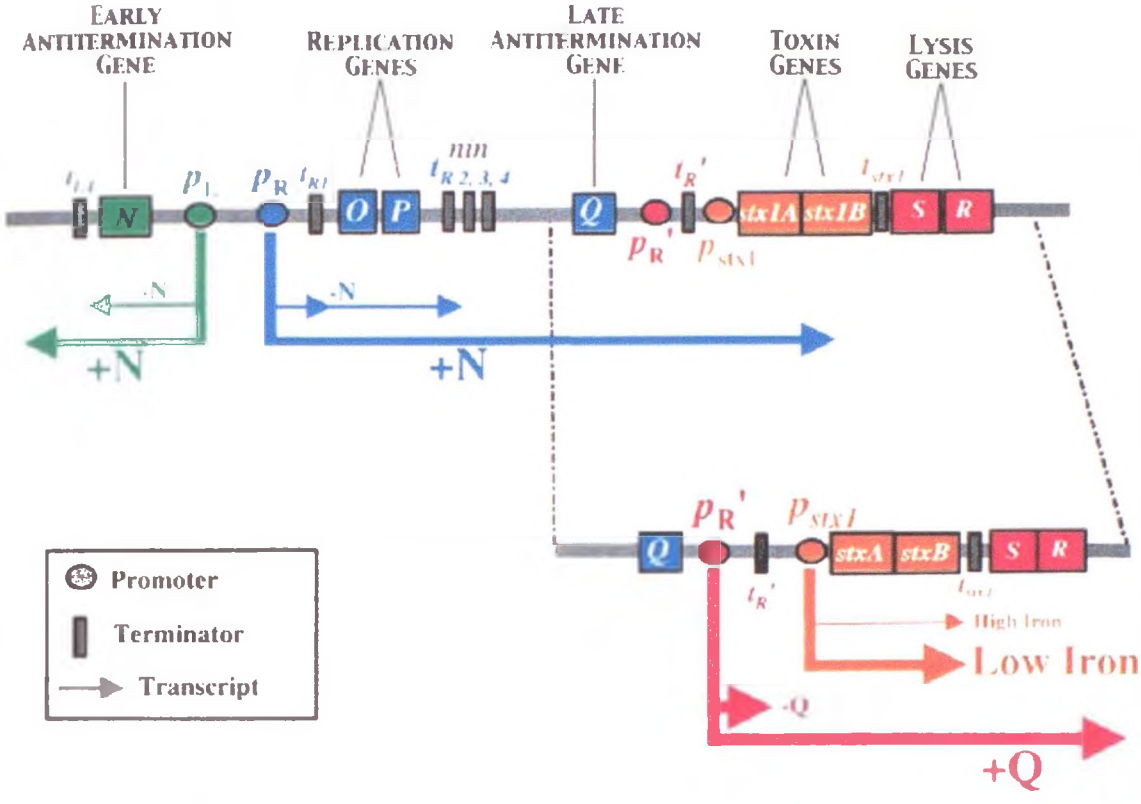


Fig 1.10 Transcriptional Regulation of *stx*

Map and transcription patterns of H-19B phage genome involved in Stx1 regulation (Neely and Friedman, 1998). Stx production occurs on induction of the lytic life cycle, where transcription through Q enables expression of the late genes including *stx*. Under lysogenic regulation, low iron concentrations can stimulate *stx1* transcription using promoter p_{stx1} terminating at t_{stx1}. Gene descriptions and description of regulation can be found in section 1.2. From Neely and Friedman (1998).

most important non-O157 serogroups that have been shown to encode Stx toxins to be O26, O111, O128 and O103. A complete list of STEC virulence factors, toxins and effector proteins identified can be seen in table 1.3.

Adhesion of STEC to the gut mucosa. STEC must adhere to the gut mucosa once it has passed through the stomach, so that it is not removed by the peristaltic flow of the gut. STEC virulence plasmid (pO157) was thought to encode a fimbrial adhesin that allowed attachment to Henle407 (kidney) cells (Karch *et al.*, 1987), although sequencing of a pO157 plasmid by Burland *et al.*, (1997) did not identify a fimbrial gene cluster. Flagellin, the main component of flagella, binds to the Toll-like receptor 5 (TLR5) on the cell, which stimulates the inflammatory response via interleukin 8 (IL-8) (Hayashi *et al.*, 2001). The following adhesins are associated with STEC; Paa, ToxB, Efa – 1 / LifA, Long polar fimbrie (LPF), Saa, OmpA and their impact is described by Kaper, (2005).

Lipopolysaccharide of *E. coli.* Lipopolysaccharide is not directly linked to adhesion, although it can bind to toll-like receptor 4 (TLR4). Bilge *et al.* (1996) identified that LPS-deficient mutants were hyper-adherent to Hep – 2 kidney cells *in-vitro*. It may then be possible that LPS plays a role in masking the adhesins (Law, 2000).

LEE Pathogenicity Island, intimate adhesion and the Creation of AE lesions

The locus for enterocyte effacement (LEE) is located on a pathogenicity island (PAI) comprising a group of genes that are involved in intimate adherence of pathogens to gut epithelial cells (McDaniel and Kaper, 1997). The LEE site is responsible for

Table 1.3 STEC virulence factors, toxins and effector proteins

Factor	Toxin Class	Target	Activity / Effect
Shiga toxin	AB subunit	rRNA	Depurinates rRNA, inhibiting protein synthesis; induces apoptosis
Urease	ABC subunit	Urea	Cleaves urease to NH_3 and CO_2
EspP	Autotransporter		Serine protease; cleaves coagulation factor V.
Cif	Type III effector		Blocks mitosis in G2/M phase; results in inactivation of Cdk1
EspF	Type III effector		Opens tight junctions, induces apoptosis
EspH	Type III effector		Modulates filopodia and pedestal formation
Map	Type III effector	Mitochondria	Disrupts mitochondrial membrane potential
Tir	Type III effector	NcK	Nucleation of cytoskeletal proteins, loss of microvilli, GAP-like activity
StcE	Type II effector	C1-esterase inhibitor (C1-INH)	Cleaves C1-INH, disrupts complement cascade
Ehx	RTX toxin	Erythrocytes / lymphocytes	Cell lysis
Efa		Lymphocytes	Inhibits lymphocyte activation , adhesion

attachment and effacement (AE) lesions of the gut as microvilli are replaced by structures (pedestals) formed out of compact microfilamentous material almost engulfing the bacterial cell (McDaniel and Kaper., 1997) (Fig 1.11 & 1.12). The LEE sites in both STEC and EPEC strains have the same function in that they aid in the intimate adherence of bacteria to gut wall, they are also located in approximately the same site in the bacterial genome, although there is an ~8 kbp difference in size. The LEE region in STEC strain EDL933 (GenBank accession number AF071034) is 43,359bp in size which, is larger than in the EPEC comparative strain E2348/69 which is 35.624bp (Perna *et al.* 1998). LEE in EPEC, EHEC, PAI-I in UPEC are all located 15 bp downstream of the housekeeping gene *selC* (Hacker and Kaper, 2001).

The formation of pedestals and effacement of the gut brush border is complex and has been described in EPEC to involve 12 genes, named *espA* to *espL*. This array of genes is essential for host signalling (Kenny *et al.*, 1996). Mutant analysis of the EPEC LEE region has categorised these 12 genes into 3 groups, allocated with respect to their role in the formation of AE lesions (Kenny and Finely, 1995). Group 1; *espA* and *espB* encode bacterial proteins from the bacteria that are transported into the host epithelial cell (Kenny and Finley, 1995). Group 2; secretory genes *sepA* to *sepJ* encode proteins that show a distinct amino acid identity to those of a type III secretory system (McDaniel *et al.*, 1997). The secretory genes are essential for the transfer of the signalling proteins into the host cell (Jarvis *et al.*, 1995). Group 3; a single adhesion gene *eae* encoding Tir (Frankel *et al.*, 1994).

The bacterial adhesin involved in intimate attachment of STEC to the gut cells is intimin (Batchelor *et al.* 2000). An intimin receptor protein (Tir) (previously Hp90) is translocated into host gut cell, phosphorylated and located at the membrane forming a hairpin loop structure (Luo *et al.*, 2000). This structure yields a surface

binding epitope for intimin (Hicks *et al.,* 1998) and the amino and carboxy terminals located in the gut cell cytoplasm (Hartland *et al.*, 1999) as a focus point for actin polymerisation and cytoskeleton rearrangement (Frankel *et al.*, 1998). Therefore, Tir, as well as binding intimin giving the close adherence exploits the actin signalling cascades within host cells upon its interaction with intimin (Campellone and Leong, 2003). Levels of intimin production have been associated with another LEE-encoded gene *ler* which when O157:H7 strain EDL933 was compared to an isolated single base pair mutant of this gene showed lower production of intimin (Ogierman *et al.*, 2000).

EPEC / EHEC target the heptameric actin-related protein 2 / 3 (Arp 2 / 3) pathway which is involved with actin nucleation (Welch *et al.*, 2002). Actin nucleating activity of the Arp 2 / 3 complex is stimulated by the Wiskott – Alderich protein (WASP) family (Campellone and Leong, 2003). Wasp proteins in turn are activated in EPEC by either rho family GTPases or adaptor proteins (Campellone and Leong, 2003). In EPEC, adaptor protein Nck binds to phosphotyrosine, which in turn stimulates N-Wasp, which initiates actin polymerisation (Rohatgi *et al.*, 2001). Fig 1.13 shows the that difference between and EPEC and EHEC is that actin polymerisation is independent of Nck adaptor protein, but not WASP.

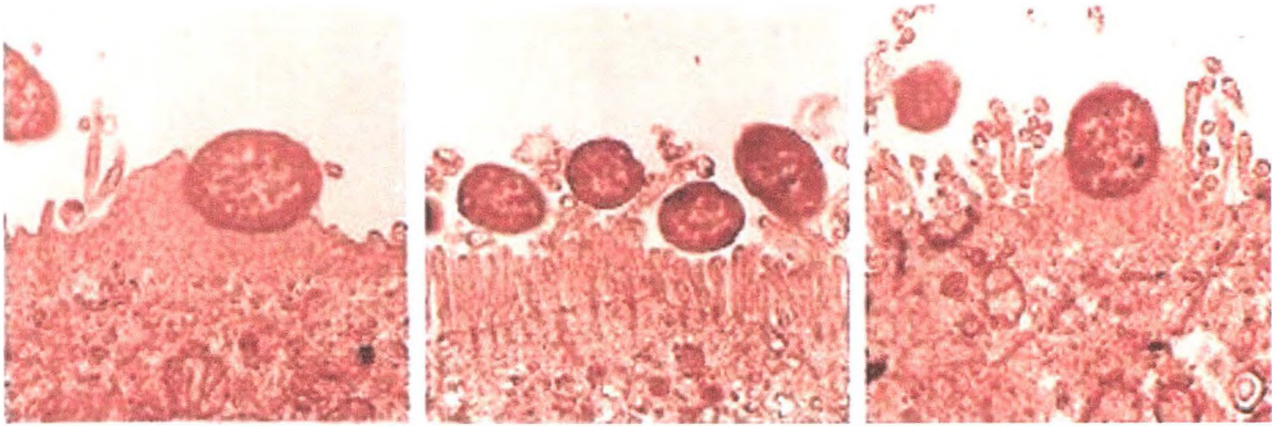


Fig 1.11 Pedestal formation: alteration of the actin cytoskeleton of gut mucosal cells during EPEC infection. Transmission electron microscopy of cultured intestinal epithelial cells infected with wild type enteropathogenic *E. coli* (EPEC), *E. coli* K-12 or recombinant clones. Left: *E. coli* K-12 strain HB101 containing cloned 35 kb Pathogenicity Island that confers attaching and effacing histopathology. Middle: *E. coli* K-12 strain HB101 without cloned pathogenicity island. Right: Wild type EPEC strain E2348/69 showing attaching and effacing histopathology (from McDaniel & Kaper, 1997)

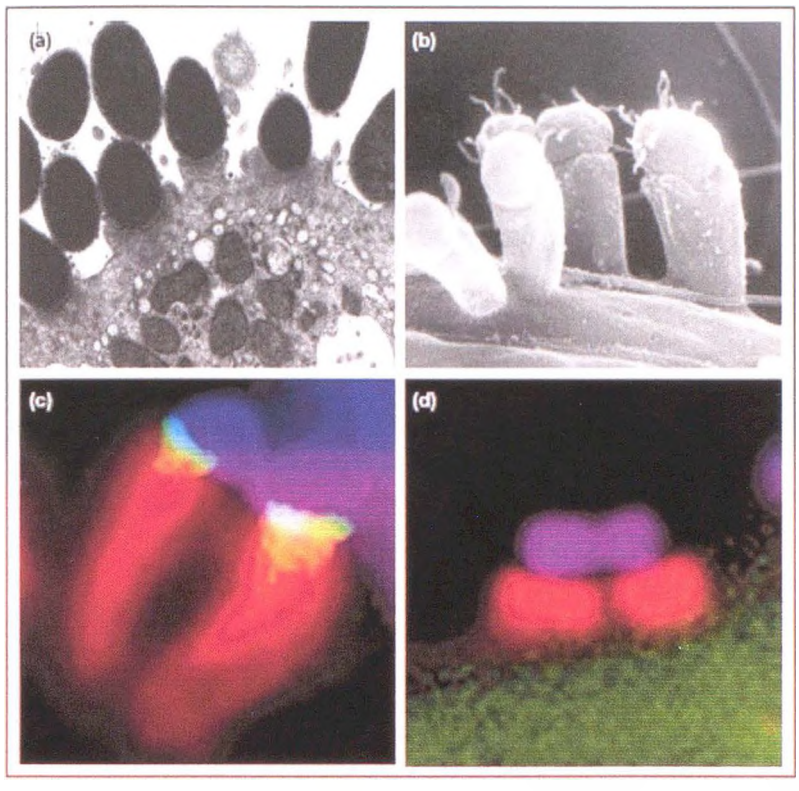


Fig 1.12 Pedestal Formation and actin rearrangement. Actin pedestals of EPEC and EHEC. (a) EPEC generates AE lesions on the intestinal epithelium after infection of gnotobiotic piglets. (b) Actin pedestals that resemble AE lesions formed *in vivo* are also generated on cultured epithelial (HeLa) cells. (c) The host adaptor protein Nck localises to the tips of actin pedestals generated by EPEC. Cultured HeLa cells were infected with EPEC and examined by immunofluorescence microscopy. F-actin is shown in red, bacterial DNA in blue and the host protein NCK in green. Co-localisation of Nck and F-actin beneath bacteria are depicted in yellow. (d) Nck is not recruited to sites of EHEC pedestal formation. Cultured HeLa cells were infected with EHEC and examined by immunofluorescence microscopy as in (c) (from Kaper, 2004).

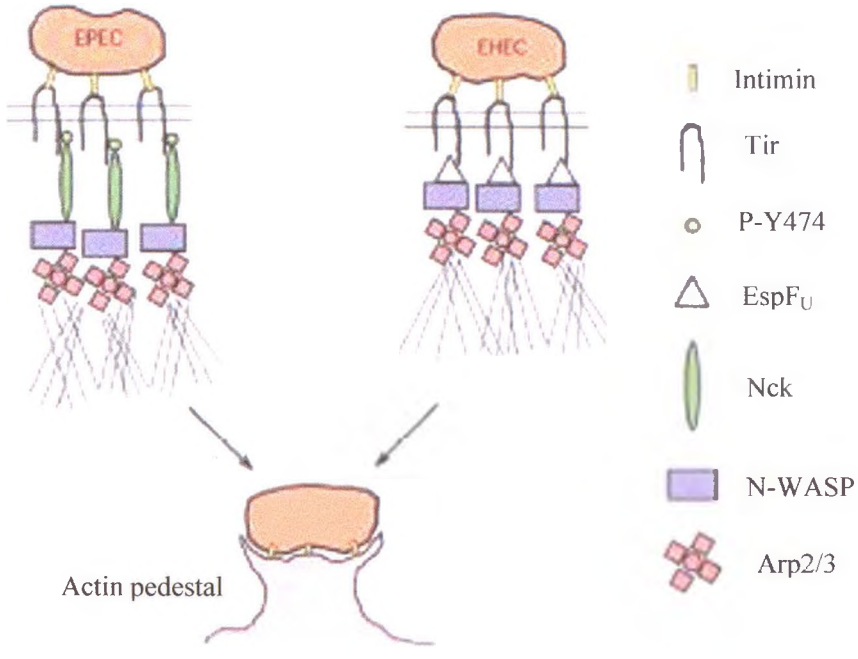


Fig 1.13 Schematic diagram of actin cytoskeleton rearrangement, including effectors of EPEC and EHEC colonisation. Diagram shows differences between pedestal formation of EPEC and EHEC (from Campellone and Leong, 2003).

(Campellone and Leong, 2003). EHEC uses a proline-rich protein $EspF_U$ an effector similar to Esp which has the ability to localise or bind N-WASP so that pedestal formation can occur (Campellone *et al.*, 2004).

Other genes have been identified in the LEE region that do not contribute in the ability of the pathogen to form AE lesions, but do have implications for initial piliated adherence of the pathogen to the gut cell. Doughty *et al.* (2002) identified long polar fimbriae genes encoded on LEE (*lpf*). Through mutant analysis it was found that reduced adherence occurred when compared to the wild type and was thus restored by complementation (Doughty *et al.*, 2002).

Plasmid pO157. There is a large F-like virulence plasmid (~92Kb) present in most VTEC O157:H7 strains called pO157 (Burland *et al.*, 1998). Study of pO157 has shown high levels of heterogeneity across STEC in both composition and arrangement (Schmidt *et al.*, 1999).

pO157-borne Enterohaemolysin. The STEC-enterohaemolysin (Ehx) was the first sequence to be determined on the pO157 plasmid and was named thus to distinguish it from alpha-haemolysin to which it is related, but not identical (60 % sequence identity) (Schmidt *et al.,* 1995, 1996). Secretion of the VTEC-haemolysin seems to be defective, as only small zones of haemolysis or a lack of haemolytic activity in the supernatant fluids occurs (Schmidt *et al.,* 1996). Production levels of Ehx between serotypes vary. For example sero-group O103 produces the same haemolytic activity as alpha-haemolysin with the *ehx* gene; it is thought that in this case that Ehx is over expressed (Schmidt *et al.*, 1999). The Ehx operon is highly conserved and so could play a role in survival or confer an advantage upon STEC infection (Law, 2000).

Schmidt and Karch (1996) described how 16 of 18 strains that were causative of HUS carried the *ehx* gene, compared to 4 of 18 of diarrhoeal related isolates. Low oxygen environments have been shown to increase the production of Ehx (Chat *et al.*, 1998), which can relate with the conditions *E. coli* that would encounter in the human gastrointestinal tract (Law, 2000). Ehx is not needed for development of sequelae such as HC and HUS, so its role in VTEC infection is unclear (Law, 2000).

pO157-borne Catalase-Peroxidase. KatP is an 82 kDa bi-functional catalase-peroxidase (Brunder *et al.*, 1996). *KatP* has been identified in ~ 66 % of O157 isolates compared to 38 % in non-O157 isolates (Brunder *et al.*, 1996). The enzyme accumulates in the periplasm of the cells and may have a role in detoxifying oxidants produced by macrophages and neutrophils, aiding the bacterial cells to evade host responses (Law *et al.*, 2000). There is no connection between the incidence of *katP* in *E. coli* isolates and the occurrence of HC and HUS, and it may, therefore, not be greatly involved in the virulence of STEC (Law, 2000).

pO157 borne *Clostridium difficile*-like Toxin. A large open reading frame (ORF L7095) on pO157 encodes a 3169 amino acid protein showing similarities to a family of toxins called the large clostridial toxins (LCT) (Burland *et al.*, 1998). The toxin includes a C-terminal domain which promotes entry into the cell and an N-terminal glucosyl-transferase that modifies proteins regulating cell architecture leading to disorganisation of the cell cytoskeleton (Law *et al.*, 2000). Similarities between the pathology of STEC infection and *Clostridium difficile*-colitis have been noted (Nataro and Kaper, 1998).

pO157-borne Extracellular Serine Protease (EspP). EspP is an autotransporter mediating its own secretion through the outer membrane (Law, 2000). EPEC and EHEC have been shown to secrete EspP into culture supernatant (Kaper *et al.* 1998). EspP is a 104 kDa protein that shows similarities to EspC of EPEC which is a secreted protein of no designated function is related to a narrow host range serine-protease family with ability to target and cleave human coagulation factor V (Law, 2000). This may influence the blood clotting cascade which would prolong bleeding and in turn allow increased haemorrhaging into the gastrointestinal tract (Kaper *et al.*, 1998).

A type II secretion pathway was found on 100% of fifty STEC 0157 strains encoded by pO157 in Burland's study (1998), although the identity of proteins secreted via this function is unknown in STEC (Law *et al.*, 2000).

Other STEC-related virulence factors. A novel prophage-encoded effector has been identified in both EPEC and STEC, named Cif (Marches *et al.*, 2003). Cif is translocated into the gut epithelial cell via the type III secretion system encoded by LEE (Marches *et al.*, 2003). Cif promotes actin cytoskeleton rearrangement and G2 cell cycle arrest, and this is identified by inactive phosphorylated Cdk1 in the cell (Marches *et al.*, 2003). Cif is the first described cyclomodulin translocated through the LEE encoded Type III secretion system that is encoded on a prophage and not on the LEE PAI.

STEC Acid tolerance. STEC must pass through the acidic environment of the stomach if they are to colonise the gut (Law, 2000). *E. coli* survives these acidic conditions by a number of responses. At pH 2.0, an arginine-dependent response

provides increased protection to STEC strains when compared to commensal *E. coli*, although, when both the commensal and STEC strain responded to the change in pH with a glutamate-dependent response results were comparable (Lin *et al.* 1996). An oxidative response that allows survival for over 2 hours at pH < 2.5 is dependent on the *rpoS* regulon (Law, 2000). O'Brien *et al.* (1997) describes how the ability of STEC to withstand acidic conditions varies from serotype to serotype. Studies have shown that when O157:H7 strains have been compared with commensal gut isolates no difference was seen with respect to acid tolerance (Foster, 2004). Heat stress has been shown to increase the resistance to low pH (Wang and Doyle, 1998), which may have implications in the re-heating of contaminated food or cooked meats (Law, 2000).

1.10e Animal Reservoir and Seasonality of STEC Shedding.

As our understanding of the STEC animal reservoir and seasonality of STEC shedding increases, it is possible to try to implement precautions to lower the opportunity for STEC infection to be acquired. STEC has been isolated and characterised from a range of livestock, predominantly cattle which is thought to be the primary reservoir of STEC. Mature cattle carry STEC, but do not show symptoms (Beutin and Muller, 1998). STEC have also been isolated from other animals such as sheep (Djordjevic *et al.*, 2004), pigs (Bonardi *et al.*, 2002), deer (Keene *et al.*, 1997), greyhound dogs (Staats *et al.*, 2003), seagulls (Morabito *et al.*, 2001) and even pigeons (Schmidt *et al.*, 2000). There have also been reports that fruit flies are able to act as vectors of STEC infection (Janisiewicz *et al.*, 1999). STEC has been reported to proliferate in the mouth-parts of house flies and is excreted for up to 3 days after (Sasaki *et al.*, 2000).

Ruminant shedding of STEC has been widely reported and studied (Hancock *et al.*, 1997). Bonardi *et al.*, (1999) in Italy and McEvoy *et al.*(2003) in Ireland in all cases showed increased shedding of STEC in the summer months. The study of Ogden *et al.* (2004) in Scotland differed from these other reports because they found increased shedding during the summer in their animals that showed high levels of shedding, but there was increased prevalence of STEC during the winter months. The explanation offered was that cattle are housed in Scotland during the winter and had levels of contact that led to increased prevalence of STEC (Ogden *et al.*, 2004).

During the summer months multiplication of STEC in water troughs and cattle feed can lead to continuous re-infection of animals (Hancock *et al.*, 2001). It has also been observed in slaughter houses that detection of STEC in the summer months can rise from an isolation rate of 2.9 % to 17. 9 % (Bonardi *et al.*, 1999). This seasonal increase has also been identified by examination of meats products (Chapman *et al.*, 2001).

1.11 Aims & Objectives

The aim of the research described in this thesis was to characterise the infection of *E. coli* with an Stx-phage ($\Phi 24_B$) in which the s*tx* gene had been interrupted with a selectable antibiotic resistance marker. The initial part of the infection process is contact between the phage tail spike and the *E. coli* surface receptor and this was the main focus. As the bacterial outer membrane protein Vpr, for which there is evidence of a role as a phage receptor, had been previously identified, an important aim of this project was to unequivocally establish that the protein was the receptor recognised by this short tailed Stx phage. The approach taken comprised the following elements:

- Development of a replicable signature phage adsorption assay to study the phage tail-receptor interaction.
- Production of a knock out mutant of the Vpr gene to demonstrate removal of sensitivity to Φ24B infection.
- Blockage of the epitope of Vpr, responsible for phage binding, using a poly-clonal antibody and detection by confocal microscopy of the localisation of Vpr
- Examination of levels of transcription of v*pr* when the *E. coli* growth environment was altered using a $_{P}$*vpr*::*LacZ* reporter gene construct.
- Comparison of the changes in transcription levels to translation and localisation by the cell using the phage adsorption assay.
- Compare tail spike genes of wild type Stx phage, to define a region of the tail fibre gene associated with phage adsorption.

Although λ bacteriophage is very well characterised, this is not true of lambdoid-like Stx-phages. The following key events in the life cycle of Stx-phage were also studied.

- Integration of the phage into the host genome
- SEM of the lytic life cycle of an induced Stx-phage lysogenised host

Chapter 2: General Materials and Methods

2.1 Materials

2.1.1 Media, Bacterial strains, Growth and maintenance

All media ingredients were obtained from Biogene except high clarity agar used for soft agar overlay (Difco) and MacConkey Agar (Oxoid). All *E. coli* strains were propagated at 37 ^{o} C unless described otherwise either in liquid or on solid Luria Bertani media (LB) (Merck). Liquid cultures were incubated in a shaking incubator at 200 rpm. Antibiotic supplements (Sigma) were stored at -20 ^{o} C and added where appropriate (Table 2.1). All containment level 2 strains were stored in 50% (v/v) glycerol at -80 ^{o} C; containment level 3 isolates were stored in 100 % glycerol at -20 ^{o} C

Table 2.1 Antibiotic supplements

Antibiotic	Solvent	Stock Conc. (mg ml^{-1})	Final Conc. (μg ml^{-1})
Ampicillin (Amp)	H_2O	100	100
Chloramphenicol (Cm)	100 % EtOH	50	50
Kanamycin (Kan)	H_2O	50	50
Norfloxacin (NFLX)	H_20	1	1
Rifampicin (Rif)	100% EtOH	34	300
Streptomycin (Sm)	H_2O	50	50

2.1.2 Bacterial Strains

Wild type STEC isolates were obtained from the University of Liverpool, Department of Veterinary Clinical Science, Leahurst. STEC strain EDL933 was obtained from the American type culture collection (ATCC) via LGC technologies, UK. Phages P27 and 3538::Cat were donated as lysogens in *E. coli* K-12 strains C600 and DH5α respectively by Prof. Herbert Schmidt, University of Dresden, Germany. *Erwinia carotovora* subsp. atroseptica (*ECC*) was obtained from the German culture collection (DSMZ). *E. coli* strain DM1187 was used as the host for all of the bacteriophage enumeration work done in this study, as it contains a mutation (*recA441*) that results in a constant state of proteolytic activation through expression of recA, thus directing lysis (see Fig 1.3). *E. coli* K-12 strain MC1061 was used as a host for the productions of lysogens. Topo (Invitrogen) *E. coli* cells were used as competent cells for cloning for subsequent nucleotide sequencing. The plasmids used in this work include are listed in Table 2.2.

Recombinant DNA polymerase was supplied by MBI Fermentas and proof reading DNA polymerase *pfx* distributed by Invitrogen by Life Technologies. T4 DNA ligase and restriction enzymes were supplied by NEB. All enzymes were used according to the manufacturers' instructions.

All glassware and aqueous solutions were sterilised by autoclaving at 121 ^{o}C for 15 minutes (15 lb in^{2}).

Table 2.2 Plasmids used in this study

Plasmid Construct	Relevant genotypes	Study
pUC19ΦR1D	*bla*, *lacZα*	Sergeant, 1998
pYY*vpr* construct using pJP5603	*aph3*, R6K	Yan Yaxian (un-published data)
pKT230	*aph3*, *Sm*, *Su*	Bagdasarian, 1991
pCR3.1	*neo*, *bla*	Invitrogen
pKT230::*vpr*	*aph3*, *Sm*, *Su*	This study

Key:

bla β-lactamase gene conferring ampicillin resistance to *E. coli.*

lacZα Encodes α-peptide of β-galactosidase

R6K Origin of Replication of origin that requires bacterial host encodes λ-pir protein, ack of *pir* leads to inability to replicate and when used in conjunction with an antibiotic selection marker can lead to cell death under the selective pressure.

aph3 Confers kanamycin resistance on *E. coli*

neo Confers neomycin resistance on *E. coli*

Sm Confers streptomycin resistance on *E. coli*

Su Confers sulphonamide resistance on *E. coli*

2.2 Bacteriophage: Propagation and Enumeration

2.2.1 Stx-phage constructs

An Stx-phage induced from strain ϕE86654, a clinical isolate of *E. coli* O157:H7 (Colindale Public Health Laboratories, CHPL) expressing the Stx2 toxin, had the *Stx*2A gene inactivated with a kanamycin resistance cassette (*aph3*) from plasmid pUC4K (Pharmacia) (Sergeant, 1998). This phage ($\phi 24_B$ (*Stx2A*::*aph3*)) was renamed $\phi 24_B$::kan. A second construct was made using the same Stx-phage wild-type including a truncated *Stx2A* gene and the inclusion of a chloramphenicol acetyl transferase gene (*cat)* from pLysS (Novagen) (Allison *et. al.*, 2003; James, 2002). This phage ($\phi 24_B$ (*Stx2AΔ*::*cat*)) was named $\phi 24_B\Delta$::cat.

Another STX2-phage was obtained from Professor Herbert Schmidt at the University of Dresden, Germany. This phage also had the s*tx2A* gene interrupted with a *cat* gene and called ϕ3538::cat (Schmidt *et al.*, 1999).

All phage stocks were stored at 4 °C in LB plus 0.01M $CaCl_2$ (phage buffer).

2.2.2 Enumeration of recombinant Stx-phages

A ten fold dilution series of lysate (0.05 ml in 0.45 ml) was mixed with mid-exponential growth phase indicator host and incubated at 37 °C for 25 min. This infection mix was then added to 5 ml molten top agar (4 % Difco Agar in LB, 0.01M $CaCl_2$), mixed and then plated on to LB agar. The pre-vented plates were left to set and then incubated at 37 °C overnight (James *et. al.*, 2001) and then the plaques enumerated.

2.2.3 Propagation of recombinant Stx-phages

To propagate a pure phage stock, a single plaque was picked using a sterile Pasteur pipette and re-suspended in phage buffer (0.1 ml). This phage suspension was then propagated and titred by plaque assay.

The surface of the resulting semi-confluent lysis plates was scraped and an equal volume of phage buffer added. This suspension was stored overnight at 4 ^{o} C to enable the phage to dissipate into the buffer, and was then membrane-filtered (0.45 μm pore size; Sartorious). This step usually yielded approximately 10^{7} - 10^{8} pfu ml^{-1}. An indicator host (DM1187) culture (100 ml) was grown to mid-exponential growth phase and infected with 10 ml of the phage suspension (10^{7} - 10^{8} pfu ml^{-1}). The infected culture was incubated at 37 ^{o} C overnight. When bacterial debris became visible, chloroform (0.1 ml) was added for a further 15 min to lyse the remaining bacterial cells. The lysate was vacuum-filtered through a manifold (100 ml, Nalgene) containing a 0.45 μm diameter pore size filter. The phage stock was titred by plaque assay and typically yielded 10^{8} - 10^{9} pfu ml^{-1}. Phage stock was stored at 4 ^{o} C.

2.2.4 Induction of the Lytic Life Cycle.

Induction of Stx-phage by stimulation of the lytic life cycle was achieved by exposing mid-exponential growth phase putative lysogen cultures to Norfloxacin (1 μg ml^{-1}) (Matsushiro, 1999) for 1 h at 37 ^{o}C. Recovery of the culture was achieved by subculturing (1 ml) into fresh medium (10 ml LB + 0.1 M $CaCl_2$) for a further 2 h at 37 ^{o} C. The lysate was filtered (0.45 μm) and titred by plaque assay.

2.2.5 Induction of wild-type Stx-phages from STEC isolates

The Containment level 3 facility in the School of Biological Sciences at the University of Liverpool does not contain a PCR machine, therefore a technique was developed to yield sufficient phage DNA template to be used for PCR amplification in the containment level 2 laboratory.

A 10 ml mid-exponential growth phase culture was subjected to Norfloxacin (NFLX) (1 μg ml^{-1}) induction 37 ° C for 1 h. The culture was recovered by sub-culturing (1 ml) into LB + 0.01M $CaCl_2$ (10 ml) to dilute the NFLX. The bacterial cells were recovered at 37 ° for 2 h. A portion of the lysate (50 μl) was serially diluted and propagated via plaque assay method described above. The host strain used for this infection was Rifampicin-resistant DM1187 (DM1187 Rif $^+$) and the antibiotic used to inhibit carry over of wild-type STEC strains used for the phage induction. The plates were incubated overnight at 37 ° C.

A single plaque was picked and propagated by plaque assay against DM1187 Rif $^+$ using the method described above. A single confluent lysis plate was scraped and an equal amount of LB +0.01M $CaCl_2$ added; this was left for approx 4-6 h to enable the phage particles to dissipate into the buffer. 1.5 ml of this lysate was then incubated at 80 °C for 20 min to kill any bacteria present, and denature the viable viral progeny. The tube was then surface-sterilised with 2 % stericol and transferred from the containment level 3 suite for the PCR analysis. This protocol was optimised using model phage $\Phi 24_B$ lysogenised in MC1061. The maximum levels of phage diffusion from the soft agar were found after ~ 6 h incubation. The 80 ° C denaturation step was shown to kill both bacteria and the virus such that the preparation could not be used to infect *E. coli i.e.* no plaques were detected by plaque assay.

2.2.6 Lysogenic infection

Lysogenic infection was stimulated by infection assays where the multiplicity of infection (MOI) was maintained at 0.1. Variation from this MOI can have an influence on the rate that lysogenic infection can occur, low MOIs have been shown to increase the possibility of lytic infection (section 1.2). Therefore, using *E. coli* MC1061 at mid-exponential growth phase OD_{600} 0.5 (~ 10^{8} cfu ml^{-1}) cells were infected with bacteriophage to give a final amount of phage of 10^{7} pfu ml^{-1} in the infection. Infection mixes were incubated for 25 min and then plated on media using the positive selective pressure of the bacteriophage, in this case either kanamycin or chloramphenicol depending on the labelled variant of $\Phi 24_B$ used.

2.4 Molecular microbiological techniques.

2.4.1 Polymerase Chain Reaction (PCR)

Two DNA polymerases preparations were used in this study; a recombinant, non-proof reading DNA polymerase (Taq) from MBI Fermentas and a recombinant DNA polymerase, proof reading enzyme called PLATINUM *pfx* (Invitrogen), which was used for amplification of DNA regions for downstream applications such as sequencing and cloning. Both enzymes were used according to manufacturers' guidelines.

The MBI Fermentas DNA polymerase reaction mix contained (per 50 µl reaction): *Taq* (1 unit); primer pair (100 nM each); dNTP mix (100 µM); 1x buffer (10 mM Tris. HCl pH 8.8, 50 mM KCl, 0.08 % Nonidet P40); $MgCl_2$ (1.5 mM). Typical thermal cycling conditions comprised an initial denaturation of 4 min, 94 °C; followed by 35 cycles of: denaturation (94 °C, 1 min); annealing (anywhere between 40 °C - 60 °C according on the Tm of the primers); extension (72 °C, 1 min per 1

Kbp product). A final extension was also added to ensure completion of amplified products (72 °C, 7 min).

PLATINUM *pfx* DNA polymerase reaction mix contained (per 100 µl reaction): *Pfx* (2.5 units); primer pair (300 nM each); dNTP mix (300 µM); 1 X PCR amplification buffer (not described, Invitrogen); $MgSO_4$ (1 mM). Cycling parameters were similar to those described above except that the elongation temp was 68 ° C and no final extension was required. All PCR amplifications were performed using Applied Biosystems (Applera Corporation) Thermal Cyclers (models: 480; GeneAmp PCR system 2400; GeneAmp PCR system 2700) or MJ research DNA engines.

All primers used in this study were derived from alignments using NCBI database sequences, or sequences determined by studies at the University of Liverpool by Sergeant (1998) and Allison *et al.* (2001). All primer sequence parameters, fragment sizes are described in Table 2.3. Templates used for PCR analysis were as follows: 1-2 µl of DNA preparations (1-100 ng); 1-2 µl of a homogenised colony from a plate incubated overnight in H_2O (50 µl); 2 µl of filtered phage lysate (10^9 pfu ml^{-1}).

Table 2.3 Oligonucleotide primers

Primer	Sequence 5' – 3'	Target gene	Tm (°C)	Reference
VTTF1-Fwd VTTF2-Fwd VTUTF-Rev	GTT GTT GTT TCG GGG ACG GTG CTG AAA GAT GGT GCG TCA TTC TCC TGT TCT GCC	Stx-phage short tail spike gene	64.0 63.2 58.8	This study
VTTF3-Fwd VTTF3-Rev	TGC AGA GGA AAG CTC GAC GCA GCC TCT TCT GCC TTT	Internal Stx-phage short tail spike gene	61.9 62.3	This study
pVprLacZ-Fwd pVprLacZ-Rev	TCG CGA GCC AGT TTT GCC G GAA TCC GTA ATC ATG GTC ATC GTT ATT ATG C	Tailed $_{p}$*vpr* oligonucleotides	70.9 70.8	This study
K12LacZ-Fwd K12LacZ-Rev	ATG ACC ATG ATT ACG GAT TC TTA TTT TTG ACA CCA GAC C	*LacZ*	58.6 54.8	This study
LacZinternal1-Fwd LacZinternal1-Rev	CAT TAC GGT CAA TCC GCC TCG AAC GCT GGA AGG CGG	Internal sequencing *LacZ*	72.1 63.9	This study
Gp1AF Gp1CR	GTT ACM GGG CAR MGA GTH GG ATG CCC GAG AAG AYG TTG AGC	*Int* of phages; λ, HK97, HK022, P434, H19J	64.0 65.0	Balding *et al.*, (2005)
Gp2AF Gp2CR	GTT ACT GGW CAR CGK TTA GG GAT CAT CAT KRT AWC GRT CGG T	*Int* of phages; P21, e14	60.0 63.0	Balding *et al.*, (2005)
Gp3AF Gp3CR	AAC ATY ATC AAY CTK GAR TGG CA CGA ACC ATT TCG ATA GAC TCC CA	*Int* of phages; P22, ST64T, SfII, phage V, DLP12, qsr	66.0 64.0	Balding *et al.*, (2005)
Gp4F	TYA CBC TRC CWA ARA CHG AMG C	*Int* of phages;	66.0	Balding *et al.*, (2005).

Gp4R	GAT AAW GMC CAG CAB GCA TAR G	P27, 933W, Gifsy-2, EH297, VT2-Sa	65.0	
Gp5F Gp5BR	AAA MMN CGH ACC GTS CCR AT TDC CBC CRT TMA TCA TRA ART G	*Int* of phages; P2, P186, WPhi, Fels-2	61.0 62.0	Balding *et al.*, (2005)
Gp6F Gp6R Gp6AF Gp6AR	CGA TRG TRR TGG YYT GTA YCT T ART CMG CCC ACC ACT GCA TC TGG RAK RAM KTC GAY TTY GAW AA ATG GWT RTA WRY CSC ACG YAC A	*Int* of phages; P4, HK620, Sf6	64.0 64.0 63.0 64.0	Balding *et al.*, (2005) Balding *et al.*, (2005)
Gp7AF Gp7CR	GST GAR MTY CGW CWK RST GA TGC CCG AGC AKC WTY TCA	*Int* of phages; phi80, Gifsy-1	55.0 58.0	Balding *et al.*, (2005)
Gp8AF Gp8CR	TGC TTA TAA CAC CCT GTT ACG TAT CAG CCA CCA GCT TGC ATG ATC	*Int* of phages; P1	64.0 68.0	Balding *et al.*, (2005)
Gp9F Gp9R	AAT GGA RAT WKC YTA TYT VTG TGC TCR TAR TCT GAR ATY CCY TTB GC	*Int* of phages; St64B	61.0 61.0	Balding *et al.*, (2005)
933Wint5' 933Wint3'	GCT GGC ACG ATA ACA GTG C GGC ACG GGC ATT AAG GAC	Phage 933W *int* gene	64.5 65.4	This study
InternalGp6F InternalGp6R	GAC ATC TCG GTT GGC ATC CTG AAT TCA TGC CTG AAT	Internal sequencing primers *int* $\Phi 24_B$	61.8 55.8	This study
24Bintprobe-F 24Bintprobe-R	CTG GAA GTA ATC CGC AGG AGC TCT TTC GTT CTT AGG	Internal probe primers *int* $\Phi 24_B$	60.8 53.6	This study
5' vprBamHI 3' vprEcoRI	CG**G GAT CC**T CGC GAG CCA GTT TT CG**G AAT TC**G TGG AGA ACA CTT AC	*Vpr*	74.0 70.0	This study

3' Stx2A 5' Stx2A	TCT GTT CAG AAA CGC TGC TAC TGT GCC TGT TAC TGG	Stx toxin gene	58.0 63.0	Allison *et al.*, 2003
5' Kan 3' Kan	AAT GTC GGG CAA TCA GG GAA TCC GGT GAG AAT GG	Kanamycin gene	63.0 63.0	Allison *et al.*, 2003
5' Cat 3' Cat	AAC TGC AGA AAT GAG ACG TTG ATC GG AAC TGC AGC CTT AAA AAA ATT ACG CC	Chloramphenicol gene	73.0 70.0	Allison *et al.*, 2003
5' Q 3' Q clone 5' ATG	CAC TGG CGA TAA AGA AGG TCT TAT CAT GAT ATG CAG ATG TTC TTA TGG TTC ACC G	Phage anti-termination gene	63.0 58.0 63.0	Allison *et al.*, 2003
5' GAPDH 3'GAPDH	ATG ACT ATC AAA GTA GGT ATC TTA TTT GGA GAT GTG AGC G	*E. coli* housekeeping genes	63.0 63.0	James, 2003

Emboldened nucleotides denote restriction enzyme recognition site.

2.4.3 Agarose gel electrophoresis.

0.75 % (w/v) agarose (Sigma) mix in 1x TAE buffer was heated to approximately 50°C. Ethidium bromide (Amersham Life Sciences) was added to a concentration of 0.5 ng ml^{-1}, and the mixture poured into a gel-forming tray and allowed to set for 30 minutes. The gel was then immersed in a volume of 1x TAE buffer sufficient to cover the wells and electrodes. DNA samples were mixed with 5x loading dye (Bioline) and run alongside molecular weight marker Hyperladder I (Bioline), λ *Hind* III (MBI Fermentas). The gels were run at 60mA for 1-1.5 hours and were viewed under UV transillumination and the presence of DNA recorded using a Syngenta imaging system using GeneSnap software.

2.4.4 DNA Sequencing.

DNA sequencing was carried out by the University of Liverpool, School of Biological Sciences Sequencing Service using a Beckman-Coulter CEQ2000, and by MWG Biotech. The subsequent sequences were extrapolated using BLAST against the National Centre for Biotechnology Database (Altschul *et al.*, 1997).

2.4.5 DNA-DNA Hybridisation

The protocol for Southern blotting described by Southern (1975) was used. A DIG Nucleic Acid labelling and detection system (Roche) was used following the System Users guide for Filter Hybridisation © 1995.

Bacterial genomic DNA was digested using restriction enzymes according to manufacturer's instructions (New England Biosciences (NEB)). The digested preparation was separated by agarose gel electrophoresis (0.75 % w/v) at 70 V. cm^{-1} for 3-4 h. A partial depurination was applied in 0.25 M HCl (10 min), to increase the

efficiency of DNA transfer. Denaturation and neutralisation of the restricted DNA was achieved according to the guidelines described in the DIG Systems user guide (Roche). The DNA was transferred to a nylon membrane (Hybond) by capillary transfer using 20 x SSC with 3MM (Whatman) chromatography paper. The capillary transfer was allowed to proceed for ~ 16 h. The DNA was then covalently linked to the membrane using a DNA cross linking system (302 nm UV irradiation; 1 min).

2.4.6 Digoxigenin (DIG) labelled probes.

Digoxigenin probes were synthesised by integrating a DIG-11-dUTP into the PCR reaction and thus the target amplification product. All other parameters in the original PCR reaction specific for that target sequence were unaltered (Table 2.3).

Hybridisation of the DIG-labelled probes was achieved by following the protocol in 'The System Users Guide for Filter Hybridisation ©1995' (Roche). Nylon membranes were equilibrated with pre-hybridisation solution (50 % formamide) for 2 h at 42 ^{o}C. The labelled probes were denatured in a heating block at 100 ^{o}C for 10 min and then mixed with pre-warmed pre-hybridisation buffer and then applied to the membranes. The hybridisation process was left overnight at 42 ^{o}C in a carousel hybridisation oven (Hybaid).

Subsequent to the hybridisation step, the unbound probes were washed from the nylon membrane as per the DIG System Users guide for Filter Hybridisation © 1995 (Roche). The nylon membranes were then equilibrated and placed into blocking solution for 1 h.

Detection of the probe was achieved using an antibody conjugated with alkaline phosphatase (anti-Digoxigenin-AP (Roche)) that has the ability to react with a luminescent substrate (CPD* or CSPD (Roche)); these components were used

according to the DIG Systems User guide (Roche). The chemiluminescent signal was visualised by exposing the membrane to imaging film (Biomax light-1, Kodak). The film was then developed, and analysis completed using the GeneGenius BioImaging and Gel documentation System (Syngene).

2.4.7 DNA purification via agarose gel extraction

DNA fragments of interest were excised from agarose electrophoresis gels after separation and purified using the QIAquick Spin Gel Extraction Kit as per the manufacturer's instructions (Qiagen).

2.4.8 Plasmid DNA purification

Plasmid DNA (5-10μg) was extracted from small overnight cultures (1-10 ml) using Qiagen's Mini and Midi plasmid preparation kits. Large scale plasmid DNA preparations were achieved using alkaline lysis and PEG precipitation as described by Sambrook *et al.* (1989).

2.4.9 Endonuclease Restriction Digests

All endonucleases were used according to manufacturer's instructions (NEB). Digest reaction mixes ranged in volumes from 20 μl to 400 μl depending on the application.

2.4.10 Bacterial chromosomal DNA extraction

Bacterial chromosomal DNA was extracted using a method described by Ausubel (1992). 1.5 ml mid-exponential growth phase cells were harvested by centrifugation, lysed by re-suspending the harvested cells in TE (10 mM Tris-HCl; 1

mM EDTA; pH 8) containing SDS (0.5 % w/v) and proteinase K (100 μg. ml^{-1}) and incubated for 1 h at 37 ° C. The cell debris was removed by precipitation using a CTAB solution (10 % CTAB (w/v) 0.7 M NaCl) with incubation at 65 ° C (10 min). Removal of proteinaceous material and thus purification of DNA was achieved using a phenol: chloroform: isoamyl alcohol (25:24:1) extraction and the DNA was recovered by ethanol precipitation (Sambrook *et al.*, 1989). Samples were resuspended in H_2O (37 ° C) and stored at – 20 ° C.

2.4.11 Transformation of bacterial cells

Preparation of competent *E. coli* cells was achieved using 2 methods depending on the method of transformation. Chemically competent *E. coli* cells were prepared using a method described by Hanahan (1983). The cells were grown to mid-exponential growth phase in 20 ml TYM broth (2 % (w/v) Bacto tryptone, 0.5 % (w/v) yeast extract, 100 mM NaCl, 100 mM $MgSO_4$). This starter culture was resuspended in 100 ml TYM and grown until it reached an OD_{600} 0.5 - 0.9. This culture (100 ml) was re-suspended in 500 ml TYM broth and incubated at 37 °C until it reached OD_{600} 0.6. The culture was then rapidly chilled on ice, the cells harvested and washed using ice-cold buffers TfBI (30 mM CH_3COOK, 50 mM $MnCl_2$, 100 mM $CaCl_2$, 15 % (v/v) glycerol) and TfBII (10 mM Na-MOPS (Sigma) pH 7, 75 mM $CaCl_2$, 10 mM KCl, 15 % (v/v) glycerol). Competent cells were pelleted by centrifugation and re-suspended in 20 ml of ice cold TfBII and subsequently aliquoted (200 μl), frozen in liquid nitrogen and stored at -80 ° C.

Electro-competent *E. coli* cells for transformation by electroporation were prepared according to the method in Sambrook *et al.*, (1989). The cells were made competent by a sequence of washing steps of a mid-exponential growth phase culture

with ice cold H_2O, eventually yielding a 100-fold concentration of the original culture. The electro-competent method previously described was also used to create electro-competent *E. carotovora* sub species *atroseptica*, except the ice-cold water washes were supplemented with 10 % (v/v) glycerol.

Electroporation. Plasmid DNA was transformed into *E. coli* K-12 strains MC1061 and DH5α using a Bio-Gene Gene Pulser Electroporation instrument. 0.1 – 1 μg of DNA was added to 100 μl of competent cells and the electroporation equipment set to 2.5 V (actual), 100 Ω resistance, 25 μFD capacitance. Putative transformants were recovered by re-suspending in 1 ml of pre-warmed (37 °C) LB, for 1 hr.

Transformation of *Erwinia carotovora* sbsp. atroseptica (*ECC*) was achieved by charging the Bio-gene Gene-pulsar between 1.25 and 2.5 V (actual), 200 Ω resistance and 25 μF capacitance. The cells were recovered in 1 ml pre-warmed LB (25 °C) for 1 h, plated on selective media and incubated overnight at 25 °C. The recovered *ECC* cells (100 μl) were plated on LB containing the appropriate antibiotic selection.

Heat shock. Transformation of Plasmid DNA into chemically competent *E. coli* was achieved using the method described by Hanahan (1983). A 100 μl aliquot of competent cells was allowed to thaw on ice. Plasmid DNA (0.1 – 1 μg) was added to the cells and allowed to incubate for a further 30 min on ice. The cells were heat shocked at 37 °C for 5 min and then placed on ice for a further 5 min. The cells were recovered by 10-fold dilution in LB broth and incubation at 37 °C for 90 min on a shaking platform (120 rpm). The recovered cells were plated on LB containing the appropriate antibiotic selection.

2.5 Bacterial protein Analysis

Bacterial cultures were grown to OD_{600} 0.55 (10 ml) and the cells harvested by centrifugation (2300 x g, 5 min). The cells were then resuspended in 1 ml of ice cold PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4). The culture was subjected to sonication (13 micron amplitude, 3 x 10 s) placing the preparation on ice in between sonication steps. Protease inhibitor cocktail (Sigma) was added to enable storage potential of the preparation. The cocktail contains: AEBSF (23 mM), to inhibit serine proteases such as trypsin and chymotrypsin; EDTA (100 mM), to inhibit metalloproteases; Bestatin (2 mM), to inhibit aminopeptidases; Pepstatin A (0.3 mM), to inhibit acid proteases such as pepsin and rennin: E64 (0.3 mM), to inhibit cystein proteases such as calpain and papain.

Quantification of the total protein was achieved using the Bradford method (Ausabel *et al.*, 1992). A standard curve was prepared using bovine serum albumin (BSA) with Bradford's reagent to give a colorimetric change (A_{595}).

SDS-Page. SDS-page is one dimensional gel electrophoresis under denaturing conditions. All polyacrylamide (30 % acrylamide / 1 % bisacrylamide) gels were prepared according to guidelines in the Mini-Protean® 3 handbook (Bio-Rad) using the Laemmli (1970). In this study, 6 μg of bacterial total protein was loaded on SDS-page gel (7.5 %) and run using the Mini-Protean® 3 gel rig and running tank at 200V for 40 min. The protein gel was stained using COomassie blue (20 min) and destained for approximately 5 h, changing the de-staining buffer (10 % methanol, 10 % acetic acid in H_2O) every two hours.

Western blotting. Western analysis was used to determine the presence of bacterial outer membrane protein Vpr. 0.6 μg of total bacterial protein was loaded on a 7.5 % polyacrylamide gel electrophoresed as above. The gels were carefully removed form the Mini-Protean® 3 tank and gel rig. Prior to blotting, the PVDF membrane (Roche) was soaked in 50 % methanol (3 s), H_2O (2 min) and blotting buffer (20 mM Tris base; 150 mM glycine: 20% (v/v) methanol) (10 min). The blotting sandwich was arranged on the Electroblotter (Bio-Rad), as follows; 3 pieces of gel sized 3MM Whatman paper pre-soaked in blotting buffer; gel sized PVDF membrane; SDS-page gel; 3 pieces of pre-soaked 3MM Whatman paper. Excess buffer was wiped from the lower Electroblotter. The upper plate was put in place and the Electroblotter leads connected. The blotting current needed for efficient transfer can be calculated by; surface area of the gel multiplied by 1.6 mA; for mini-gels as used in this study the current needed was 76.8 mA. If multiple gels were to be blotted this required compensation by increasing the current. After 1 h, the PVDF was allowed to air dry at room temperature for ~ 10 min promoting uniform adsorption of proteins and stopping incompletely bound proteins from washing away. The blot was then placed in 50 % methanol for 30 s and then transferred to H_2O for 15 min. The blot was then transferred to blocking solution (5 % dried skimmed milk in 1 x TBST (20mM Tris-HCl (pH 7.6), 150 mM NaCl, 0.05% (v/v) Tween® 20)) for at least 1h at room temp or at 4 ^{o} C overnight. The blot was washed in TBST for 5min and then transferred to blocking solution containing 1 ^{o} antibody (for example polyclonal antibody anti-rabbit anti-Vpr 1:50,000 dilution of stock sera) under gentle agitation for 2 h at room temperature or 4 ^{o} C overnight. The blot was washed twice (10 min) in TBST and incubated with the 2 ^{o} antibody (alkaline phosphatase conjugated anti-rabbit IgG whole molecule (Sigma) 1:30,000 dilution) 1:30,000 dilution in blocking solution or

TBST. The blot was washed twice (10 min) in TBST and equilibrated in alkaline phosphatase buffer for 5 min.

Detection. There are two methods of detection to visualise the western blot, the first protocol is chemical development. After equilibrating the membrane in alkaline phosphatase buffer as above, the blot was incubated in 10 ml of alkaline phosphatase buffer with 15 µl of 10 mg ml^{-1} BCIP in dimethyl formamide (DMF). The blot was then stored in the dark until the background developed a pink colouration, or, bands had stained to the required degree. The reaction was stopped by the addition of 2 ml each of 1 M Tris-HCl (pH 7.5) and 0.5 M EDTA (pH8).

The second method is a chemilluminescent approach. Detection of the blot was achieved by adding a chemilluminescent substrate (CPD* or CSPD (Roche)); these components were used as per The DIG Systems User guide (Roche) used for detection of Southern hybridisation. The chemi-luminescent signal was visualised by exposing the membrane to imaging film (Biomax light-1, Kodak) and the film developed according to manufacturers' guidelines.

2.6 Scanning Electron Microscopy (SEM)

Samples for SEM analysis were prepared by Dr. Caes Veltkamp (School of Biological Sciences, The University of Liverpool). Samples were dehydrated by adding excess (5 ml) ice cold (- 18 °C) absolute ethanol to the sample followed by incubation at – 18 °C (1 – 2 h), then transfer to 4 °C (overnight). As the sample had sunk to the bottom of the ethanol it was possible to replace the top 2 -3 ml of ethanol to make sure complete dehydration had occurred. It was then incubated at -4 °C for a further 4 h. Samples were dried in carbon dioxide using a Polartron E3000 critical

point drier, attached to stubs and sputter coated with 60 % gold-palladium in a Polaron E5100 coater. The SEM images were produced using a Philips 501B scanning electron microscope at accelerating voltages of 7.2 and 15 kV (Veltkamp *et al.* 1994).

2.7 Laser scanning confocal microscopy

E. coli strain MC1061 cells were grown to OD_{600} 0.55, 0.5 ml of the cells were washed three times with PBS, and resuspended in 0.5 ml of PBS. To the washed cells a 1: 25,000 dilution of the primary anti-Vpr antibody was added to the cells, followed by the addition of the secondary antibody (anti-rabbit IgG conjugated with Alexafluor 433 (Molecular probes) in a 1 : 25, 000 dilution).

Samples were visualised by Dr. David Spiller, Centre for Cell Imaging, University of Liverpool, using a Carl-Zeiss LSM510, laser scanning confocal microscope.

2.8 Sequence analysis

Identification. Nucleotide and protein sequences were analysed using the BLAST facility of the National Centre for Biotechnology Information (NCBI) bioinformatics webpage (Altschul *et al.* 1997).

Alignment. Protein and nucleotide alignment were achieved using ARB (Ludwig *et al.* 2004) and ClustalX.

Phylogenetic Analysis. Protein sequence alignments derived from both ClustalX and ARB were used to infer parsimonius phylogenetic analysis on the bacterial receptor for $\Phi 24_B$ infection. Homologues of Vpr were identified by blastp analysis

of the complete amino acid sequence of the protein. Homologues identifying greater than 50 % (Fig 4.2) or 30 % (Appendix 3) amino acid sequence identity were aligned using ARB. Parsimony analysis of the alignment was achieved using the protpars algorithm of ARB weighted by bootstrapped values derived from 100 analyses.

Parsimony analysis is a good tool as it infers phylogeny by analysing the shortest evolutionary distance between the sequence it is analysing, forming nodes and clades which form the phylogenetic tree. Parsimonius trees may not always infer true phylogeny as the algorithm depicts the least evolutionary distance between numbers of amino acid changes between sequences. Bootstrapping using parsimonius analyses resamples the alignment to derive alternative analyses, which are finally compounded into a consensus cladogram or phylogenetic tree. Frequency of occurrence of groups/clades adds weight to each individual node or clade offering a figure for this weighting.

Because the parsimony algorithm is based upon identifying amino acid differences between the sequences in question, it is possible to analyse sequences of different lengths that harbour varying levels of sequence diversity. It is for this reason why parsimony analysis was used to form representative cladograms for the host recognition proteins of both short and long tailed lambdoid-like bacteriophage. 933W / λ related tail spikes were identified using blastp analysis. The amino acid sequences derived were aligned using ARB and parsimony analysis achieved using the protpars facility in ARB weighted by bootstrapping of 100 analyses. The cladograms that were produced are a representative of the amino acid difference between the bacteriophage tail spikes idenitified. These cladograms cannot infer phylogeny as they are derived from sequences of different sizes.

Chapter 3: Electron Microscopy of bacteriophage induced from *E. coli*

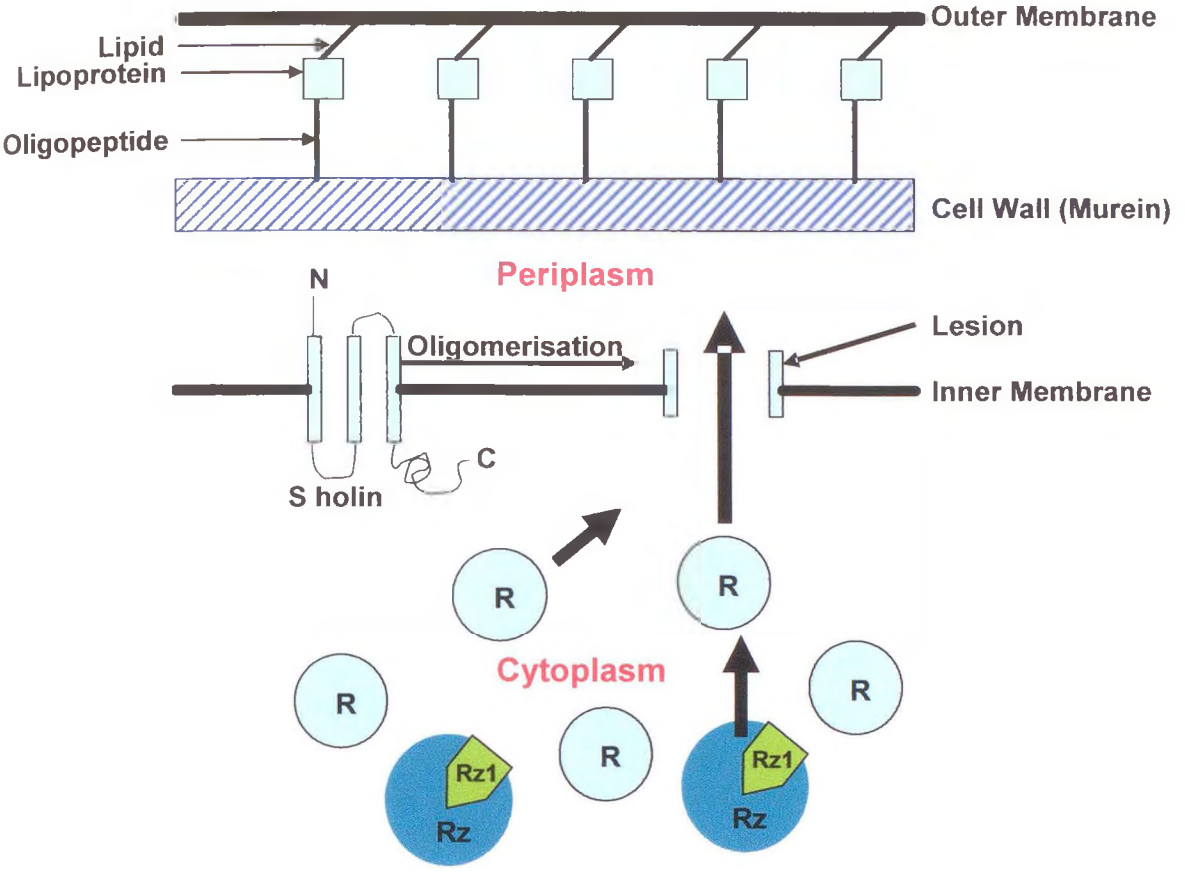
3.1 Introduction

The initiation of the lytic life cycle of lambdoid phages can be achieved by stimulating the bacterial SOS response via DNA damage, resulting in auto-cleavage of the phage repressor protein and transcription and translation of the phage late genes. A range of mutagens can be used to induce the bacterial SOS response, including Norfloxacin a member of the quinolone family of anti-microbial compounds (Matsushiro *et al.*, 1999), UV irradiation (Lwoff *et al.*, 1950) and Mitomycin C (Peterson *et al.*, 1981). High hydrostatic pressure has also been shown to induce Stx-phage (Aertsen *et al.*, 2005).

In order for phages to enter the environment, coliphage such as λ and Stx-phages must traverse both the bacterial cytoplasmic membrane and the mesh-like structure of the peptidoglycan layer. Double stranded DNA phages produce endolysin, a soluble muralytic enzyme (Young *et. al.,* 2000). These phage-encoded endolysins are so named due to their ability to degrade bacterial cell walls. Endolysins target the murein layer in the cell wall, either by transportation through the inner membrane or by accumulation in the cytoplasm, with access via lesions to the inner membrane (Bernhardt *et al.*, 2002).

Figure 3.1 shows how λ drives lysis of the cell using this dual enzymatic approach and this is much more complex than the method used by ssRNA or small DNA (> 6 Kbp) bacteriophage which adopt a single protein strategy *e.g.* ΦX174 (*Microviridae*) and MS2 (*Leviviridae*) (Bernhardt *et al.*, 2002).

Figure 3.1 The Holin-Endolysin Lysis System of Bacteriophage λ



The R endolysin accumulates in the cytoplasm and the S holin protein accumulates and oligomerises in the membrane. At a genetically programmed point during the lytic cycle, the S holin suddenly forms lesions in the inner membrane. These disruptions allow the R endolysin to diffuse into the periplasm. Through this lesion and the subsequent damage to the murein layer, λ proteins Rz and Rz1 are able to attack the oligopeptide links to the outer membrane, ultimately leading to cell lysis. Produced from Bernhardt *et al.* (2002) and Young *et al.* (2000).

The lysogen cell is ruptured to release the mature phage particles in a regulated manner as it would not be profitable for the phage to lyse the cell to expose partially constructed immature virions (Bernhardt *et al.*, 2002). It is possible that this timing is determined by levels of the key components of active holin production. The λ S holin protein (S^{λ}) actually encodes 2 proteins S^{λ} 105 and S^{λ} 107, via translational alterations. Under standard growth conditions, the ratio of S^{λ} 105 to S^{λ} 107 is 2:1 respectively. Increase in S^{λ} 107 leads to the formation of the 3 domain transmembranal protein in (Fig 3.1) which is an active holin protein (Young *et al.*, 2000).

In this chapter, the lytic burst is visualised and described in relation to the growth cycle of the $\Phi 24_B$ lysogen of *E. coli* strain MC1061. Induction of lytic life cycle is achieved by adding the antibiotic norfloxacin (see 2.2.4). Briefly, the norfloxacin (NFLX) stimulates the bacterial SOS response, which in turn activates the auto-proteolytic cleavage of cI and subsequent anti-termination and transcription of the late genes (see 1.2). The $\Phi 24_B$ lysogen was grown to OD_{600} 0.5 (~ 3 x 10^{8} cfu ml^{-1}) before induction, which is equivalent to the mid-exponential growth phase of the lysogen culture (Fig 3.1A)

3.2 Results

Fig 3.2 images a – d give a representational view of the healthy lysogen cells at mid-exponential growth phase. High proportions of the cells are in the process of septation and cell division. As the cells are actively dividing, the cell size varies between 1 and 2 μm in length.

Fig 3.3 images a - d show lysogen cells that have been exposed to the DNA gyrase inhibiting antibiotic NFLX (1 μg ml^{-1}). The cells have deep indentations in

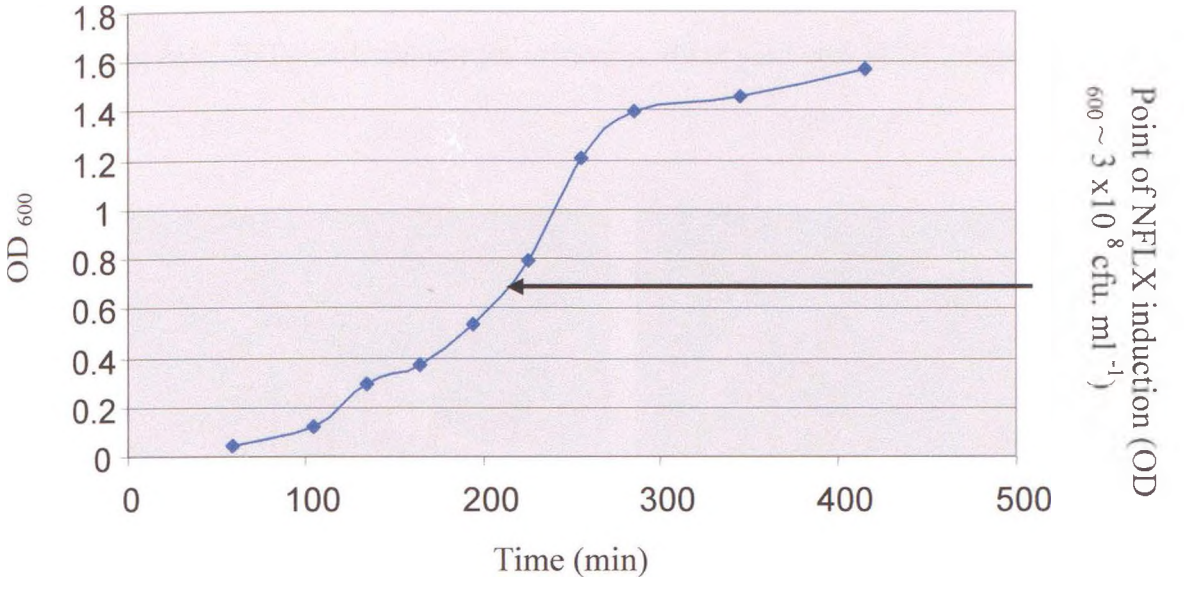


Fig 3.1A $\Phi 24_B$ Lysogen of *E. coli* strain MC1061 Growth curve. Lysogen cultures were induced when the cell optical density reached OD_{600} 0.5, which correlates to approximately 3 x 10^{8} lysogen cells ml^{-1} and is located at mid-exponential growth phase of the lysogen growth cycle. Sample data points are the mean of duplicate samples.

Figure 3.2 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogens harvested at Mid-Exponential Growth Phase 0.55 (OD_{600})

1 ml of cells was washed twice in ice cold H_2O and resuspended finally in 5 ml of ice cold 200 proof ethanol for critical point drying and SEM imaging.

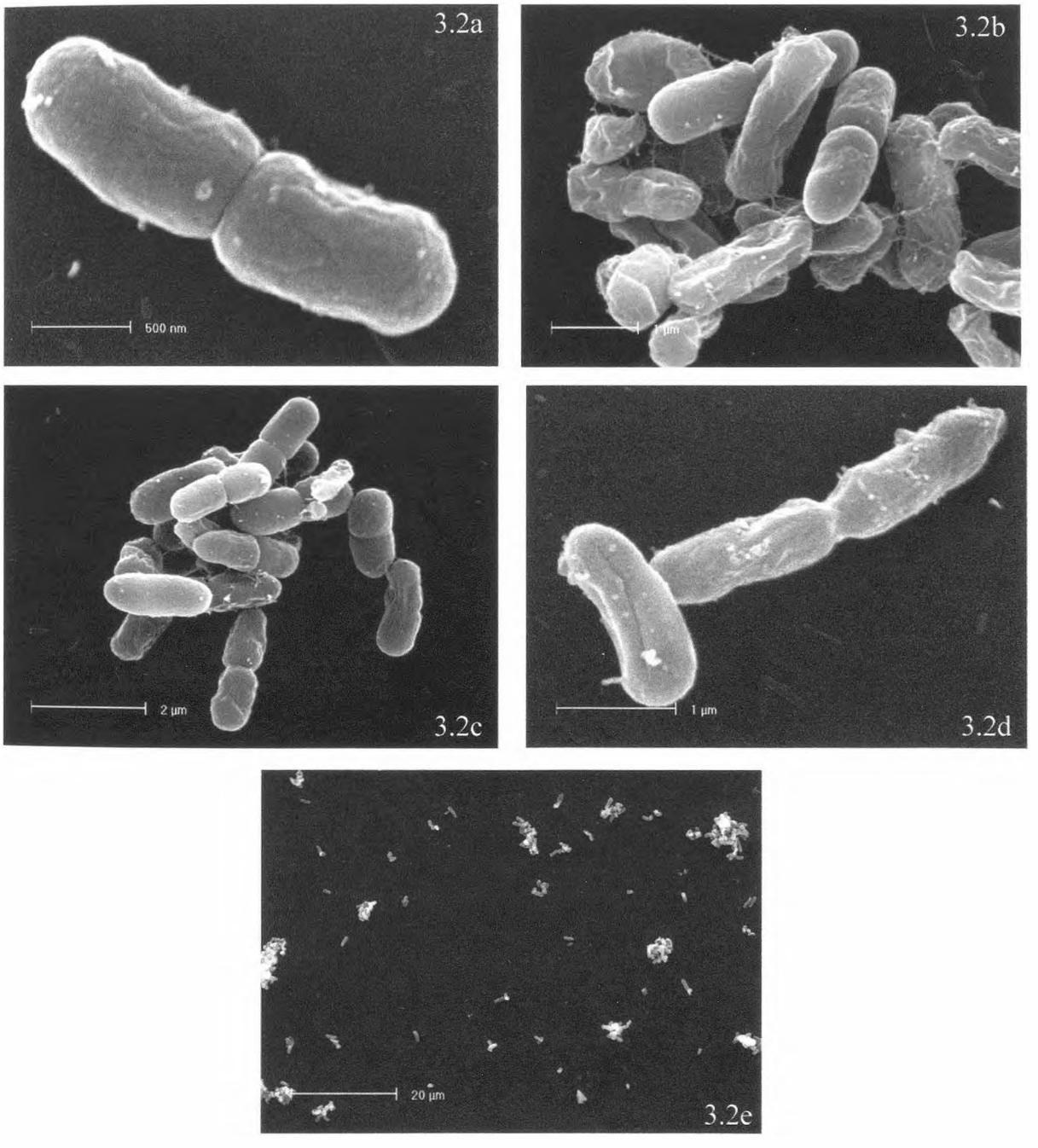


Figure 3.3 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen after 1 hr incubation with 1μg.ml^{-1} NFLX

1 ml of cells was washed twice in ice cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging.

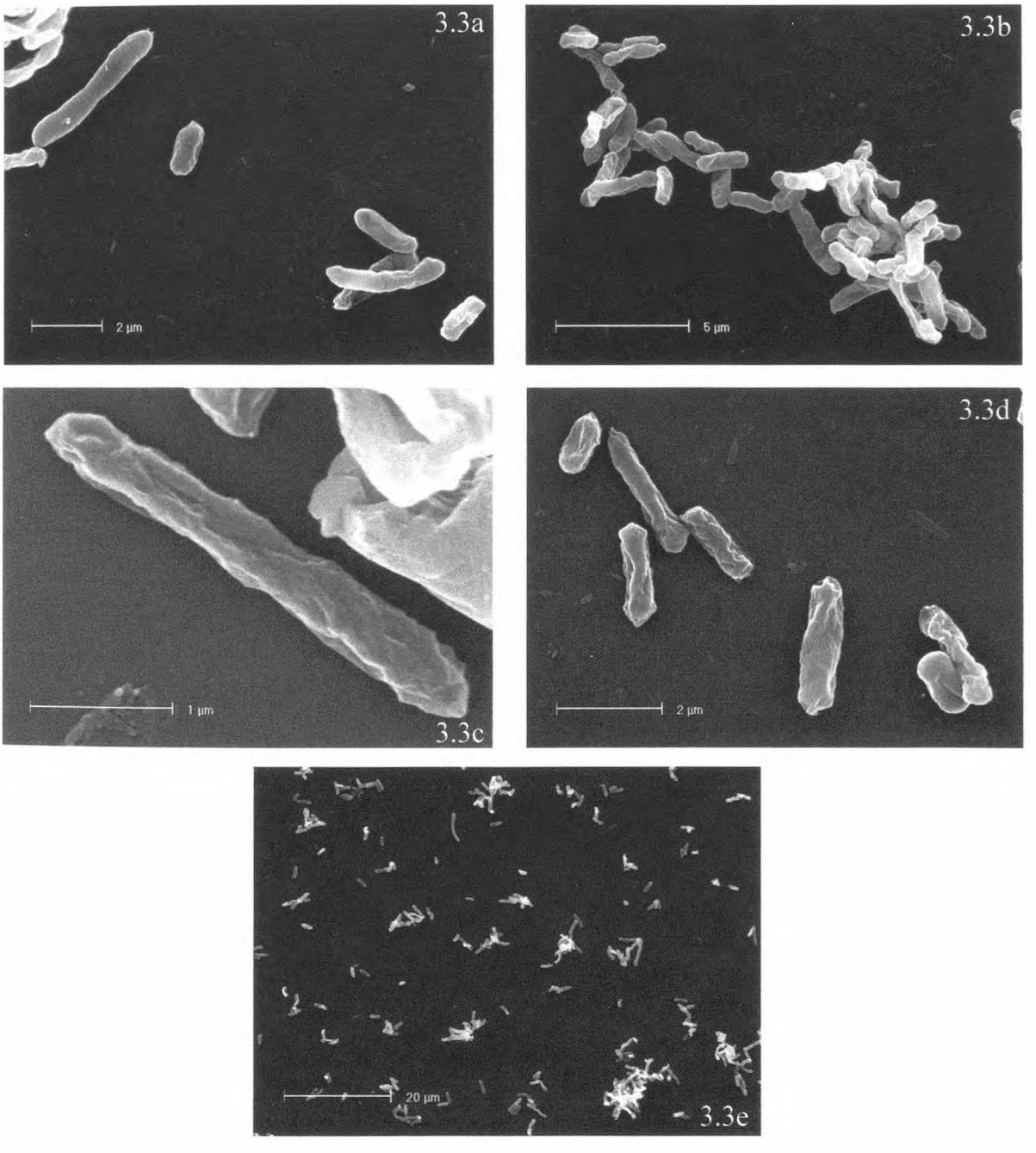
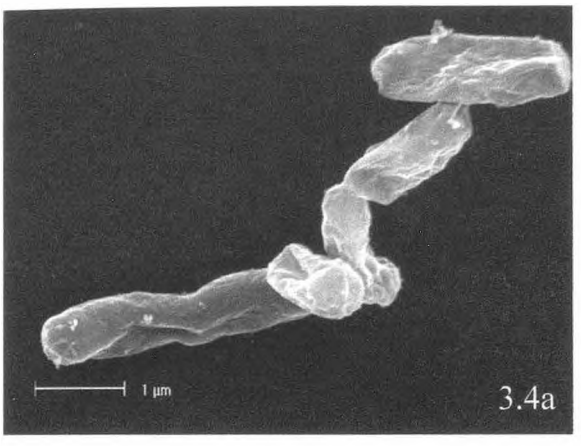
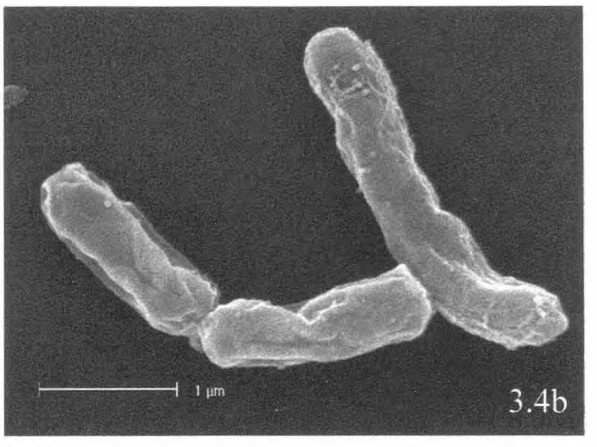
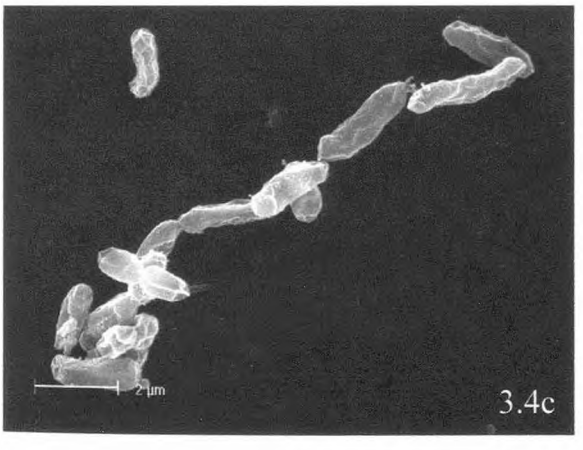


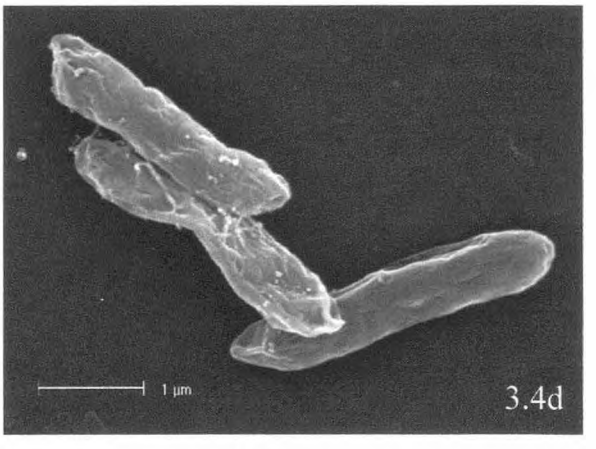
Figure 3.4 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen, 15 min after dilution of NFLX.

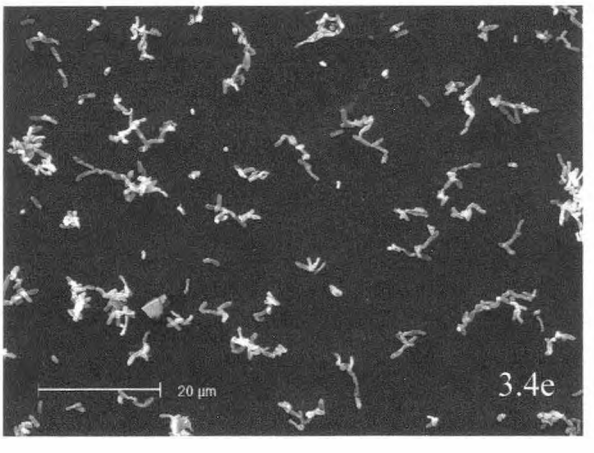
1 ml of cells was washed twice in ice cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging.











the cell surface (Fig 3.3c and 3.3d) with a dehydrated appearance. The cells varied in size, but were often more elongated than the control cells (Fig 3.2), in the absence of NFLX.

Fifteen minutes after the dilution of NFLX, the lysogen cells still had a collapsed appearance and were often elongated (Fig 3.4). The lysogen cells observed after 30 min of recovery had lost their collapsed appearance, but cell division did not appear to be occurring as elongated cells (up to *ca* 4 μm) only predominated (Fig 3.5 a – e).

After 45 min of recovery, some of the lysogen cells were still elongated, but there was also some form of off-centre septation at the apices of some cells (encircled with white rings in Fig 3.6c).

At time points 60 and 75 min into the recovery period (Fig 3.7a – b and 3.8a respectively), the lysogen cells were still elongated and similar in appearance to the cells at the 45 min time point. The effect of norfloxacin, as revealed by electron microscopy, therefore appears to be a collapsing of the cell, followed by inhibition of cell division. Cells seem elongated, but have recovered a more healthy appearance ca 30 – 45 min after recovery by dilution of the NFLX.

After 90 min of recovery it was possible to identify some cells either at the beginning or at the end point of phage-mediated lysis. Figure 3.9a shows an extremely elongated cell of approx 20 μm in length; Fig 3.9b places that elongated cell within the field of view previously used to provide a representation of overall cell length. Figure 3.9c is another view showing an elongated cell of ~ 20 μm. Figure 3.9d shows an almost flat cell at either the beginning or the end of phage mediated lysis. Fig 3.9f shows increased magnification of the lysing cell that can be seen in Fig 3.9d and e, showing a phage and intracellular material being externalised.

Fig 3.9g shows an either off-centre budding effect seen in Fig 3.6c, or possibly the initial step in phage-mediated cell lysis.

After 105 minutes of recovery, Fig 3.10 a-e show cells in various stages of phage-mediated lysis. Fig. 3.10c shows lysis of a cell where there seems to be a large plume-like effect of intracellular matter. In many cells, released phage and intracellular material may have been dislodged during the centrifugation and washing steps required for SEM preparation.

After 120 min of recovery, Fig 3.11 a and b show more examples of lysing cells mixed with what seem to be actively growing and dividing cells, more similar to those observed prior to NFLX induction. There also seems to be a lot of small particles bound to the cell surfaces; it can be speculated that these may be viral particles binding to the cell surface, as at this time point it was possible to recover 3 x 10^{8} pfu ml^{-1} from the sample prior to the absolute ethanol dehydration step.

Figure 3.5 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen – 30 min after dilution of NFLX (recovery period)

1 ml of cells were washed twice in ice-cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging

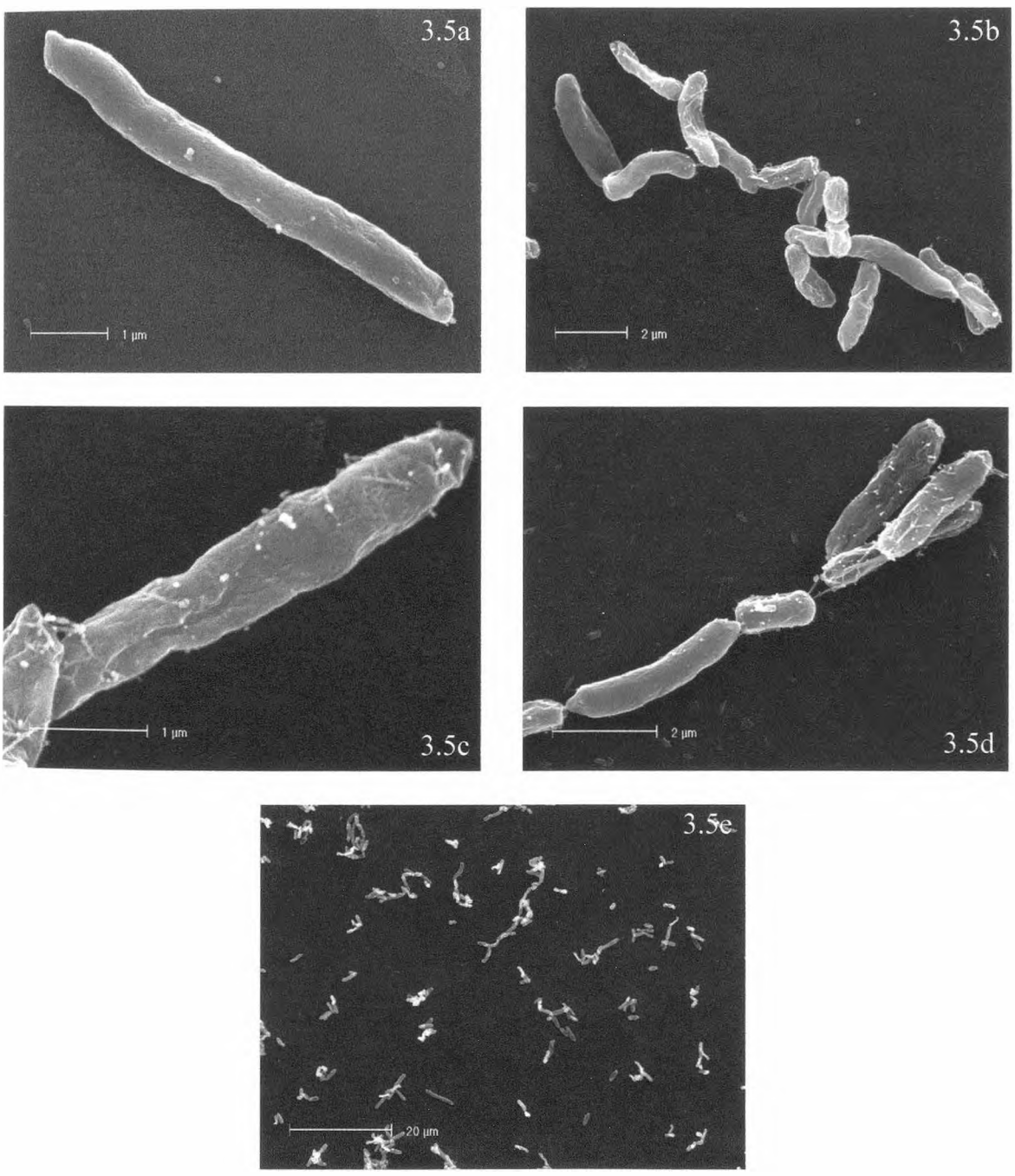


Figure 3.6 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen – 45 min after dilution of NFLX (recovery period)

1 ml of cells were washed twice in ice-cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging. White rings in Fig 3.6c highlight the occurrence of off-centre septation in some lysogen cells.

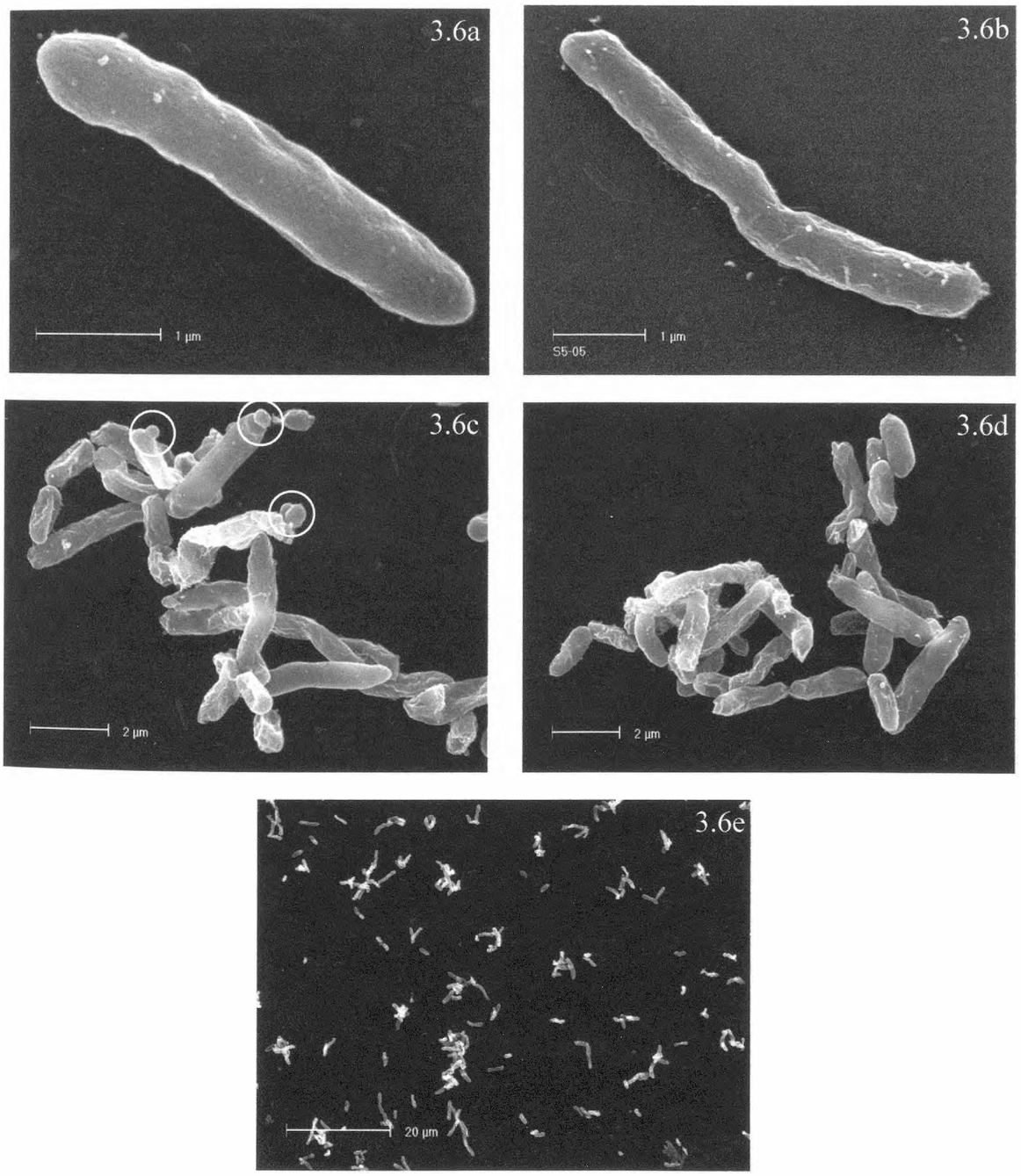
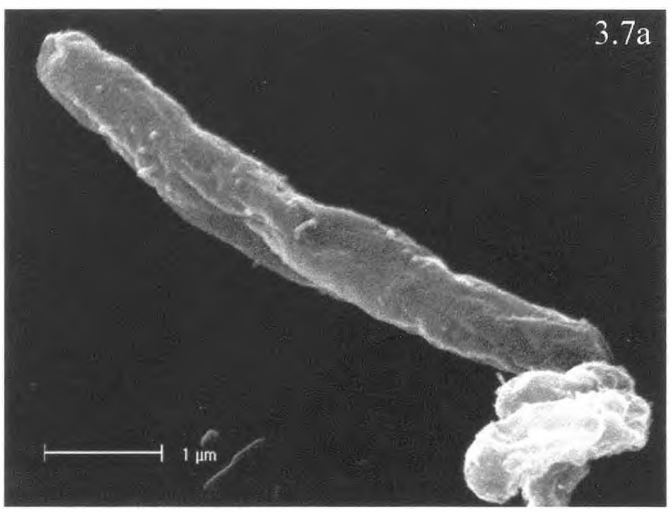


Figure 3.7 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen – 60 min after dilution of NFLX (recovery period)

1 ml of cells were washed twice in ice cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging



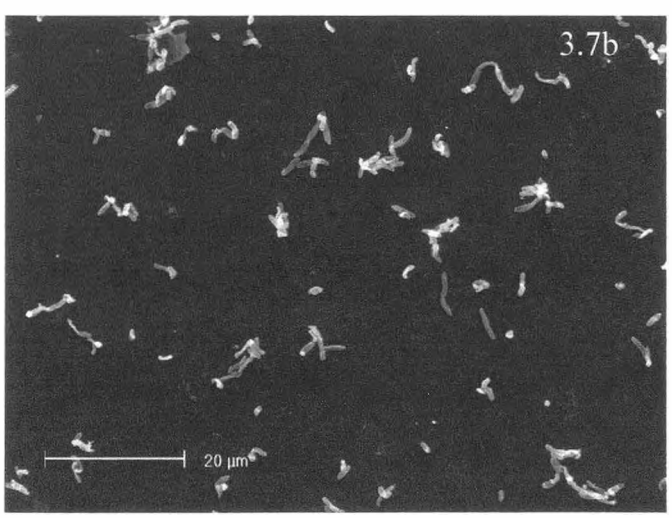
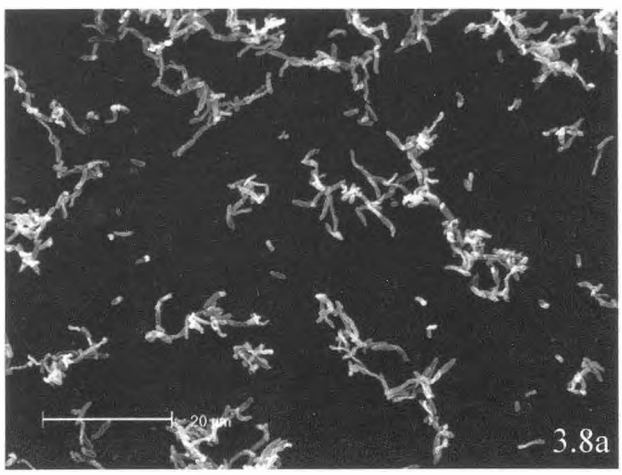


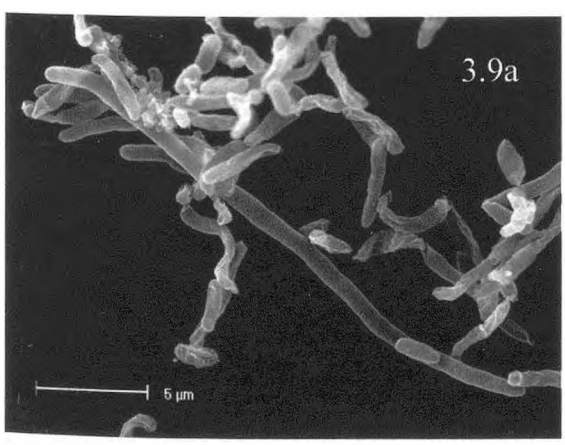
Figure 3.8 SEM Analysis of Φ24$_B$::Δ*cat* MC1061 lysogen – 75 min after dilution of NFLX (recovery period)

1 ml of cells were washed twice in ice-cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging



3.9 SEM Analysis of Φ24$_B$::Δ*cat* MC1061 Lysogen – 90 min after dilution of NFLX (recovery period)

1 ml of cells were washed twice in ice-cold H_2O and resuspended finally in 5 ml of ice-cold 200 proof ethanol for critical point drying and SEM imaging



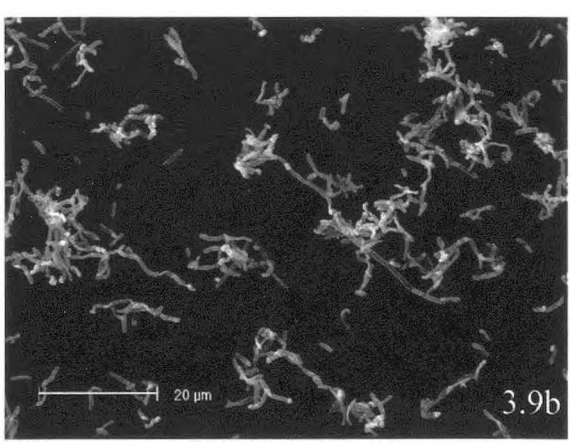
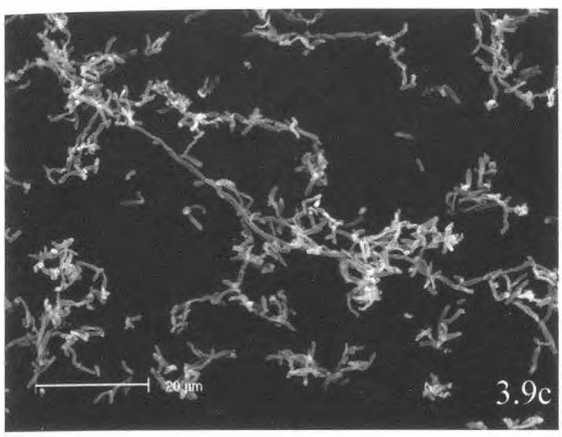
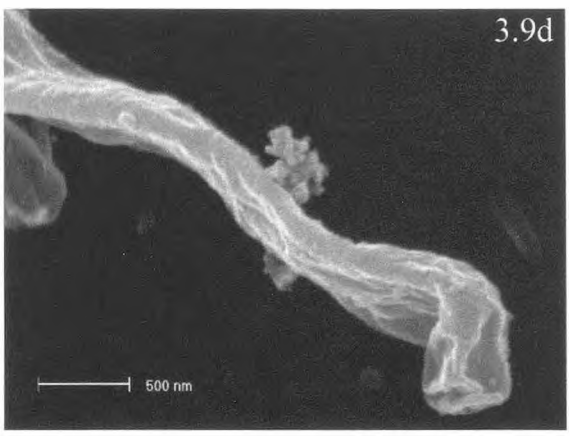
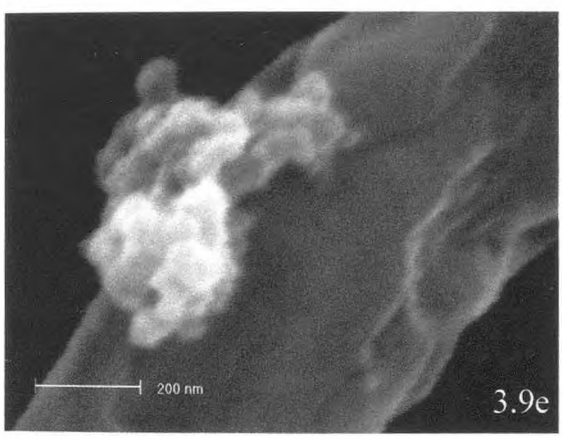
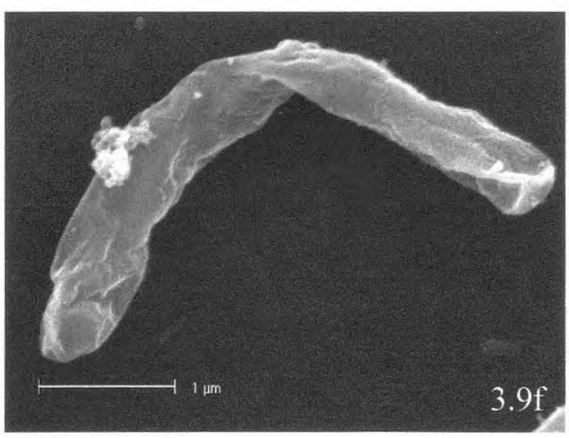


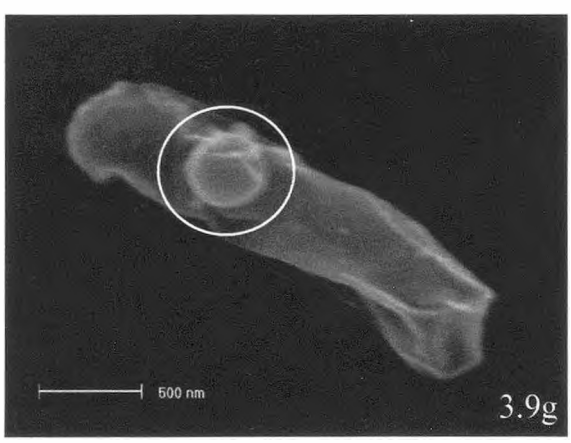
Figure 3.9 contd.



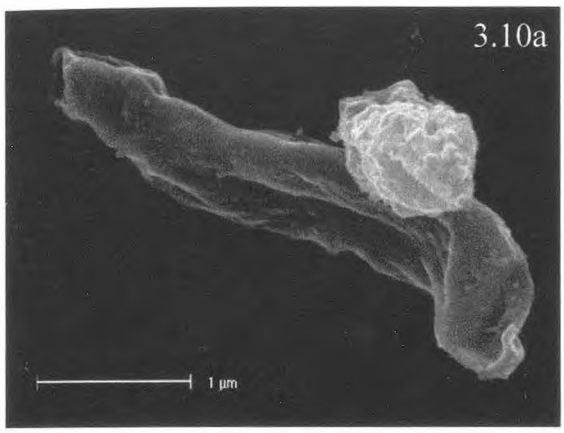


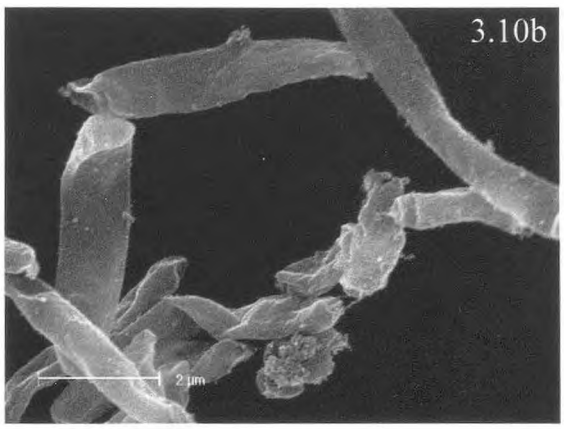


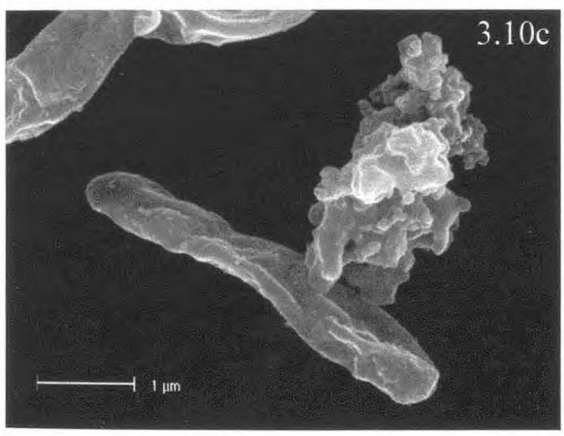


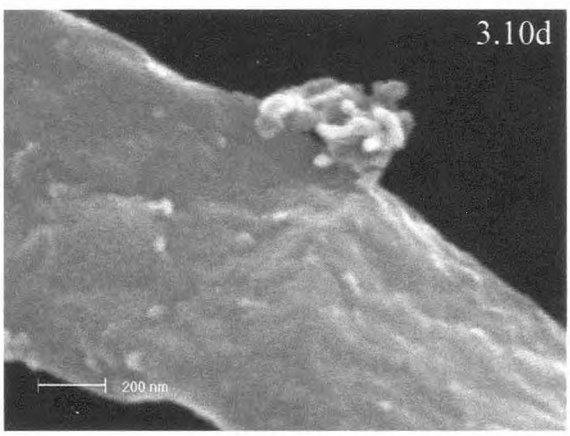


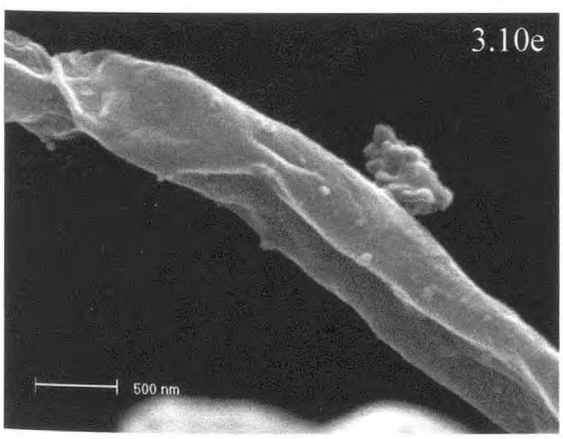
3.10 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen – 105 min after dilution of NFLX (recovery period)











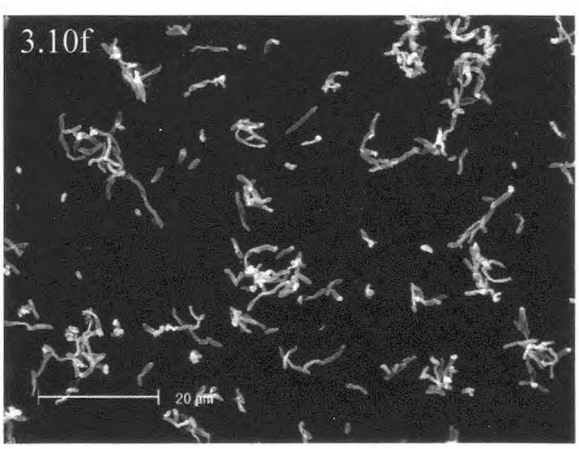
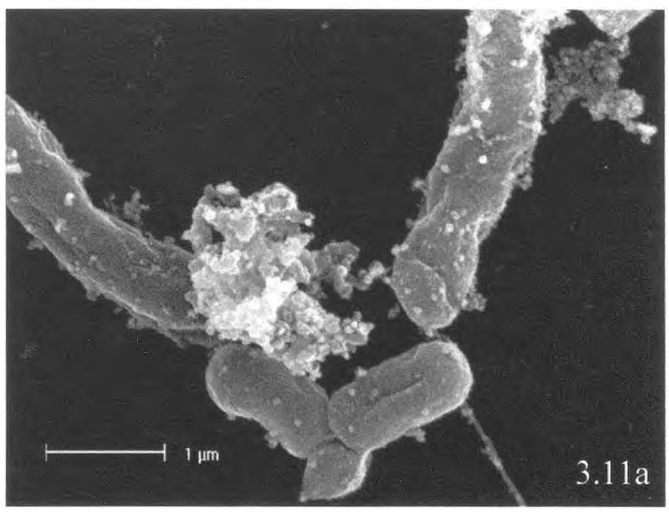


Figure 3.11 SEM Analysis of $\Phi 24_B$::Δ*cat* MC1061 lysogen – 120 min after dilution of NFLX (recovery period)



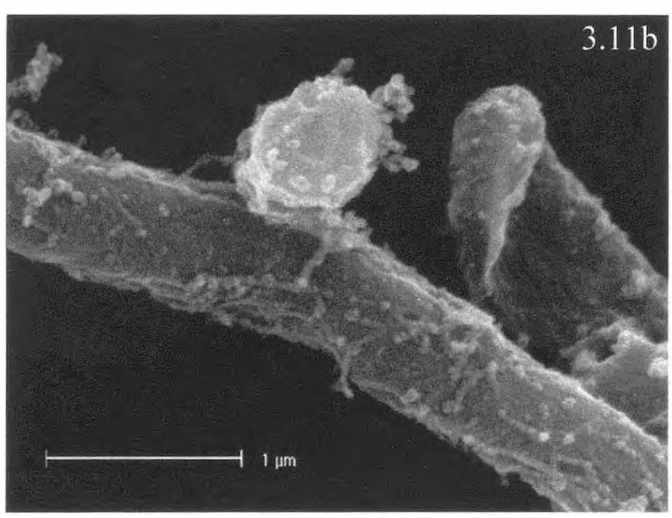
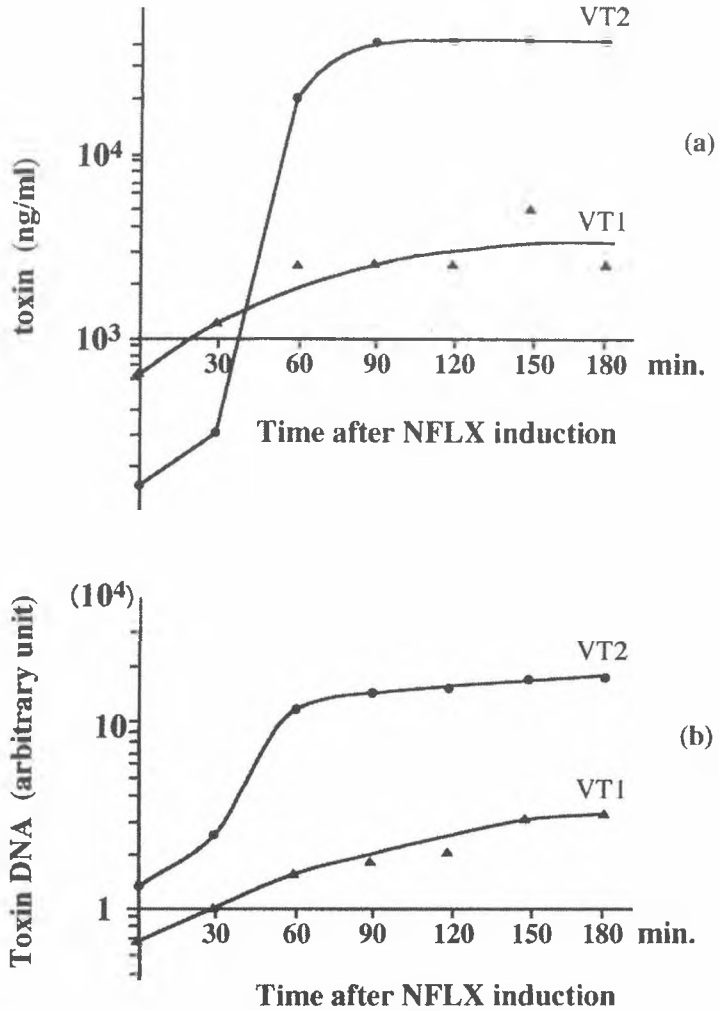


Figure 3.13 Levels of Stx Toxin production by *E. coli* under induction using norfloxacin



Time course experiments using EHEC O157 strain RIMD0509893. Panels: a, toxins VT1 and VT2; b, VT1- and VT2-specific DNAs. The VT1 and VT2 toxin DNA assay method used is as follows. After induction for 30 min with 1 μg/ml NFLX, 1 μl aliquot of a culture was spotted onto each of two nitrocellulose filters (Nitroplus 2000; Micron Separations Inc., Westborough, Mass.) every 30 min. One filter was hybridized with fluorescein-labeled fragments containing VT1 gene DNA (100 ng), and the other was hybridized with VT2 gene DNA in 10 ml of 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-5× Denhardt's solution-0.5% sodium dodecyl sulfate-4 mM EDTA-100-μg/ml salmon sperm DNA. After being washed in 0.5× SSC at 60°C, the filter was incubated with an anti-fluorescein-alkaline phosphatase conjugate and visualized with the fluorescence substrate Atto Phos (Amersham Pharmacia Biotech, Ltd., Uppsala, Sweden). The fluorescence signals were detected with a FluorImager (Molecular Dynamics Inc., Sunnyvale, Calif.). The amount of VT1- and VT2-specific DNA was expressed in arbitrary fluorescence intensity units as measured at 540 to 560 nm (excitation at 488 nm).

From Matsushiro *et al.*, (1999)

3.3 Previous SEM image analysis of $\Phi 24_B$ induction

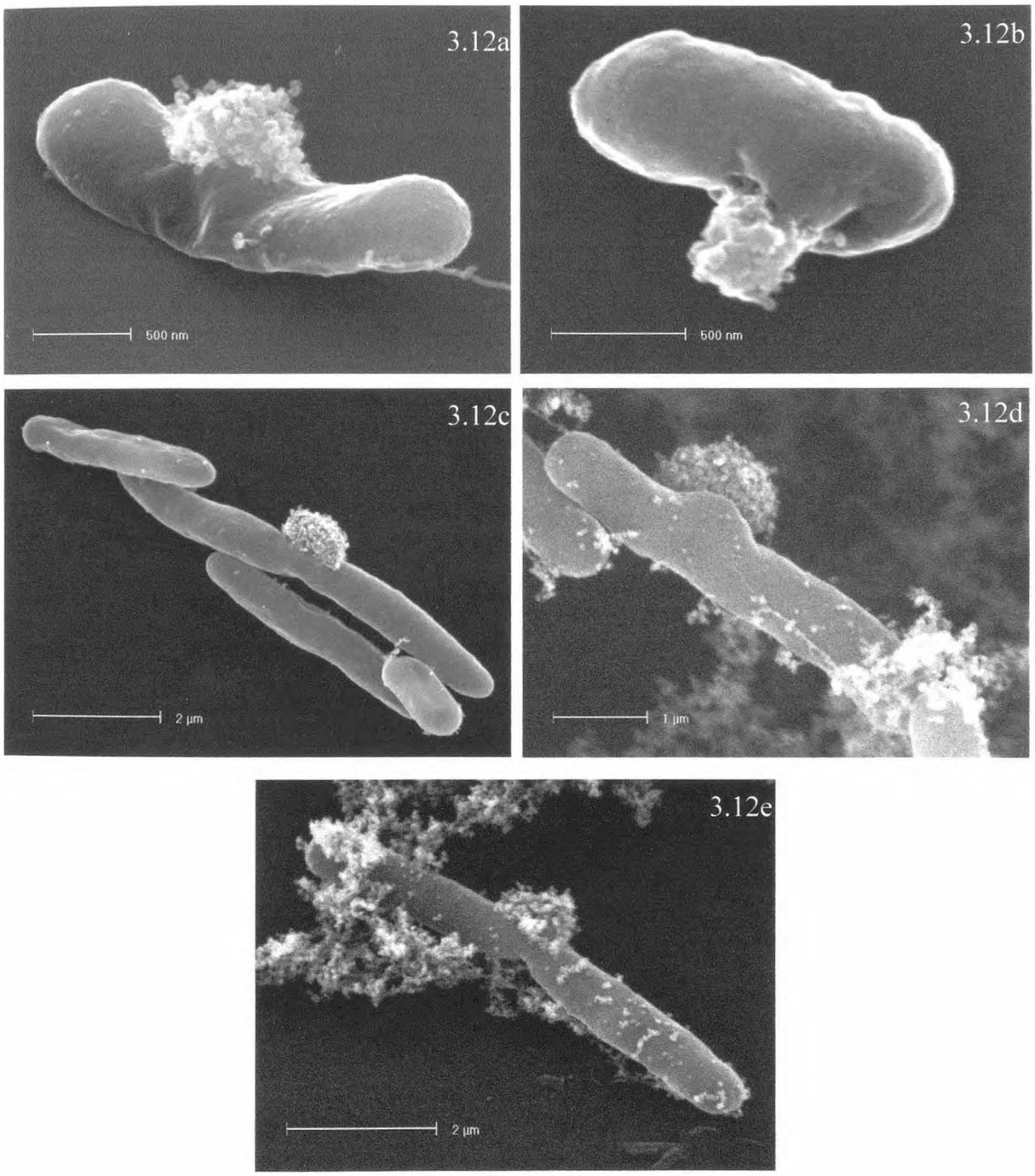
Fig 3.12 shows a series of images of *E. coli* MC1061 releasing Stx-phages and were produced in collaboration with Chloe James (2002) and an undergraduate research student (Donna Tagg). They are presented here as they show particularly good images of bacteriophages released from *E. coli* after induction with norfloxacin.

3.4 Discussion

NFLX inhibits the A subunit of DNA gyrase (Hirai *et al.*, 1986) arresting DNA synthesis and permitting accumulation of single-stranded DNA, which stimulates the RecA-mediated SOS-response (Craig *et al.*, 1980). RecA cooperatively binds to single-stranded DNA and then pairs it to a homologous DNA complex (Courcelle *et al.*, 2003) ready for the DNA repair process to begin. Fig 3.3 shows that NFLX treatment has a profound effect on the appearance of cells after one hour exposure. The cells appear almost flattened / dehydrated with heavy ridging across the cell surface. As the recovery period proceeds elongated cells sometimes over 20 μm in length, are common. This inability of the cell to divide upon interference with DNA synthesis has been described previously (Mount, 1975). During stimulation of the SOS-response, the single-stranded DNA-bound RecA acts as a stimulus for the upregulation of over 40 genes (Courcelle *et* al., 2001) with the aim to stabilise the DNA replication fork, repair lesions and carry out translesion synthesis. Immature cell division is also prevented (Courcelle *et al.*, 2003). This filamentous growth stimulated by DNA-replication arrest was first identified by Cohen and Barber (1954) and further characterised by Mount (1975). As the time course moves past the 30 min recovery period, the cells begin to recover their normal

Figure 3.12 Image of phage release from MC1061 by $\Phi 24_B$

Images 3.2 a - e display a range of cells releasing bacteriophage after norfloxacin induction.



appearance and after ca. 120 min the cells are approximately the correct size for an actively growing *E. coli*. The titre of free phages increases upon induction with NFLX until ca. 120 min when numbers slightly decrease due to adsorption to lysed cellular debris (Matsuhiro *et al.,* 1999). Identification of increased bacteriophage lysis correlates with the increased levels of Stx toxins detected by of Matsushiro *et al.*, (1999) and reproduced in Fig 3.13. Thus, induction by NFLX results in the entry of bacteriophages into the lytic life cycle and subsequently increases levels of Stx expression, which is as expected due to the location of *stx* is in the "late gene" region of the bacteriophage genome.

The preliminary SEM work of James (2002) suggested that the point of lysis was at a distinct position on the cell surface, located in the region of the cell that undergoes septation. The SEM data presented here, using the same phage/host system, show that lysis can occur at a number of points at the cell surface (Fig 3.12d, 3.10a, 3.10e). Phage-mediated lysis occurs via the phage-encoded holin proteins, which are regulated to lyse the cell to yield maximum viable virion release (section 3.1). It has not been determined previously if there are specific sites at the cell surface where lysis occurs. The point of septation could be a weak point at the cell surface where lysis therefore occurs preferentially and then could be the explanation for the observations of James (2002). It may be that holin poration events are occurring throughout the cell envelope and that the point of lysis is simply the point at which the first hole-forming event occurs first.

4.0 Identification and characterisation of Vpr

4.1 Background: Initial identification of the $\Phi 24_B$ Receptor

In a previous study, Sergeant (1998) putatively identified a gene encoding an outer membrane protein associated with a short tailed phage, $\Phi 24_B$, infection of *E. coli* using a strategy previously described by Kiino and Rothman-Denes (1989), in which introduction of the wild-type gene into the resistant bacterial host restores susceptibility to bacteriophage infection. Initially the resistant mutant was isolated by subjecting the *E. coli* K-12 host strain (MC1061) to infection with bacteriophage ΦE8664-Stx2. Subsequent bacterial colonies were mainly lysogens, but some were found to be resistant to Stx-phage infection. To confirm that these actually lacked the receptor required for Stx-phage infection, Sergeant (1998) then bombarded the resistant hosts with a cocktail of Stx-phage comprising ΦE86654-Stx1, ΦE86654-Stx2, ΦE85539-Stx2a, ΦE83819-Stx1 and ΦD155-Stx1. Phage adsorption assays on surviving bacterial colonies confirmed that the target receptor was absent for $\Phi 24_B$ to adsorb to and thus infect.

Fig 4.1 follows the progression through creation of the $\Phi 24_B$-resistant MC1061 to complementation with *Sau3A* I partially restricted *wt* genome carried on pUC18/pUC19. Selection for transformants was achieved using the plasmid-borne ampicillin resistance marker, and they were then subjected to infection with $\Phi 24_B$::Kan. Colonies that had grown under the kanamycin and ampicillin selective pressure were deemed to be putative lysogens. They were confirmed as lysogens if it was possible to induce bacteriophage from them. The lysogens therefore contained both the phage and plasmid harbouring the bacterial receptor for $\Phi 24_B$ infection.

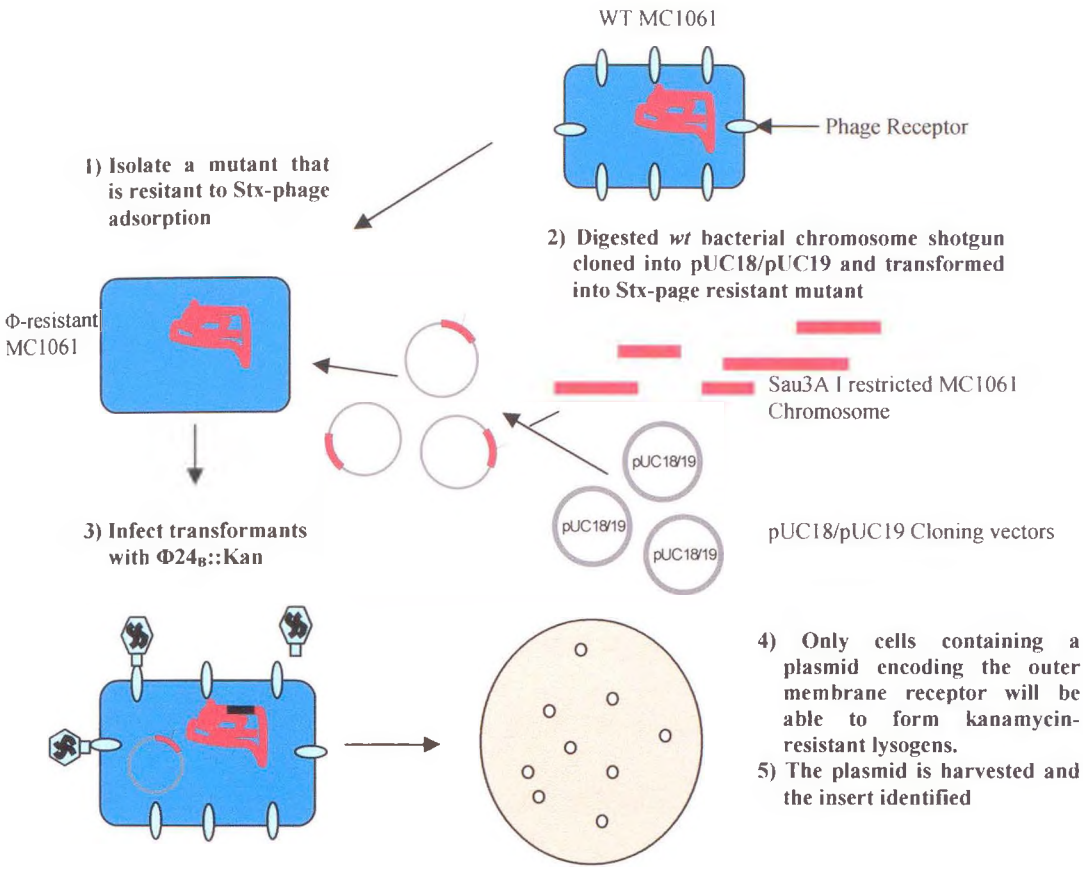


Fig 4.1 Identification of $\Phi 24_B$ outer membrane receptor by complementation. 1 and 2). The phage-resistant mutant was transformed with a shotgun cloned library of the *wt* genome (MC1061) on cloning vectors pUC18 / pUC19. Transformants were selected using pUC vector-borne ampicillin resistance. **3)** Transformants were subjected to infection with $\Phi 24_B$::Kan. **4)** Lysogens were selected on LB with ampicillin (100 μg ml $^{-1}$) and kanamycin (50 μg ml $^{-1}$), incubated ~ 18 hr, 37 ^{o} C. **5)** Vector from ampicillin / kanamycin resistant lysogens was harvested and the insert sequenced to identify the gene that restored phage infection sensitivity; this gene was designated *vpr*.

Two clones (pUC19ΦR1 (4.5 Kb insert) and pUC18ΦR2 (1.6 Kb insert)), were identified that conferred phage sensitivity to the Φ-resistant MC1061. Sub-cloning of pUC19ΦR1 insert digested with *Bt* I (2.5 Kb) yielded a plasmid pUC19ΦR1D that fully complemented the Φ-resistant mutant. Nucleotide sequencing and alignment of this fragment identified an 810 aa protein designated Vpr (**V**erocytotoxin **P**hage **R**eceptor). This *vpr* gene is located at minute 4.4 on the *E. coli* chromosome.

4.2 Background: Verocytotoxin Phage Receptor (Vpr)

Subsequent to Sergeant (1998), v*pr* homologues have been identified in all of the *Enterobacteriaceae* sequenced thus far (James *et al.*, 2002). *vpr* has been identified as an essential gene in *E. coli* (Gerdes *et* al., 2003); Dartiganlongue *et al.* (2001) had previously described a viable knockout mutant SR4455 (MC4100 EcfK::Ωkan), through personal correspondence with Dartigalongue revealed that *vpr* had been complemented on a plasmid vector prior to the subsequent knockout of the chromosomal copy. Thus *vpr* is an essential gene for *E. coli*. Homologues of *vpr* in sequenced *E. coli* strains have been reported under many different names *ecfK* (Dartigalongue *et al.*, 2001), *yaeT* (Blattner *et al.*, 1997). All *E. coli* genomes sequenced thus far reported identical *vpr*-encoding genes. A phylogenetic tree (Fig 4.2) has been derived using parsimony analysis and bootstrapped on a hundred analyses. The tree shows how related the orthologues of Vpr are and this can also be seen in the protein sequence alignment of Vpr homologues (Appendix 1). In further analyses, Vpr orthologues were identified by blastp searching against the Vpr protein sequence and all proteins possessing > 50 % protein sequence identity were included in both the alignment and phylogenetic tree.

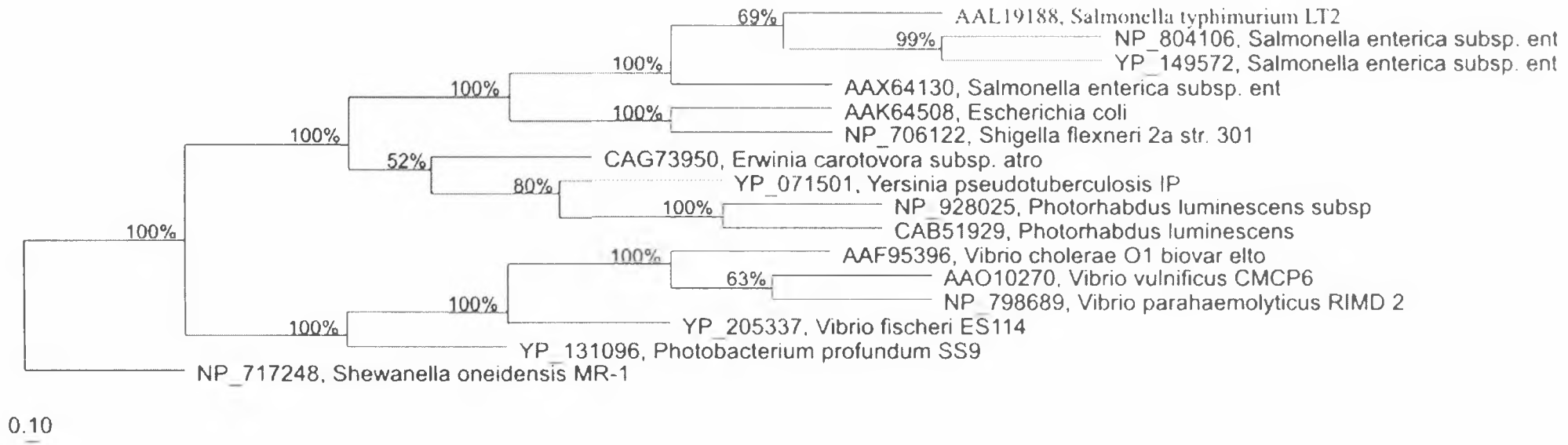


Fig 4.2 Phylogenetic tree of Vpr homologues possessing > 50 % protein sequence identity as identified by blastp analysis of Vpr. The *E. coli* node shows one protein but represents all *E. coli* Vpr sequences up to 19/08/05, as all have identical protein sequence. The tree was derived by parsimony algorithm with branches bootstrapped from 100 analyses. The tree shows relatedness between homologues of Vpr. Scale bar equal to 0.1 % amino acid change. Alignment from amino acid position 18 – 803 of Vpr (785 amino acids aligned).

The protein sequence alignment in Appendix 1 reveals large regions of sequence identity between bacterial families that have been sequenced thus far (including the *Enterobacteriaceae*). Vpr has been identified as being part of a hetero-oligomeric complex with three other proteins, YfgL, YfiO and NlpB, which previously had no designated function (Wu *et al.*, 2005). The interaction was identified using His-tagged YfiO and NlpB, which, when co-precipitated showed the presence of YaeT and YfgL by western analysis (Wu *et al.*, 2005). Vpr (YaeT) has been identified as the β-barrel in this complex which is responsible for lipid transfer from the cytoplasmic side of the inner membrane to the OM (Wu *et al.*, 2005; Ruiz *et al.*, 2005; Doerrler *et al.*, 2005). YfgL has been identified as an OM lipoprotein that is involved in the assembly of OM β-barrels, either through being a non-essential part of construction or by regulating associated machineries for the assembly of the OM (Ruiz *et al.*, 2005). YfgL and YaeT have been shown to be involved in the regulation / production of both LamB, the bacteriophage λ ligand and OmpA (Wu *et al.*, 2005) the T-even phage ligand. Through mutation or insertion of an inducible promoter for YfgL an effect on the production of all β-barrels proteins was observed (Wu *et al.*, 2005), thus underlying the essential role of Vpr in the biogenesis of the OM. This essentiality is further corroborated by reports by Gerdes *et* al. 2003 and Wu *et al.* 2005 and our attempts to make a viable Vpr mutant. An orthologue of Vpr (Omp85) has been identified in *Neisseria meningitidis*. Genevrois *et al.* (2003) showed that by placing the Omp85 gene under the control of an inducible promoter it was possible to identify by electron micrograph an accumulation of material in the inner membrane. On fractionation this material was found to be composed of phospholipids and LPS

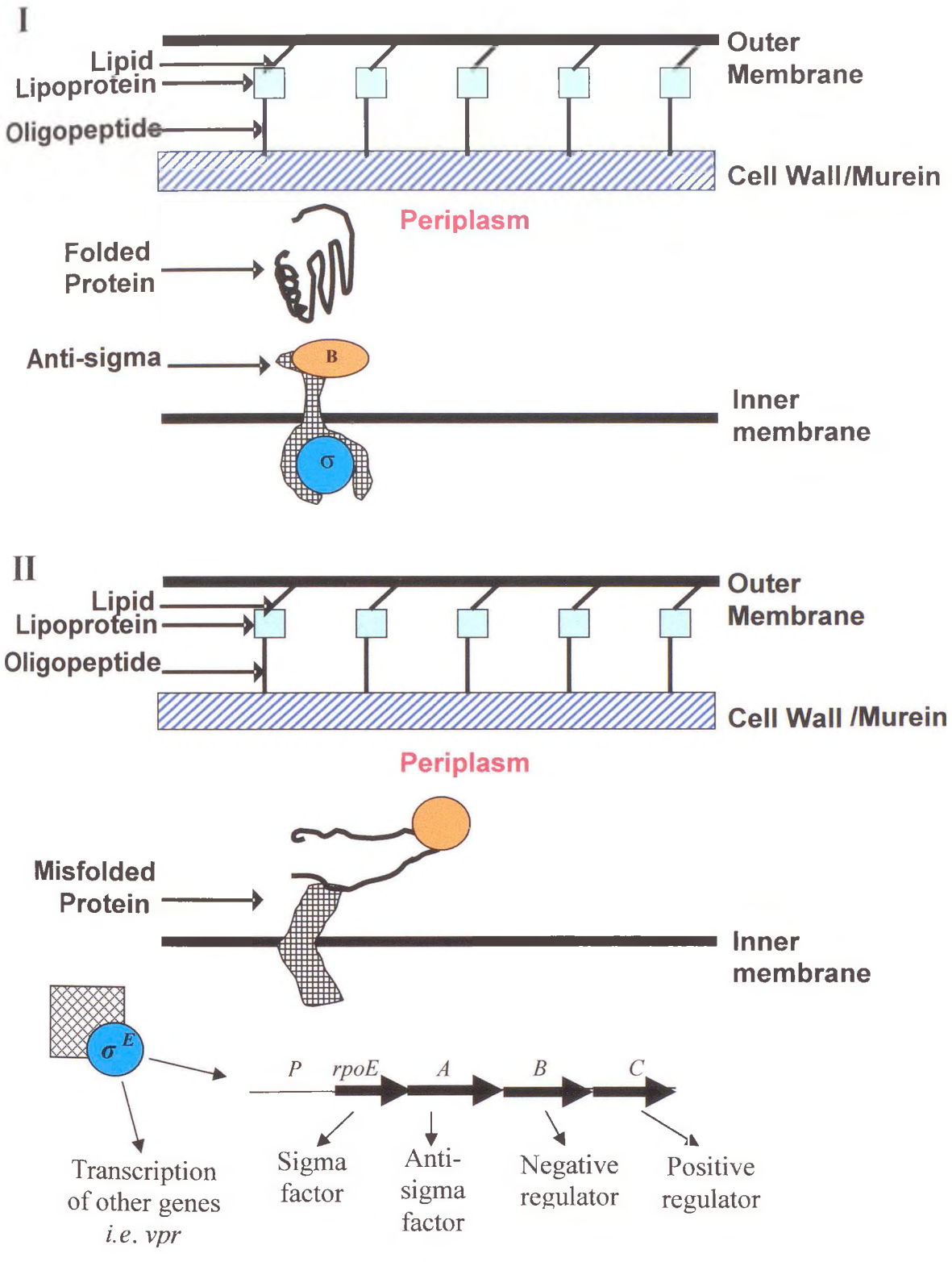


Fig 4.3 Functional Model of σ^E regulation. Under non-modifying growth conditions (**I**) the anti-sigma factor is held within the periplasm with no effect. When environmental conditions alter to stimulate the mis-folding of proteins (**II**), the release of the sigma factor initiates transcription of the *rpoE* regulon; the sigma factor (σ^E) is the first gene to be transcribed and subsequently autoregulated by the operon. Modified from Missiakis *et al.* (1998).

Dartigalongue *et al.* (2001) identified Vpr (*ecfK*) as part of a transcriptional regulon that is stimulated by misfolding of proteins in the periplasm and named the σ^E regulon. It is a member of the σ^{70} sub family of sigma factors, and σ^E is the product of the *rpoE* gene (Missiakis *et al.*, 1998). Fig 4.3 shows how the *rpoE* (σ^E) regulon is stimulated. Environmental changes that would cause the mis-folding of proteins in the periplasm inaugurate the release of the σ factor, which in turn initiates transcription of this autoregulated regulon. This response can up-regulate the production of genes such as *vpr*. To date, 47 σ^E-dependent promoters, controlling-expression of ~ 100 genes, have been identified (Onufryk *et al.*, 2005). Most of these proteins are involved in production of the cell envelope, including chaperones *e.g.* Skp (skp is found just upstream of *vpr*); (Chen *et al.*, 1996) and proteases that target folded and misfolded periplasmic proteins, transporter proteins, as well as phospholipid and LPS biosynthesis (Onufryk *et al.*, 2005). Changes in the production and localisation of Vpr at the cell surface would increase or decrease the chances of phage adsorbing to a host. Using parameters relevant to the mammalian gut environment, the levels Vpr production in relation to phage adsorption was examined using both a promoter-reporter gene fusion described in section 4.5.3 and phage adsorption assays.

4.3 Background: Creation of *vpr*-knockout mutant using insertional mutagenesis by double recombination.

Using a classical molecular genetic approach it was decided that a logical step was to knock out the Vpr gene and demonstrate phage resistance. This combined with complementation would unequivocally demonstrate that Vpr is the ligand required for $\Phi 24_B$ adsorption and infection. Initially Sergeant (1998) tried to

knock out *vpr* using a pUC18-derived approach in which 1.8 Kb of *vpr* was interrupted on the vector at the *Nru* I restriction site with *aph*3 encoding kanamycin resistance. Selection for kanamycin-resistant, ampicillin (carried on pUC18 backbone)-sensitive cells would indicate that a double recombination event had occurred. Multiple attempts failed to produce this mutant. James (2002) used both single and double homologous recombination approaches to knockout *vpr* including; the transformation of linear DNA, single and double homologuos approaches using SacB-mediated suicide vector (pMCS-1-*sacB*ΦR1C (Sergeant, 1998)), double homologous recombination approaches using plasmid shuttle vectors, but the approach that was persisted with was a single homologous recombination approach using a λ-*pir* suicide vector derived from pJP5603 (pYY::*vpr*).

Explanations that can be offered to support Sergeant's findings is that the use of a cocktail of Stx-phages has identified a strain containing spontaneous point mutation/s in *vpr*, which not only makes it resistant to $\Phi 24_B$, but has no influence on essential gene function. It also may be possible that a pleotropic mutation in *vpr* may have an influence on adsorption/infection due to its regulation of another genes function.

4.4 Objectives

The aim of the work reported in this chapter is to unequivocally determine that Vpr is the receptor required for $\Phi 24_B$ infection and to further characterise the interaction between this bacteriophage and its respective bacterial ligand. As Vpr is also associated with a bacterial stress response (σ^E regulon), this research reports on how transcription levels of *vpr* can affect levels of the protein localised at the bacterial cell surface under different growth conditions. Immunofluorescence

confocal microscopy was also used to show that Vpr is localised at the outer membrane. Previous research into phage – bacteria interactions has had the advantage that the host receptors were non-essential for bacterial survival. This study aims to prove that this essential protein is the bacterial ligand responsible for $\Phi 24_B$ infection and determine if this ability can be transferred to naturally resistant species. The level of infection of *E. coli* in the animal gut (*in vivo*) with Stx bacteriophages could be significantly affected by the levels of expression of Vpr. Since this gene is now known to be part of the σ^E regulon responsive to environmental conditions, Vpr expression in relation to phage adsorption under different conditions was examined.

4.4.1 Adsorption of bacteriophage $\Phi 24_B$

As attempts to knockout *vpr* using molecular methods were unsuccessful, an alternative strategy was required to characterise the interaction between $\Phi 24_B$ and Vpr. An adsorption assay to measure adsorbed bacteriophage would provide data on the interaction between the bacteriophages and their bacterial receptor. Previously, Sergeant (1998) had used a phage adsorption assay designed by Kilno and Rothman-Denes (1989) for bacteriophage N4 adsorption. Briefly, 5 ml of an overnight culture of *E. coli*, concentrated to 1 ml by centrifugation (2300 x g, 5 min) was then incubated for 20 min at 37 $^{\circ}$ C with 0.1 ml of bacteriophage suspension (~ 10^{8} pfu ml^{-1}). The infection mix was subjected to centrifugation (16100 x g, 1 min) and the unadsorbed phage particles were recovered. These phages were subsequently titred by plaque assay from the supernatant following passage through a 0.45 μm filter. This method was able to demonstrate adsorption of bacteriophage, but the quantitative reproducibility was very poor. Most infection assays, such as plaque

assay, use bacterial cultures that are in mid-exponential growth phase. The Kilno and Rothman-Denes (1989) method used overnight cultures. Stationary phase cultures contain much capsular material which may result in localised differences in receptor presentation due to an extra-polysaccharide masking effect, thus distorting phage adsorption rates. The numbers of cells will have an effect on the amount of phage that can adsorb, as the higher absolute amount of receptor results in higher numbers of adsorbed bacteriophages. Furthermore, the presence of large numbers of cells may lead to the formation of a cellular matrix-like structure that will also lower the probability of the phage being able to access the receptor or artificially trap phages during centrifugation. The final centrifugation step in the Kilno and Rothman (1989) procedure to pellet the adsorbed phage and bacteria is very strong (16100 x g). This may dislodge phage particles that are at a reversible point of adsorption (*i.e.* prior to injection of the phage genome into the sensitive bacterial host). Due to the doubling time of actively growing *E. coli* (~ 20 min), the contact time during the adsorption step should be less than the doubling time, as an increased number of cells may increase the number of available receptors and therefore introduce discrepancies to the quantitative adsorption assay. With these parameters in mind, it was possible to modify the Kilno and Rothman-Denes (1989) method to yield a more reproducible adsorption assay. The protocol developed is presented in Fig 4.4.

Enumeration of bacteriophage adsorbed was ascertained by calculating the percentage of phage particles that were adsorbed (%-adsorption) using the following calculation:

% -adsorption = (100 / no. of phage in control) x no. of phage in sample)

% adsorption inhibition = 100 - (100 / no. of phage in control – no. of phage 100 % adsorption) x no. of phage in sample)

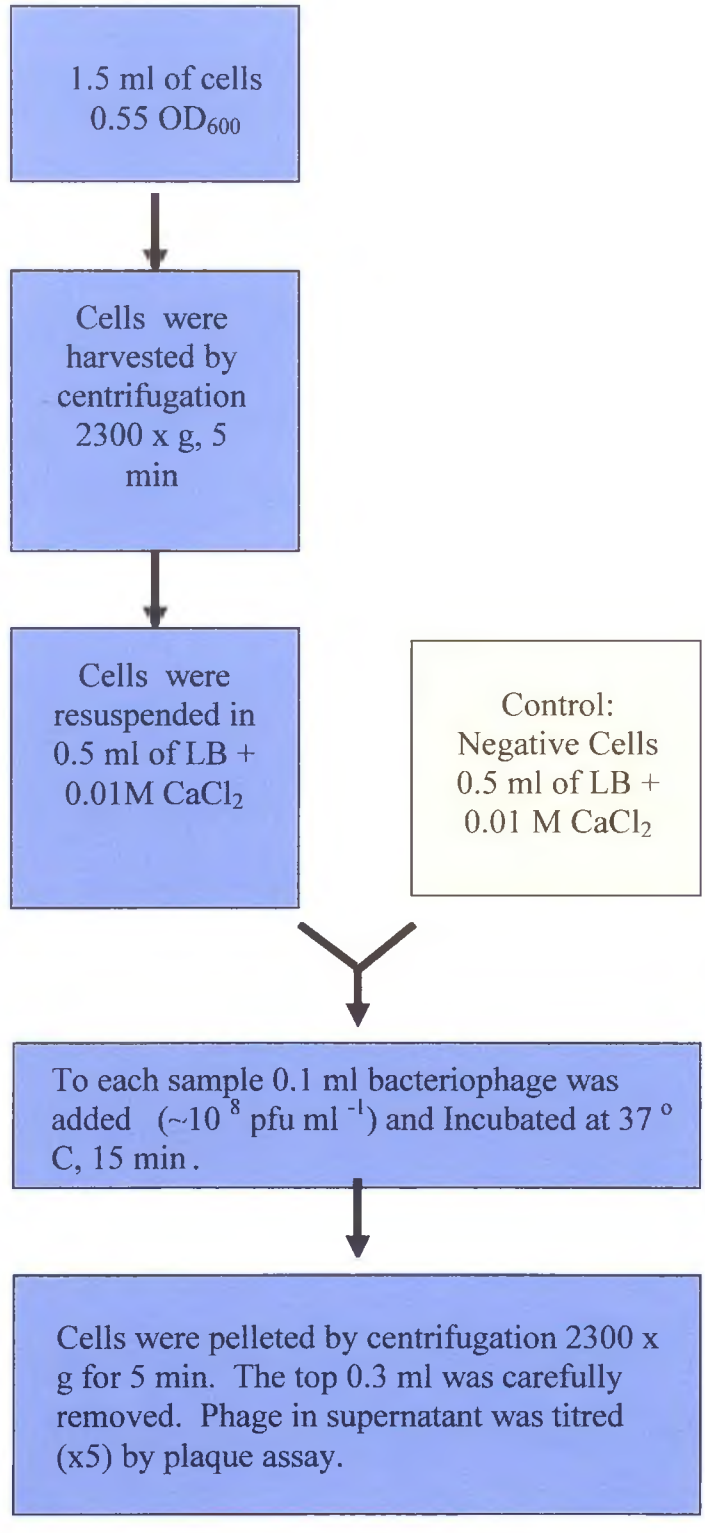


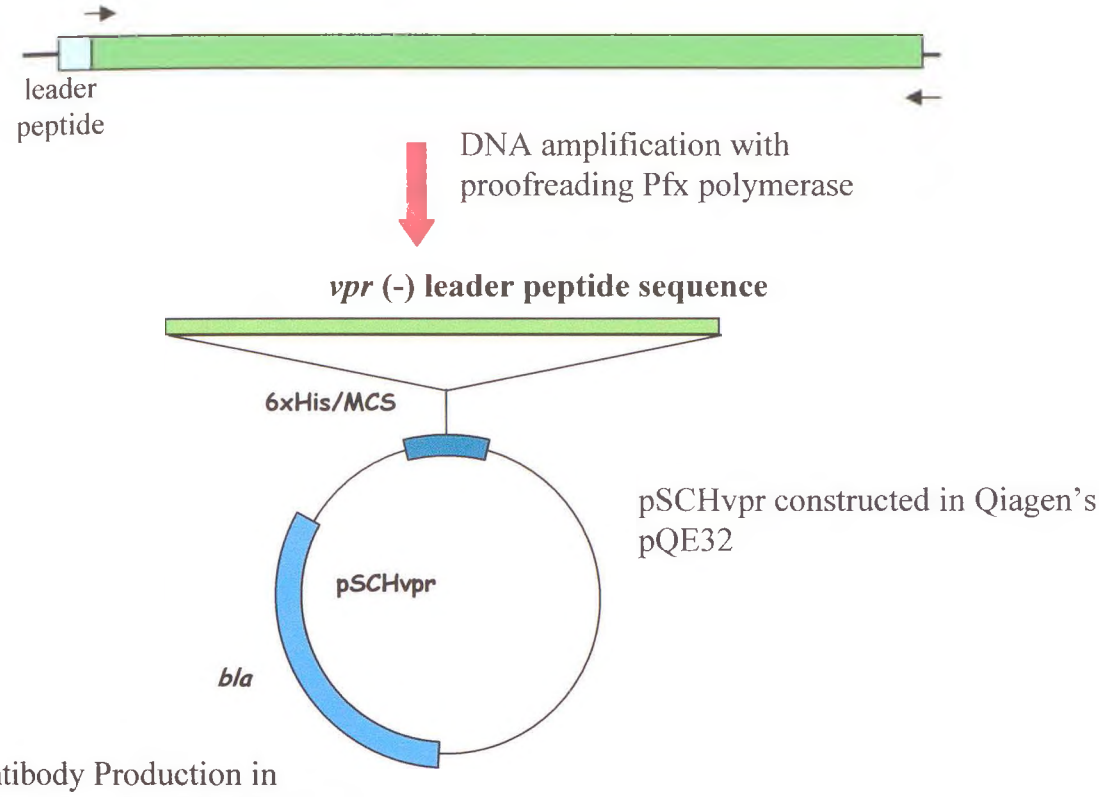
Fig 4.4 **Adsorption assay protocol developed and applied to $\Phi 24_B$ and *E. coli* cells.** Each adsorption assay was performed in triplicate, and each replicate was subjected to five plaque assays.

4.4.2 Inhibition of adsorption using a poly-clonal anti-rabbit anti-Vpr antibody.

Previous to this study, antibody was raised in New Zealand white rabbits to recombinant-Vpr. As previously mentioned in the introduction, antisera was used to inhibit bacteriophage infection of *E. coli* using antibodies specific to both the λ tail spike and its primary *E. coli* ligand, LamB (the maltose transport protein); (Wang *et al.*, 2000). Allison (personal communication) produced a recombinant Vpr protein via the Qiagen expression system using vector pQE32 (Fig 4.5). The His-Vpr protein was purified by affinity chromatography and eluted under denaturing conditions.

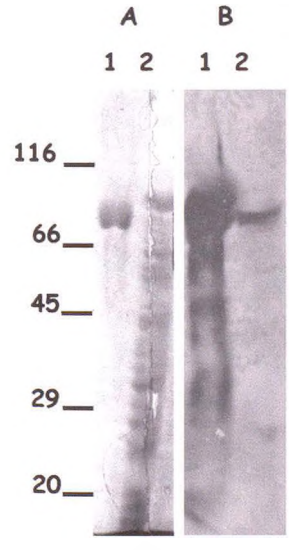
The bacteriophage adsorption inhibition assay protocol developed here is presented in Fig 4.6. These data in Fig 4.7 demonstrate that by increasing the dose of anti-Vpr antibody, it was possible to increasingly mask the epitope that $\Phi 24_B$ used to infect *E. coli*. Furthermore, these data in Fig 4.8 show that the pre-immune serum of the rabbit had no influence on phage inhibition. As the polyclonal antibody was raised to a denatured protein, it was possible that the antibody would not have the ability to inhibit adsorption. Clearly, however, the anti-Vpr antibody has the ability to mask the Vpr epitope required for $\Phi 24_B$ adsorption and infection. Adsorption is a preferable marker of interaction to infection, as it directly examines the protein-protein interaction between the bacteriophage tail spike and the bacterial ligand. Adsorption assays are a good tool, although discrepancies between individual experiments can be quite high. For example it was found that multiple experiments will show the same overall effect, although absolute adsorption

I)



Antibody Production in
New Zealand White Rabbits

II)



Western blot analysis of anti-Vpr sera.

Panel A) 7.5% SDS-PAGE,

Panel B) western blot using 1:10,000 rabbit anti-Vpr and AP-labelled goat anti-rabbit.

Lane 1) Affinity chromatography purified His-Vpr eluted under denaturing conditions

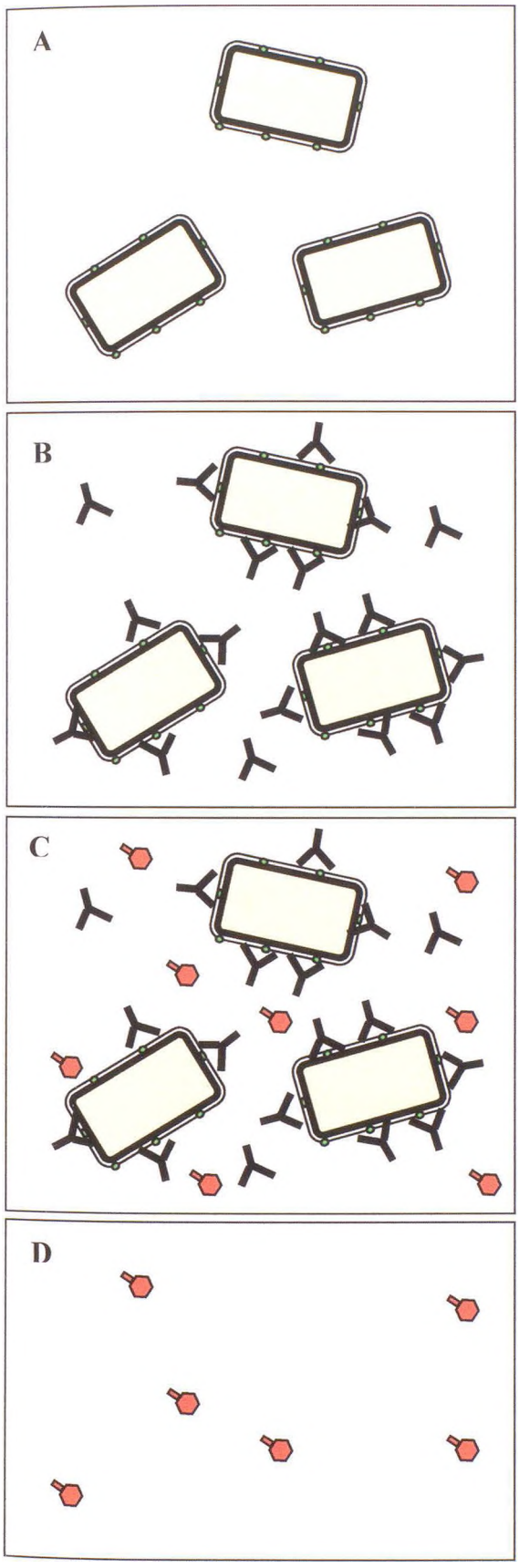
Lane 2) whole cell lysate of w.t. *E. coli* strain MC1061.

Fig 4.5 Production and Validation of Rabbit anti-Vpr antibody

I) Vpr minus its leader peptide was cloned into expression vector pQE32 for recombinant Vpr production.

II) Panel showing purified His-Vpr and rabbit anti-Vpr recognition of Vpr by SDS-page and Western analysis.

(Diagram modified from H. E. Allison, personal communication)



A) 1.5 ml of exponential growth phase cells (~ 0.55 – 0.6 OD_{600}) harvested by centrifugation (2300 x g, 5 min) and resuspended in 1 ml of LB, 0.1 M $CaCl_2$.

B) Rabbit anti-Vpr antibody added and allowed to bind to its cell surface target at ~ 18 °C for 5 min.

C) 0.1 ml of bacteriophage suspension (~ 10^8 pfu ml^{-1} added to suspension and allowed to adsorb for 15 min at 37 °C.

D) The cells were pelleted by centrifugation (2300 x g, 5 min), removing adsorbed phages from the preparation. Remaining unadsorbed phage were enumerated by plaque assay.

Fig 4.6 Inhibition of bacteriophage adsorption using a polyclonal rabbit anti-Vpr antibody.

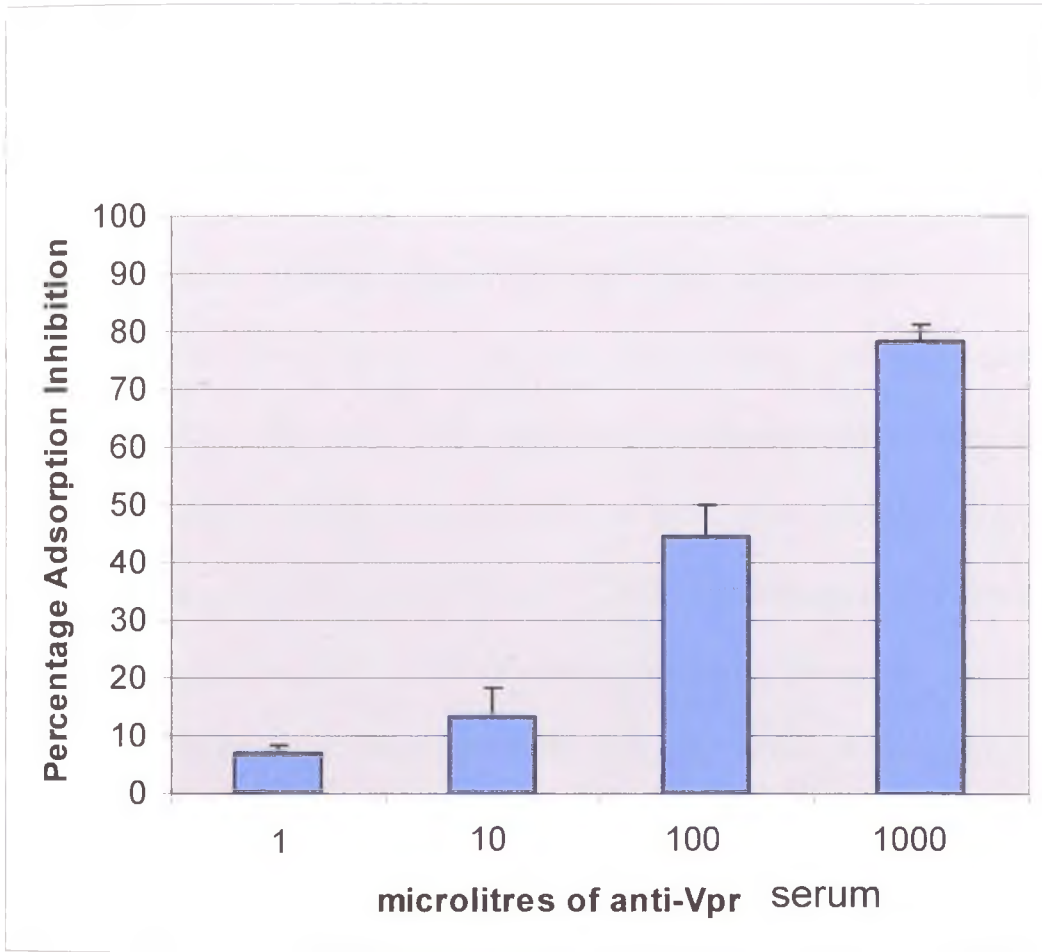


Fig 4.7 Inhibition of phage adsorption using rabbit anti-Vpr antibody

A sequential 10-fold increase in the addition of anti-Vpr Ab yields a positive dose response determined as percentage inhibition of phage adsorption. Therefore, an increase in anti-Vpr yields increased phage recovered resulting from inhibition of phage adsorption expressed as a percentage. Samples were assayed in triplicate and and subjected to plaques assay 5 times. Error bars: SEM, n = 15.

inhibition rates may vary by up to 20 %. There are a large number of parameters in this experiment and the dynamics of phage binding to the host receptor are likely to be complex. In order to limit cell growth cycle differences and ensure that the cell numbers were identical for the inhibition assays, bacterial cells were harvested from the same culture throughout the experiment.

4.4.3 Restoration of phage adsorption ability to a resistant host

Vpr has an essential function for bacteria grouped within the *Enterobacteriaceae* (see 4.2). This may mean that the use of this outer-membrane protein as a target receptor would result in a broad host range for $\Phi24_B$, as has indeed been demonstrated by James *et al.* (2001). Homologous Vpr proteins have been found in all members of the *Enterobacteriaceae* sequenced thus far. Vpr homologues that have been identified in the *Enterobacteriaceae* possess a conserved nature which may account for the broad host range of $\Phi24_B$. The alignments in Fig 4.9 show that *Erwinia carotovora* sbsp. *atroseptica* (ECA) has a Vpr with 87 % protein sequence identity to that found in *E. coli* K-12. It was demonstrated that this *Enterobacteriaceae* species was unable to support $\Phi24_B$ adsorption and thus could be used as a host to which phage sensitivity might be conferred by introduction of the *E.* coli Vpr gene. Furthermore, ECA is a plant pathogen that cannot colonise the mammalian gut, therefore conferring sensitivity to a Shiga toxin-encoding bacteriophage was regarded as a "safe" and ethically uncontentious experiment. Adsorption assays using *ECA* in place of *E. coli* showed that $\Phi24_B$ did not adsorb to *ECA*, meaning either that the Vpr orthologue in *ECA* was masked in some way by the cell surface topography or did not contain the epitope required for $\Phi24_B$

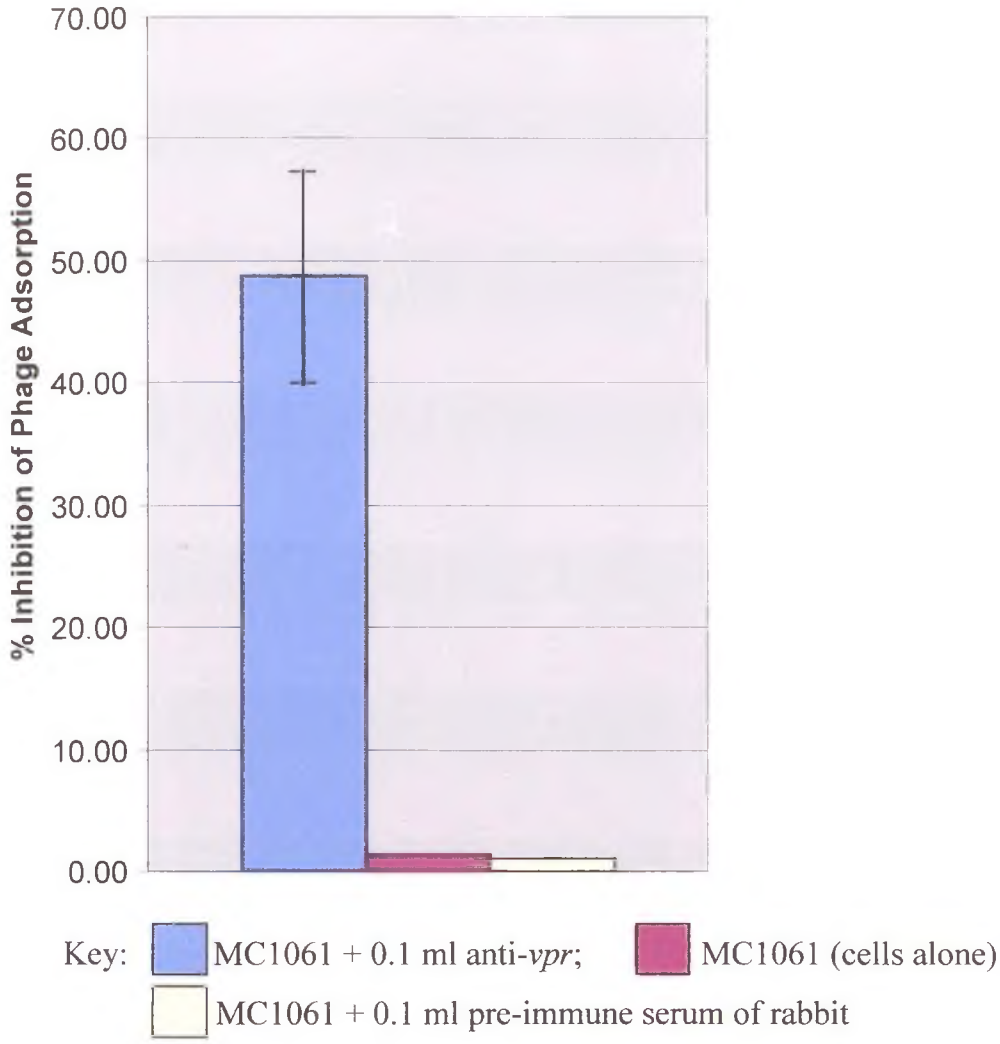


Fig 4.8 Inhibition of $\Phi 24_B$ adsorption using rabbit anti-Vpr antibody and rabbit pre-immune sera. Comparison of the percentage inhibition of phage adsorption to *E. coli* K-12 MC1061 cells treated with both the serum of the pre-immune rabbit and the serum of the His-Vpr immunised rabbit. These data show that the pre-immune serum does not contribute to inhibition of $\Phi 24_B$ adsorption and the results are comparable to *E. coli* (MC1061) cell suspension to which antiserum has not been added. Error bars: SEM, n = 15

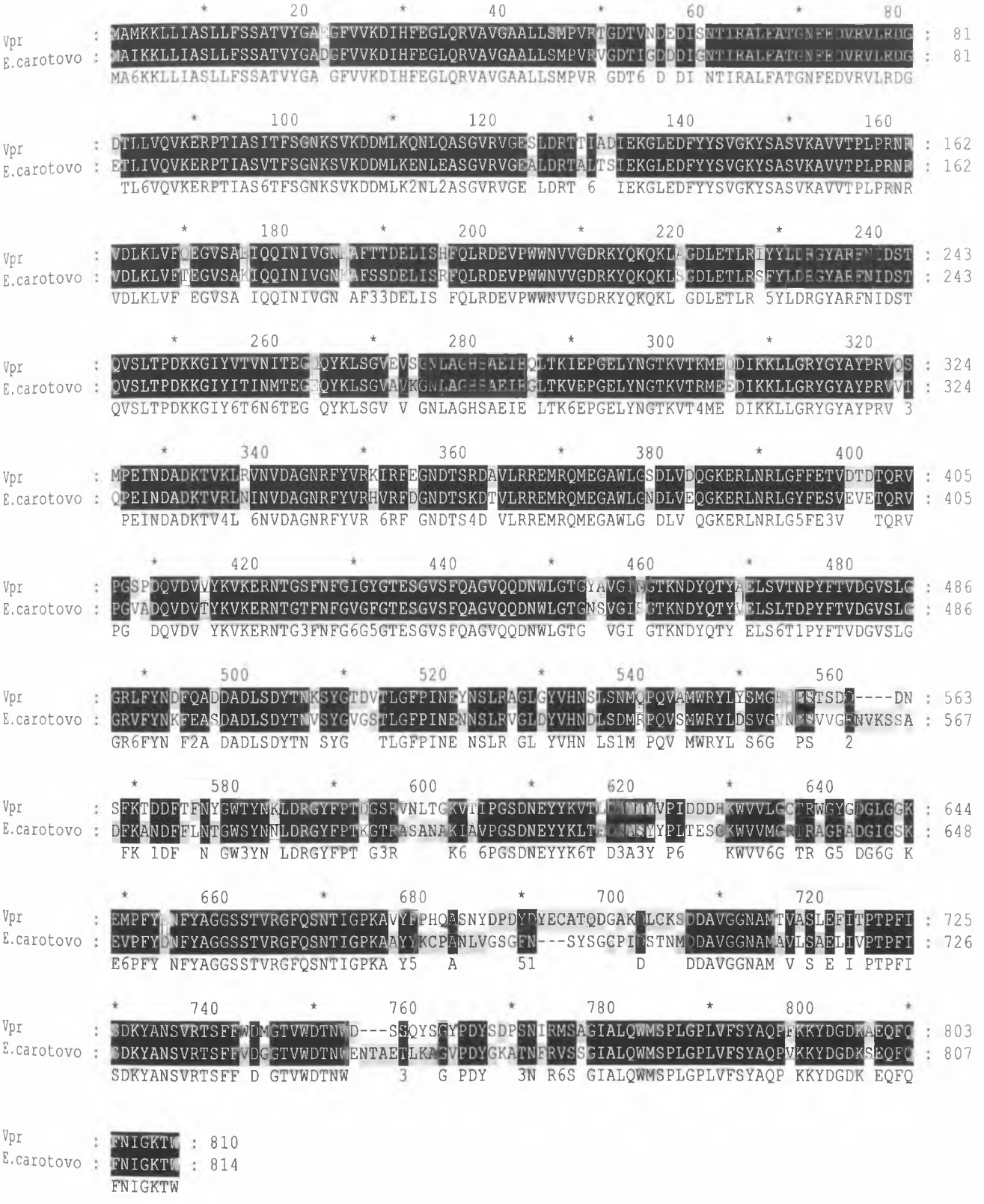


Fig 4.9 Protein sequence alignment of Vpr from *E. coli* (MC1061) against the Vpr orthologue found in *Erwinia carotovora* sbsp. *atroseptica*. Vpr protein sequences annotated in all sequenced *E. coli* strains thus far are identical in *E. coli*. These data show a protein alignment of 87 % protein similarity between Vpr from *E. coli* and the orthologue found in *E. carotovora* sbsp. *atroseptica* (*ECA*). Black highlighted amino acids denote an identical match, shaded amino acids denote amino acid difference between sequences and a dash identifies an insertion or deletion between sequences.

adsorption and infection. Manipulation of bacteriophage adsorption sensitivity by transferring a cloned gene between species has never been previously described.The objective therefore was to introduce *E. coli* K-12 *vpr* into *ECA* to determine whether the ability of the phage to adsorb to *E. coli* could be transferred to the plant pathogen. This would provide strong support for the hypothesis that Vpr is the receptor for $\Phi 24_B$. The vector selected for this experiment was pKT230, a broad host range plasmid previously used for general cloning in Gram negative bacteria (Bagdasarian *et al.*, 1981), a derivative of RSF1010 (Scholz *et al.*, 1989).

A number of approaches were used to clone *vpr* into pKT230, the most successful of which are presented in Fig 4.10. In the first instance, *vpr* was excised from pUC19ΦR1D with *EcoR*I, the *EcoR*I sites on pUC19 are located either side of the multiple cloning site which in, pUC19ΦR1D, includes the entire v*pr* coding region. In the second approach *vpr* was amplified by PCR, using oligonucleotide primers (5' vpr-*EcoR*I and 3' vpr-*EcoR*I, Table 2.1) and the proof-reading DNA polymerase pfx (Invitrogen). These oligonucleotide primers are tailed with the target restriction site for subsequent *EcoR*I digestion. For this second approach two methods were used to try to optimise increased *EcoR*I digestion to release the *EcoR* I sticky ended fragment. The PCR product was digested with *EcoR*I according to manufacturer's guidelines (MBI Fermentas), and the fragment was purified using a PCR clean up kit (Qiagen) and ligated into *EcoR*I-digested pKT230 using T4 DNA ligase (MBI Fermentas) according to manufacturers' guidelines. The blunt-ended *vpr* PCR fragment was also cloned into vector pCR ZeroBlunt (Invitrogen) and transformed into *E. coli* Topo cells (Invitrogen). The plasmid was harvested from cells (mid-exponential growth) using a plasmid extraction kit (Qiagen). The

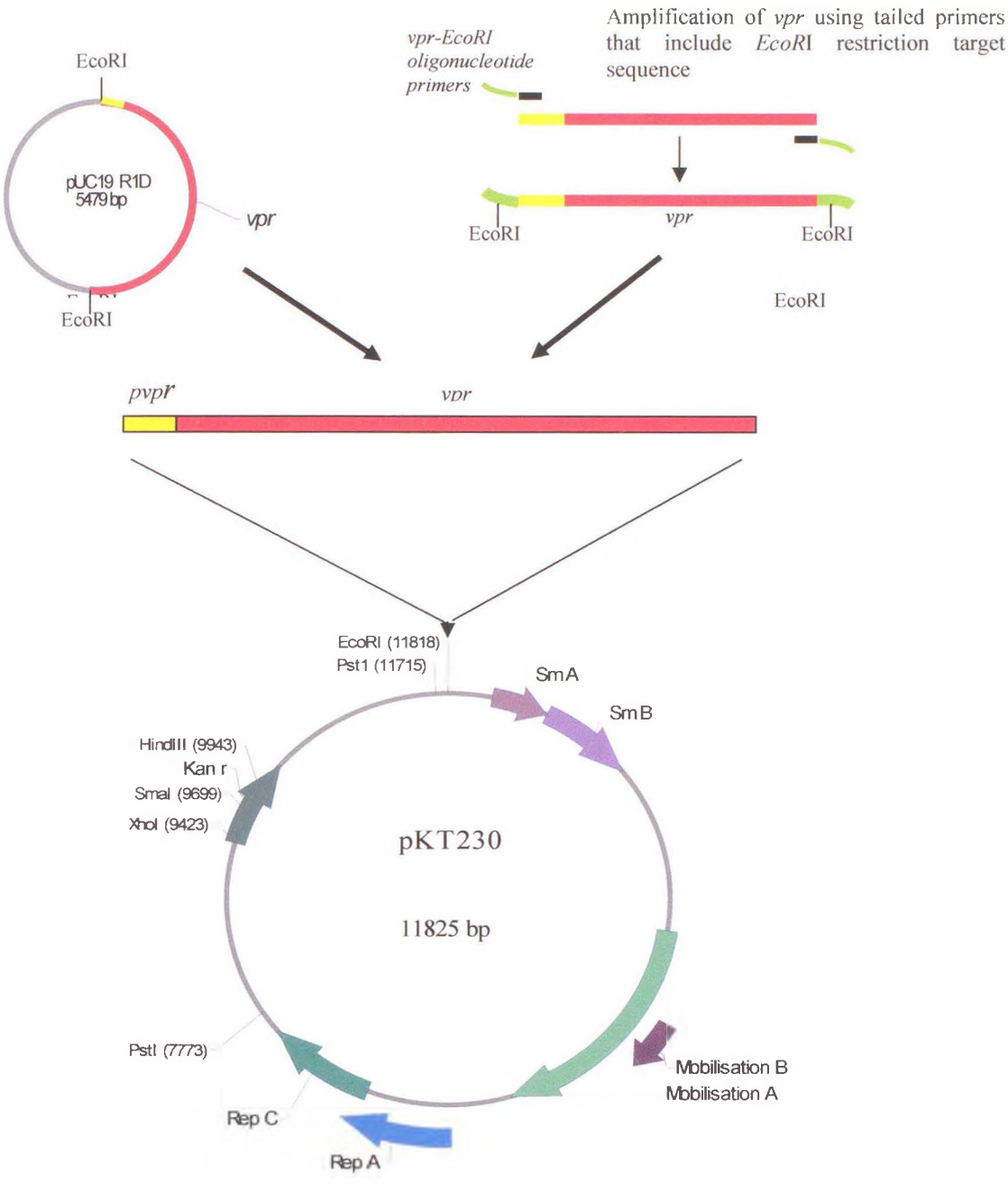


Fig 4.10 **Molecular approaches for cloning *vpr* into pKT230.** **I)** *vpr* was excised from pUC19ΦR1D using pUC19 borne *EcoR* I restriction sites and cloned into *EcoR*I-digested pKT230. **II)** *vpr* amplified using Vpr-*EcoR*I forward and reverse oligonucleotide primers that include an *EcoR*I target site for subsequent cloning into *EcoR*I-digested pKT230.

extracted plasmid was then digested with *EcoR*I and ligated into *EcoR*I-digested pKT230. Ligation reactions were transformed into Topo *E. coli* cells (Invitrogen) according to manufacturer's instructions and putative transformants selected on LB with kanamycin (50 μg ml^{-1}). These were re-plated on LB with kanamycin 50 μg ml^{-1}) and screened for presence or absence of recoverable plasmid. Ten ml of culture was required for this procedure as the plasmid copy number is low, and the plasmid has a size of ~15 Kb.

Plasmid preparations were digested with *EcoR*I, some of which yielded a ~ 3 kb fragment. Its identity was confirmed using oligonucleotide primers designed to flank the *EcoR*I site of pKT230 (Fig 4.11). The confirmed PCR (was sequenced) with the flanking primers pKT230-sequencing-For & Rev (MWG Biotech). Sequence analysis using Blast-n (NCBI) confirmed that the insert was *vpr*. pKT230::*vpr* was harvested from the Topo *E. coli* background using the Qiagen mini-plasmid DNA extraction kit and transformed into *ECA*. Putative transformants were screened using the plasmid-borne kanamycin resistance (LB + kanamycin, 50 μg ml^{-1}).

Problems of contamination with *E. coli* did occur at this point and they subsequently out-competed *ECA* causing problems when selecting for putative transformants. These contaminants were identified by SDS-page of total protein extracts that readily differentiated between *E. coli* and *ECA* (Fig 4.12). It was imperative that a protocol was put into action that gave the ability to select between *E. coli* and *E. carotovora* at a glance. Using selective medium TBX (Tryptone Bile X-glucuronide) (Oxoid) it was possible to differentiate between putative clones and *E. coli* contamination. *E. coli* carries *uidA* which encodes for beta-D-glucuronidase; this enzyme cleaves the chromogen (5-bromo-4-chloro-3indolyl-β-D-glucuronide

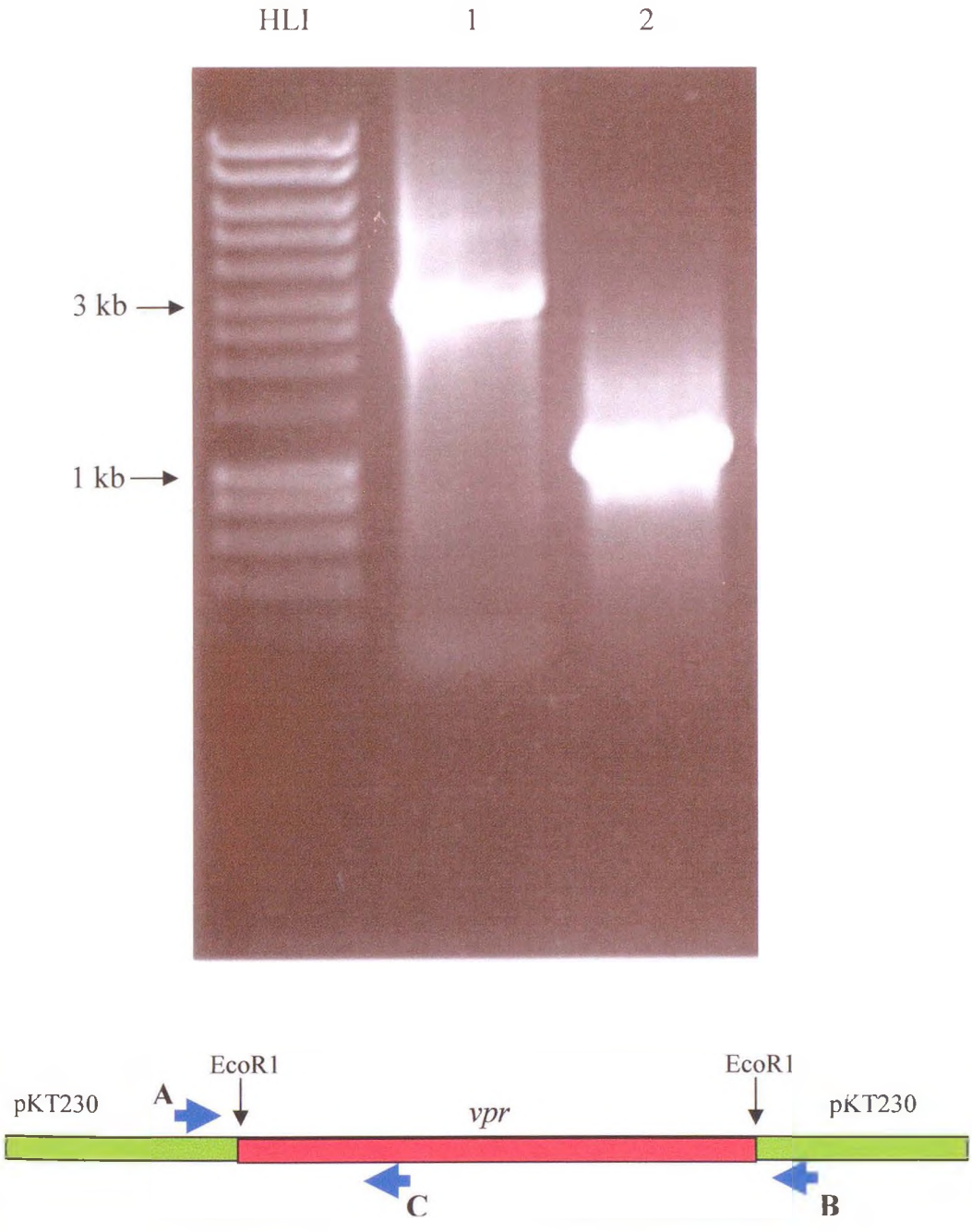


Fig 4.11 **Characterisation of pKT230::*vpr* by PCR analysis.** This shows that *vpr* is inserted into the *EcoR*I site of pKT230. Lane HL1: Hyper Ladder I (Bioline), Lane 2: ~ 3 Kb PCR product from oligo-nulceotide primers **A** (pKT230 – sequencing-For) and **B** (pKT230-sequencing-Rev), both of which anneal to regions flanking the *EcoR*I restriction sites. Lane 2: ~ 1 Kb PCR product from oligonucleotide primers **A** and **C** (vprrev1000, James, (2002).

(BCIG)) releasing a coloured chromophore which builds up in the cell and gives the colonies a blue/green appearence.

4.4.4 Confirmation of *E. carotovora* containing pKT230::*vpr* by SDS-page and western analysis

Putative clones were grown to OD $_{600}$ 0.55 at 30 ^{o} C for ~ 18 hr. The cultures (10 ml) were lysed by sonication and the levels of total cell protein assayed using the Bradford reagent (2.5). From previous SDS-page gels and western analysis loading levels of total protein had been optimised to 6 μg for SDS-page gels and 0.6 μg for western analysis. Fig 4.12 shows how the total protein profile of the clone used for further experimentation resembles the wild-type *ECA* strain. This clone was then subjected to western blot analysis using the polyclonal anti-Vpr antibody and visualised using a secondary anti-rabbit antibody conjugated to the horseradish peroxidase enzyme (anti-Digoxigenin-AP (Roche). Reactive proteins were then detected by chemiluminescence using the substrate, CPD* (Roche);(Chapter 2.4.6).

Fig 4.13 indicates that both *E. carotovora* and *E. coli* contain the epitope needed for the anti-Vpr antibody to bind. This is not surprising as they share 87 % protein sequence identity, but binding seems to be less efficient than that seen in *E. coli*, alternatively levels of production of Vpr by *ECA* are less than those in *E. coli* when equal amounts of total protein from OD_{600} 0.55 cells was loaded by western analysis. The western blot also shows that *E. coli vpr* is being transcribed and translated in the *E. carotovora* clone as the Vpr signal from the western analysis is substantially greater than the wild-type when equal amounts of protein were loaded.

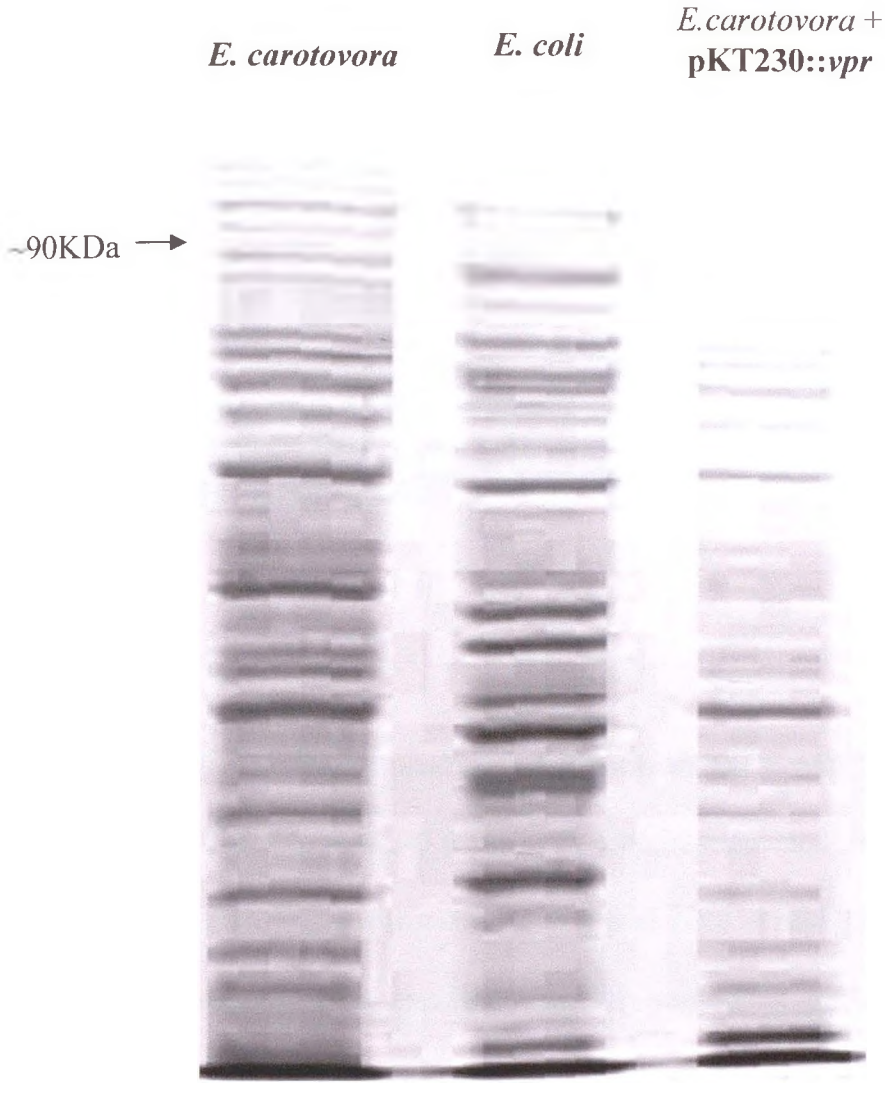


Fig 4.12 **SDS-page confirmation of *E. carotovora* transformant.** Image shows 6 µg of total protein extracted from mid-exponential growth phase cells of *E. carotovora, E. coli* and E. *carotovora* + pKT230::vpr (OD_{600} 0.55). Samples were run on a 7.5 % acrylamide gel, 200V, 40 min, and stained with Coomassie blue. Protein profiles are similar between the *ECA* wild type and *ECA* containing the pKT230::*vpr* construct.

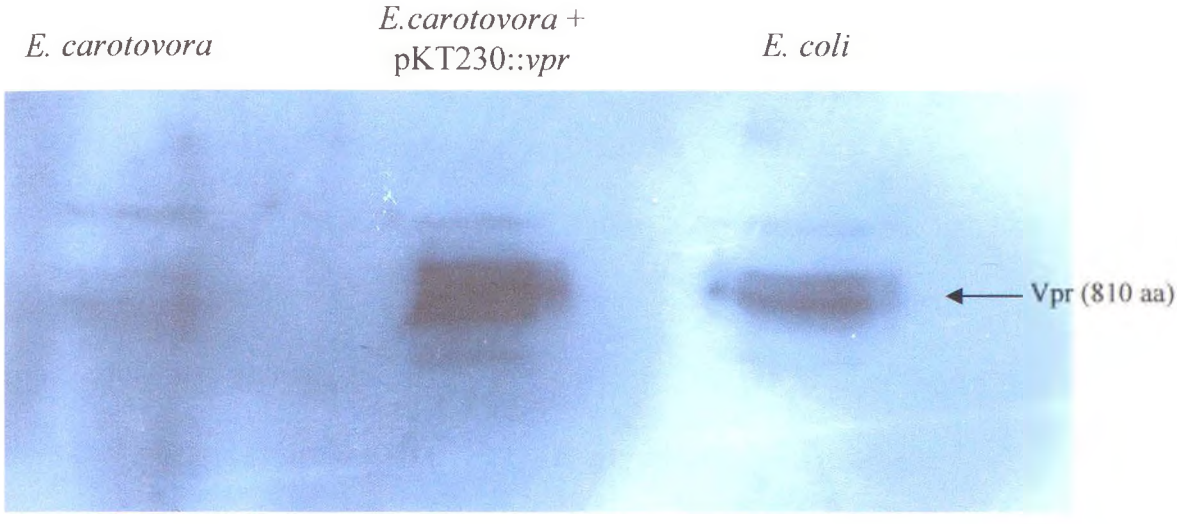


Fig 4.13 Western blot analysis of *E. carotovora* + pKT230::*vpr* using anti-Vpr antibody. Image shows 0.6 µg total protein harvested from mid-exponential growth phase cells of *E. carotovora*, *E. coli* and *ECA* + pKT230::vpr (OD_{600} 0.55). The primary antibody used was anti-Vpr, the secondary antibody was anti-rabbit whole IgG (whole molecule), conjugated to alkaline phosphatase (Sigma). CPD* (Roche) was used as substrate for chemiluminescence and visualised on X-ray sensitive film (Kodak). The *ECA* pKT230::*vpr* shows increased expression of Vpr compared to wild type ECA and *E. coli*; this may be due to accumulative expression of the chromosomal and plasmid copies of *vpr*. The difference in intensity of the bands in the blot seems to show that wild type *ECA* produced lower levels of Vpr than *E. coli* at this point in the cell growth cycle (OD_{600} 0.55).

4.4.5 Does increased expression of Vpr alter adsorption of $\Phi 24_B$

With pKT230::*vpr* transformed into *ECA*, levels of Vpr production appear to be raised (Fig 4.13). To further characterise the cells ability to regulate Vpr and localise it at the cell surface, Topo *E. coli* (Invitrogen) containing pKT230::vpr clone was used in an adsorption assay. Fig 4.14 shows how increasing the level of Vpr *in vitro* increases the level of $\Phi 24_B$ adsorption, Vpr was produced by both a chromosomal and plasmid borne copy of the gene. Fig 4.14 also shows that $\Phi 24_B$ does not adsorb to *E. carotovora* lacking the *vpr*-recombinant plasmid.

4.4.6 Adsorption of $\Phi 24_B$ to *Erwinia carotovora* sbsp. atroseptica containing pKT230::*vpr*

Fig 4.15 shows that by increasing the numbers of *ECA* cells containing pKT230::*vpr*, it was possible to increase $\Phi 24_B$ adsorption, thus suggesting that the *E. coli* Vpr was being localised at the cell surface. These data use wild type *ECA* as a negative control where there is significantly less reduction in the numbers of free bacteriophage with an increase in cell number. This can be attributed to increasing cell numbers, leading to increased probability of a filtration effect from a pelleting cell matrix action during the centrifugation step of the adsorption assay.

4.4.7 Response of *vpr* transcription to growth conditions using a reporter gene assay.

As discussed in section 4.2, Vpr has been identified as being part of a response mechanism in *E. coli* linked with biogenesis of the outer membrane. The aim of this part of the study was to determine whether altering growth conditions, linked to environmental conditions in the mammalian gut, would change the

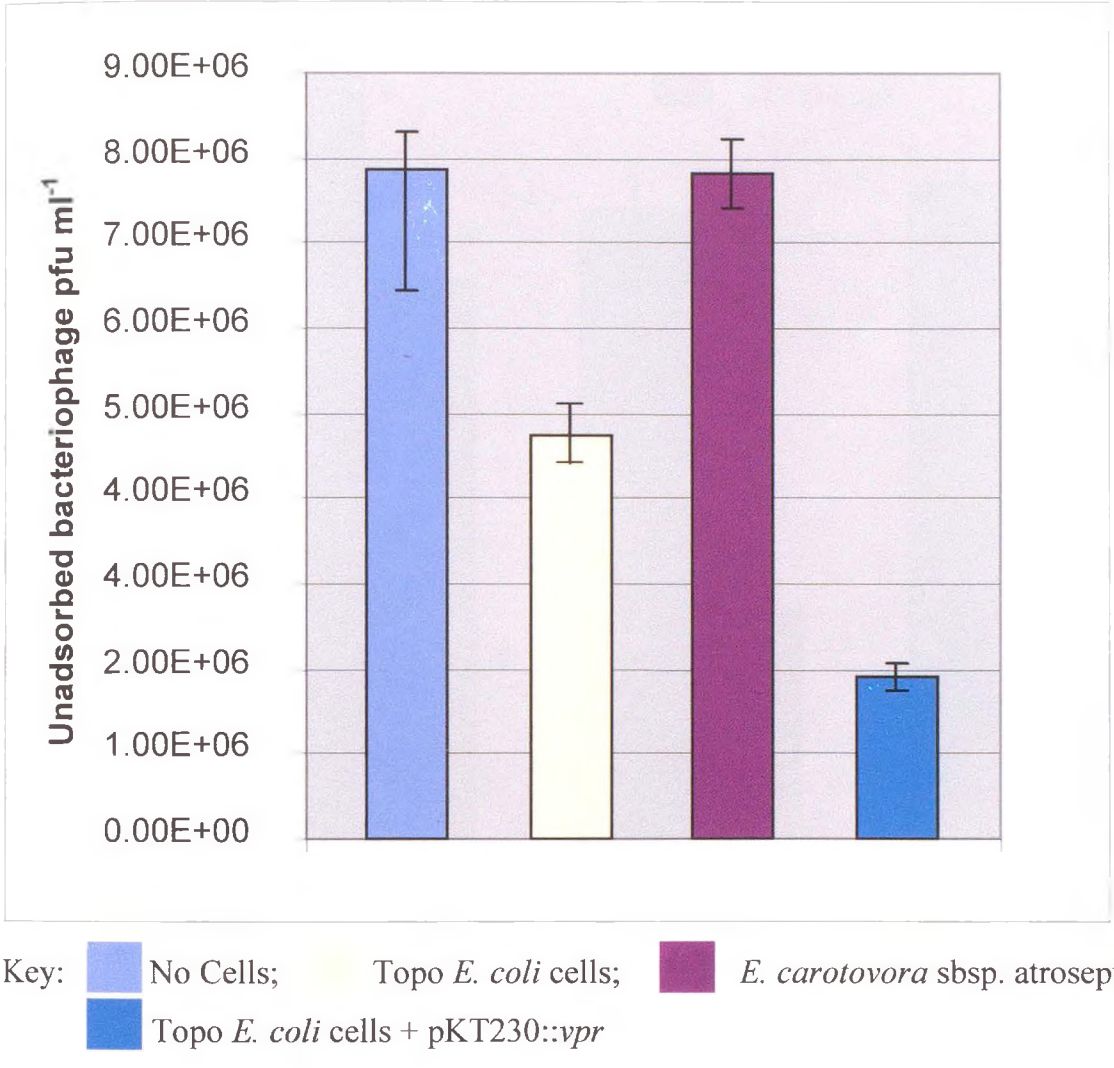


Fig 4.14 Adsorption assay to compare adsorption of $\Phi 24_B$ to wild-type *E. coli* and *E. coli* pKT230::*vpr*. The free phage particles recovered from the adsorption assay illustrate that $\Phi 24_B$ does not adsorb to *ECA* as there is no difference between free phage recovered and the control adsorption assay with no cells. $\Phi 24_B$ adsorbs to the pKT230::*vpr* *E.* coli transformants at a higher level than $\Phi 24_B$ to wild type Topo *E. coli*. (n)= 9, error bars = SEM.

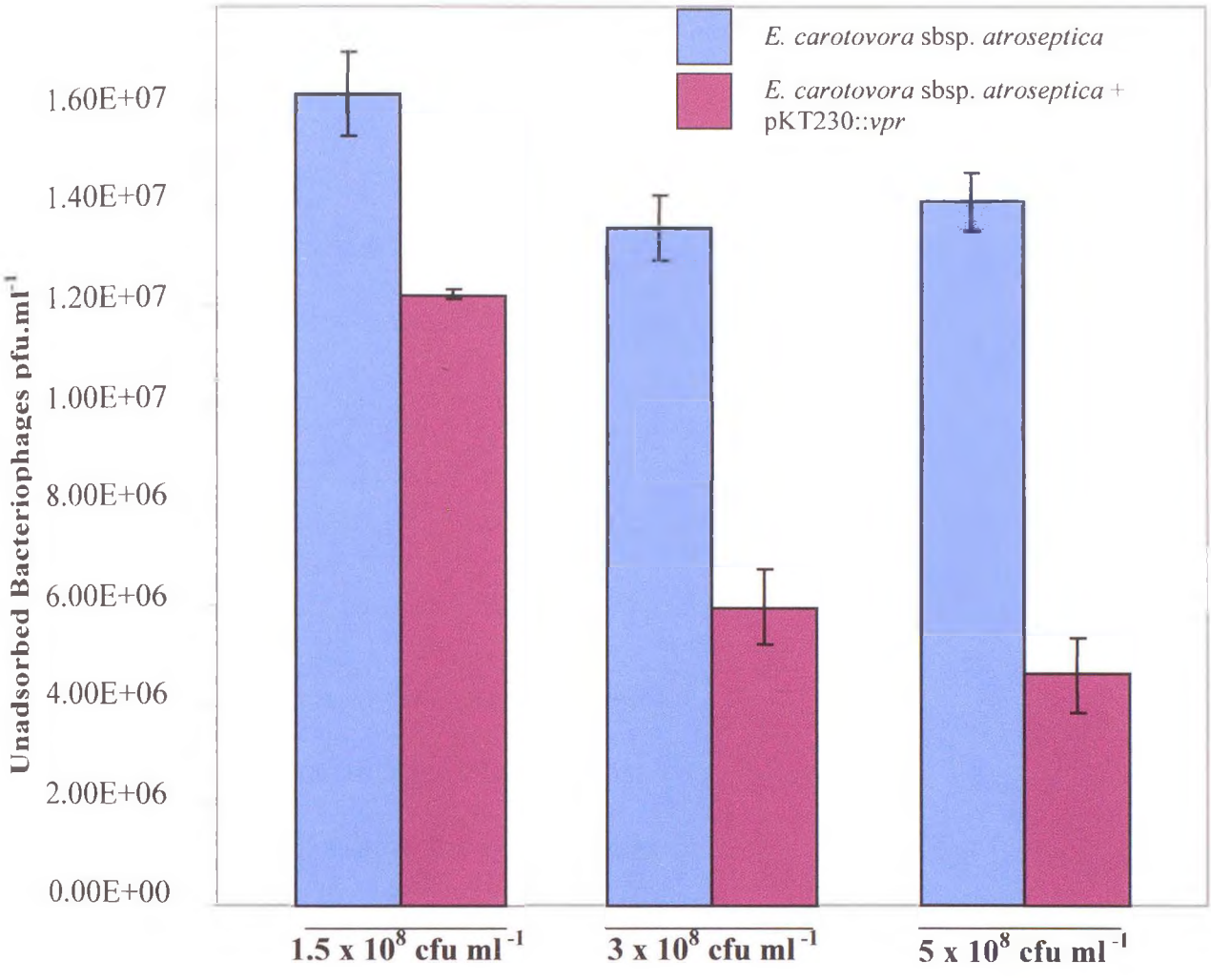


Fig 4.15 Adsorption of $\Phi 24_B$ to *E. carotovora* sbsp. *atroseptica* containing pKT230::*vpr*. In the presence of 1.5 x 10^8 cells of *ECA*, there was a 25 % reduction in the number of free bacteriophages recovered from the adsorption assay with the transformant and the *wt* ($P < 0.005$). In the presence of 3 x 10^8 cells, there was a 56 % reduction in the number of free bacteriophage particles recovered from the adsorption assay ($P < 0.004$). In the presence of 5 x 10^8 cells there was a reduction of 67 % in the number of free bacteriophages ($P < 0.0007$). These data show that by increasing the cell number of *ECA* + pKT230::*vpr*, the level of free phage particles recovered decrease thus showing increased phage adsorption. These data also show that this does not occur in phage adsorption assays using the ECA parent strain. These data were confirmed in an independent experiment. Sample number (n) =5, error bars = SEM.

galactosidase reporter gene (Fig 4.16). The activity of $_{p}vpr$ was determined by the (β-galactosidase) cleavage of o-Nitrophenyl-beta-galactopyranoside producing a colour change. Promoter activity is directly proportional to the amount of o-nitrophenyl produced measured spectrophometrically and expressed in Miller units (Miller, 1972). Changes in the level of transcription (Miller units) can be related to levels of Vpr production under different growth conditions.

$$\text{Miller Units} = \frac{OD_{420} \quad X \quad 1000}{\text{Volume of sample (0.5 ml) X Time (35 min) X } OD_{600} \text{ of Cells}}$$

4.4.8 Construction of the $_{p}vpr$::*LacZ* fusion

The oligonucleotide primers (Fig 4.16) amplify the promoter region containing a 19 bp tail of the first 19 bases of *lacZ*. It is this tail that makes it possible to anneal the *lacZ* gene onto the promoter in the correct orientation and identical reading frame to the *E. coli vpr* gene. It was possible to create the fusion by placing molecular equivalent amounts of both the promoter region and the *lacZ* gene into a PCR reaction using the external oligonucleotide primers (1 and 4, Fig 4.16 ; 94 ^{o} C denaturation, 5 min, cycle step 1 [15 cycles]; 94 ^{o} C, 2 min; 40 ^{o} C, 1.5 min; 68 ^{o} C, 4.5 min. Step 2 [35 cylces]: 94 ^{o} C, 2 min; 45 ^{o} C, 1.5 min; 48 ^{o} C 4.5 min, no final extension). This protocol amplifies any complete fusions between promoter region and reporter gene. Fig 4.17 shows the size increase between *lacZ* and the fusion product, an increase of ~ 150 bp, equivalent to the promoter region. Using standardised M13 primers, it was possible to sequence the gene fusion whilst

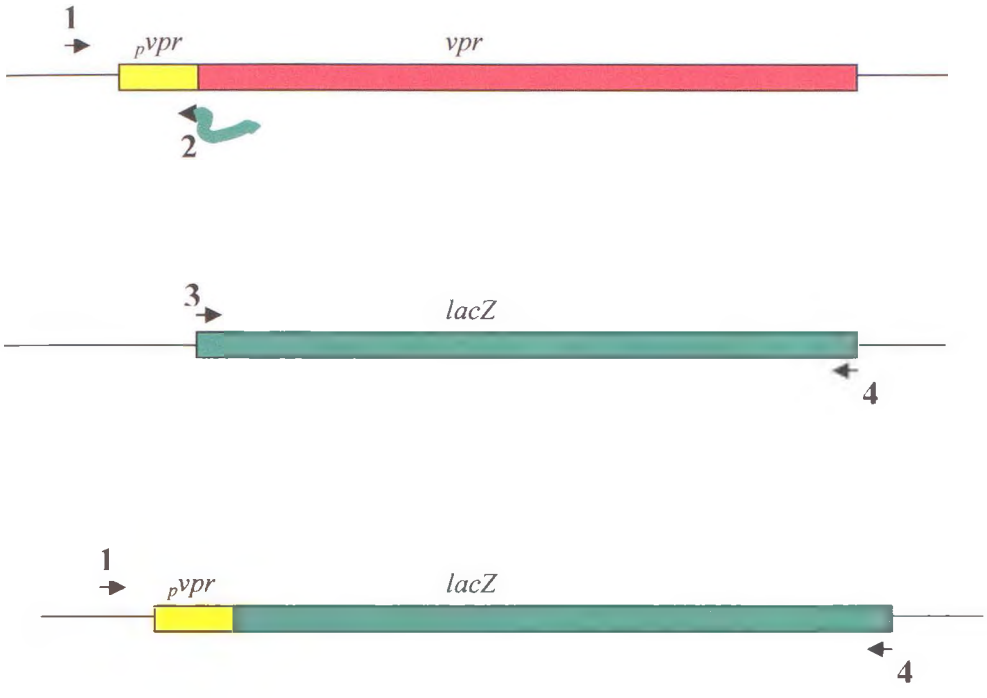


Fig 4.16 Construction of a *vpr* promoter reporter gene construct to quantify transcriptional regulation of *vpr* under different conditions. Oligonucleotide **1** (18 mer), 5' *pvprlacZ* anneals 42 bp upstream of -35 site, oligonucleotide **2** (31 mer), anneals directly upstream of the *vpr* atg start codon (12 bp), the oligonucleotide also contains a tail which will be utilised to anneal to the first 19 bp of the *lacZ*. Oligonucleotides **3 & 4** (K-12 l*acZ* forward 20-mer and reverse 19 mer) amplify the complete β-galactosidase gene (*lacZ*) annealing at both start and stop codon respectively. Creation of the fusion occurred by amplifying the fragment by PCR using oligonucleotides **1 & 4**. The design of the oligonucleotide primers means that the reporter gene is held in exactly the same orientation that *vpr* would be held in the chromosome. The fusion insert was cloned into pCRZeroBlunt (Invitrogen) and transformed into *E. coli* K-12 strain MC1061.

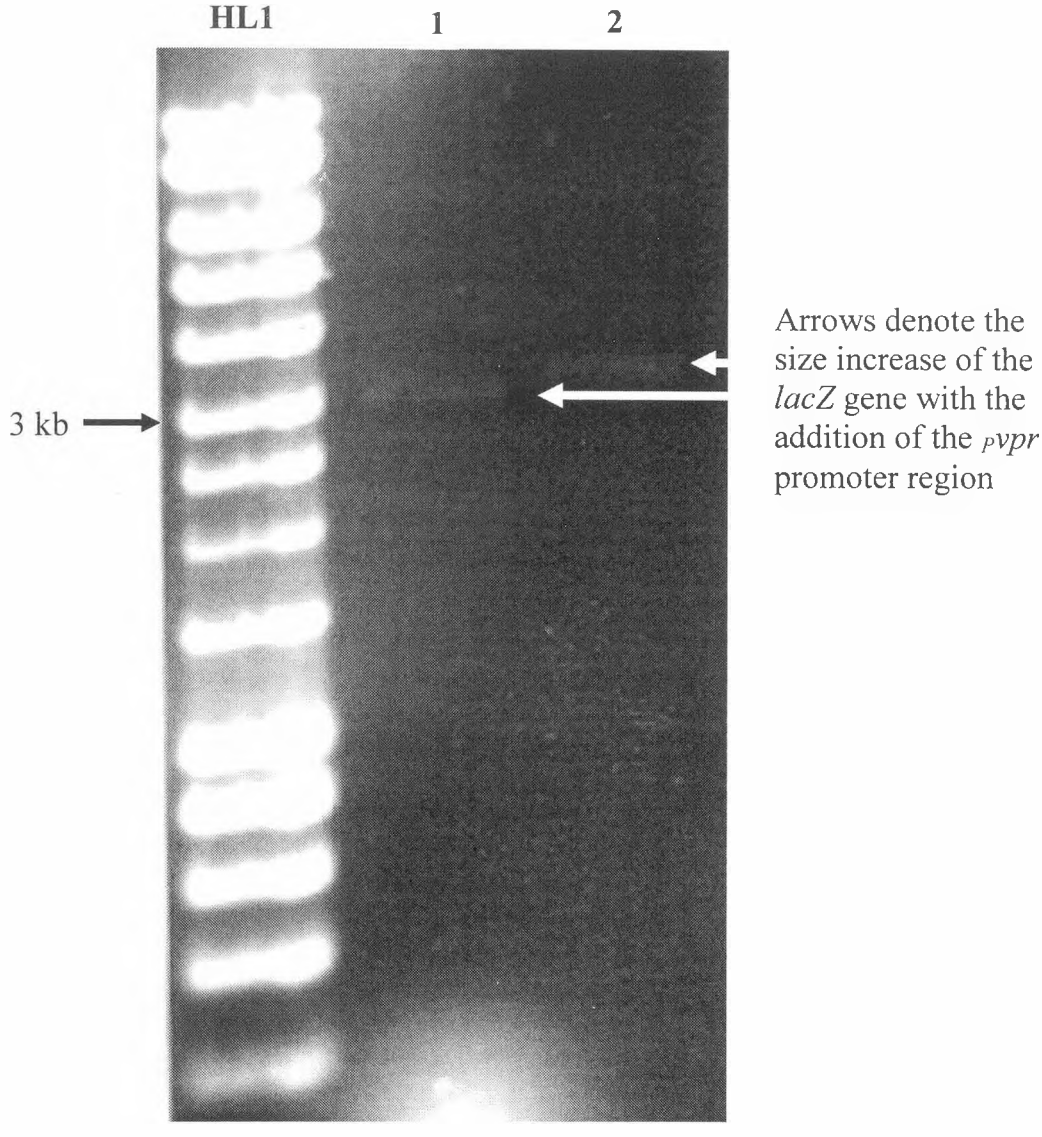


Fig 4.17 Confirmation of the *pvpr*::*lacZ* fusion. This shows that when the putative *pvpr*::*lacZ* (lane 2) was compared to the β-galactosidase gene PCR amplification product (lane1) that there is a shift in size that would denote the attachment of the *pvpr* region. Lane HL1; Hyperladder I

cloned into pCRZeroBlunt; sequence analysis confirmed the junction between $_{p}$*vpr* and *lacZ*. The objective of this part of the study was to determine if transcription levels change by altering environmental growth conditions, potentially to those associated with the mammalian gut environment.

4.4.9 Transcriptional regulation of *vpr* under discrete environmental growth conditions.

As previously discussed in section 4.2, Dartigialongue *et al.* (2001) made a similar reporter gene construct to that described here, although the orientation of the promoter to the reporter gene was such that the reporter gene was not in the identical orientation with respect to the *in situ* promoter-gene spacing. Using their $_{p}$*vpr* reporter gene construct ($_{p}$*ecfK::lacZ*) Dartigalongue *et al.* (2001) reported increased levels of transcription at 14 ^{o} C and 42 ^{o} C and so these were repeated along with growth at 37 ^{o} C, 37 ^{o} C + 1.5 g l^{-1} bile salts (Oxoid), 37 ^{o} C anaerobic growth (anaerobic gas jar, Oxoid). All incubations under aerobic conditions were performed in baffled 200 ml flasks containing 50 ml of LB plus associated supplements shaking at ~ 150 rpm. The data in Fig 4.18 show that there is variation in the levels of transcription under different growth conditions. The results show that increasing the growth temperature of *E. coli* from 37 ^{o} C to 42 ^{o} C results in a marked increase (>2-fold) in the level of transcription of the reporter gene. Under anaerobic growth conditions an increased level of transcription greater than the effect of increasing the temperature to 42 ^{o} C was observed. Although increased levels of expression were observed using $_{p}$*vpr*::*lacZ* at 42 ^{o} C, as reported by Dartigalongue *et al.* (2001), growth at 14 ^{o} C lowered expression (Fig 4.18), which is contrary to this previous

work. Physiological differences between *E. coli* strains (here, MC1061 cf. MC4100 used by Dartigalongue *et al.*, 2001) may be the explanation.

In these experiments all cells were washed twice in PBS before the reporter gene assay was performed to control for the inhibitory effect bile salts have on the enzymatic action of β-galactosidase. The addition of bile salts seems to attenuate the level of response of the regulation of the promoter. This attenuation can be seen at both the 42 ^{o} C and under anaerobic growth conditions. Phage 933W, the best characterised short tailed Stx-phage, which has an identical tail spike to $\Phi 24_B$ (see Chapter 5), is reported to infect *E. coli* through FadL a fatty acid transporter protein (Watarai *et al.* 1998). Watarai *et al.* (1998) showed that by increasing the osmotic stress, in the media, had a limitng effect on phage 933W infection. This concurs with the findings in this study as increasing osmolaric stress using bile salts lowers the expression of Vpr (Fig 4.18) and thus would lower infection rates. Watarai *et al.*, (1998) also inactivated *fadL* to produce a host that was resistant to phage infection, which could be restored by complementing the gene back into the phage-resistant strain. Knocking out a gene and looking for infection is not a true reflection of the protein – protein interaction between the tail spike and the outer membrane receptor; this can be only achieved by adsorption to determine if the phage can actually bind but not infect. Problems can also occur in mutation analysis, as the fluidity of the cell surface may vary, which could lead to the masking of the phage-binding epitope. Vpr is integral to the biogenesis of the outer membrane and, as part of its association with YfgL (a component of the YaeT outer membrane biogenesis complex), has been associated with the regulation of other outer membrane proteins such as LamB and OmpA (Wu *et al.*, 2005). For this association between OMPs, is possible that a masking effect, due to a change in the fluidity of the OM may explain why Watarai

et al., (1998) observed phage infection resistance when they inactivated *fadL*. Watarai *et al.* (1998) were unable to neutralise 933W by adding fractionated FadL from a 2 %-Triton treated preparation of the outer membrane, and were still able to recover 98 % of 933W present, which could not be explained. This study has identified Vpr as the outer membrane receptor for $\Phi 24_B$ and 933W as their host recognition proteins (tail spike) are identical. Therefore contrary to these findings of Watarai *et al.* (1998) Vpr appears to be the outer membrane protein associated with 933W and $\Phi 24_B$ infection.

Ruminant animals have a core body temperature of 39 ^{o} C, 2 ^{o} C higher than that in humans, also core body temperatures can be raised or lowered at different points of the day, increasing by 1.5 ^{o} C in the evening when livestock have been out to pasture (Piccione *et al.*, 2003). If these increased temperatures raised the display of the bacteriophage ligand on the cell surface *in vivo,* it could increase the probability of phage infection and in a 100 - 150 litre environment such as the bovine rumen this could lead to significant increases in phage infectivity and spread of *stx* genes throughout *E. coli* populations *in vivo* or in *animalia*. As *vpr* is essential and regulation can be altered by changing the bacterial growth conditions, this raises the question of whether Vpr is being expressed constitutively during the cell growth cycle and whether increased regulation of Vpr leads to increased adsorption of bacteriophage $\Phi 24_B$?

4.4.10 Expression of Vpr during the bacterial growth cycle.

If Vpr is constitutively expressed during the bacterial growth cycle and localisation and masking effects do not occur, it would mean that the ligand would be available for phage infection throughout the growth of the bacterial cell. We know

that this is not the case and fluidity of the bacterial cell surface can inhibit the phage from finding its recognition site. Also in stationary phase of the cell cycle the bacteria produce capsular proteins that also may inhibit phage infection. It may also be that expression varies throughout the bacterial growth cycle and optimum phage infection rates occur at specific points in the cell growth cycle. The data in Fig 4.19 is a representative data-set of the expression of Vpr by measuring promoter activity during the bacterial growth cycle. Fig 4.19 shows initially, that Vpr expression increases with growth at both 37° C and 42° C. The difference between the data sets (A and B) seen in Fig 4.21 is that at 42 ° C (B) the promoter activity is double that at 37° C, with the highest activity reaching ~ 300 Miller units per cell in mid-exponential growth phase. The 42° C data set does not cover the complete bacterial growth cycle, but it would be expected to follow the same trend as what was observed at 37° C growth in that levels of expression would increase through to stationary phase and then begin to drop away. Reporter gene assays can also cause other problems because β-galactosidase can accumulate in the cell during growth and can sometimes artificially indicate levels of expression beyond what is really present. Then main observation during these experiments is that expression occurs in the same trend, *i.e.* with growth, but because of the increased regulation due to the σ^E regulon promoter activity is doubled at the 42 ° C temperature.

4.4.11 Adsorption of bacteriophage related to growth of *E. coli* at 42 ° C

The data in Fig 4.18 show that by increasing the growth temperature by 5° C *E. coli* can double the transcription levels of *vpr*. Because of this increase in promoter activity the next natural question is whether this increased expression has a bearing on Vpr surface localisation, which would give the phage an increased chance

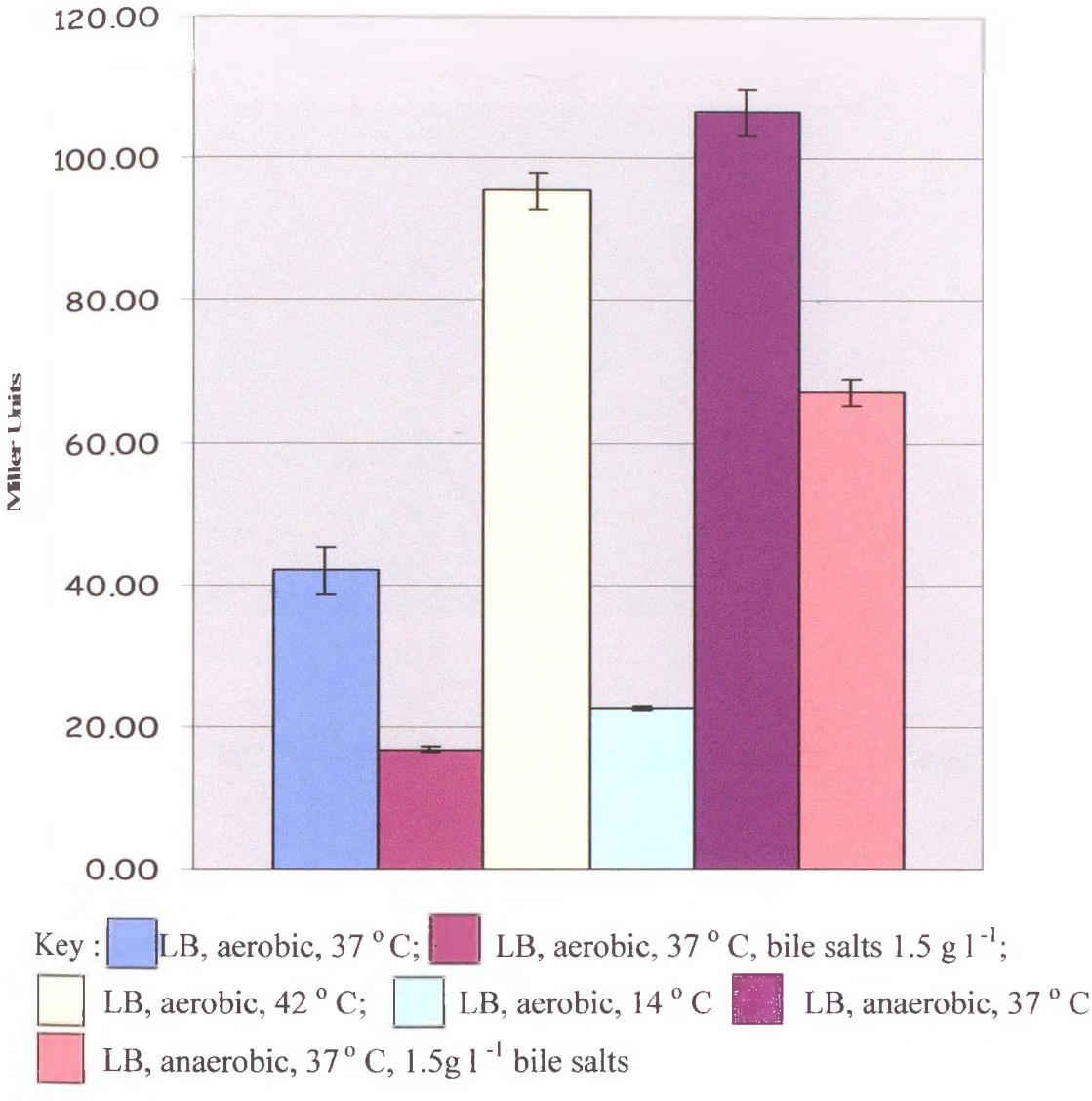
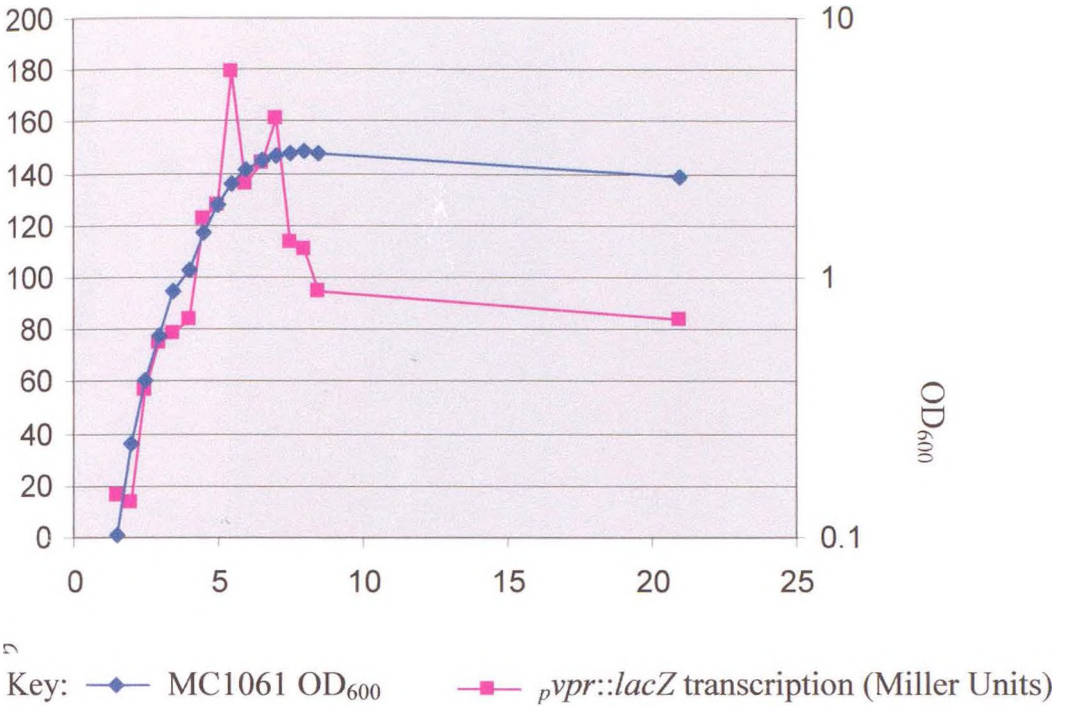


Fig 4.18 Transcriptional regulation of $_{p}vpr$ under different bacterial growth conditions using a reporter gene assay. MC1061 containing $_{p}vpr$:.*lacZ* construct grown under different conditions. Increasing the growth temperature from 37 ° C to 42 ° C more than doubled transcription of *vpr*. Anaerobic growth at 37 ° C increased transcription levels even greater than oberved at 42 ° C. Bile salts generally attenuated *vpr* transcription; this can be observed with growth at 37 ° C under aerobic / anaerobic conditions (plus bile salts) when compared to their untreated counterparts. Lowering the growth temperature to 14 ° C decreased the activity $_{p}vpr$ below that observed at 37 ° C, with aerobic conditions. Error bars = SEM, n = 9.

A)



B)

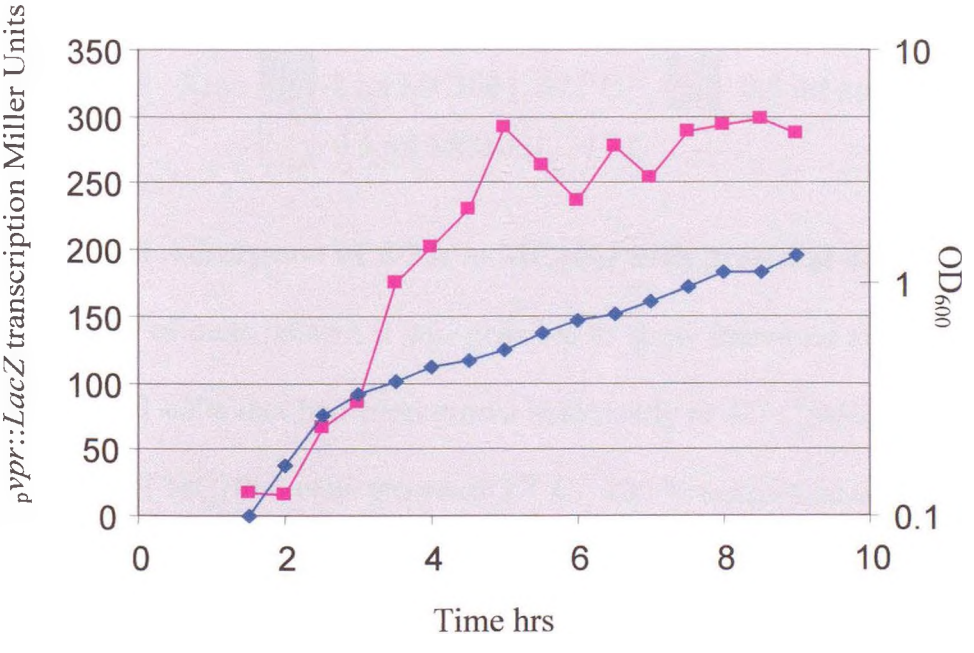


Fig 4.19 Expression of Vpr throughout the *E. coli* growth cycle. Using the $_{p}vpr::lacZ$ reporter gene construct it was possible to measure promoter activity throughout the cell growth cycle when grown at **A)** 37° C and **B)** 42° C. At the highest peak, there was almost double the level of β-galactosidase produced in cells grown at 42 ° C. Samples are means of duplicate samples.

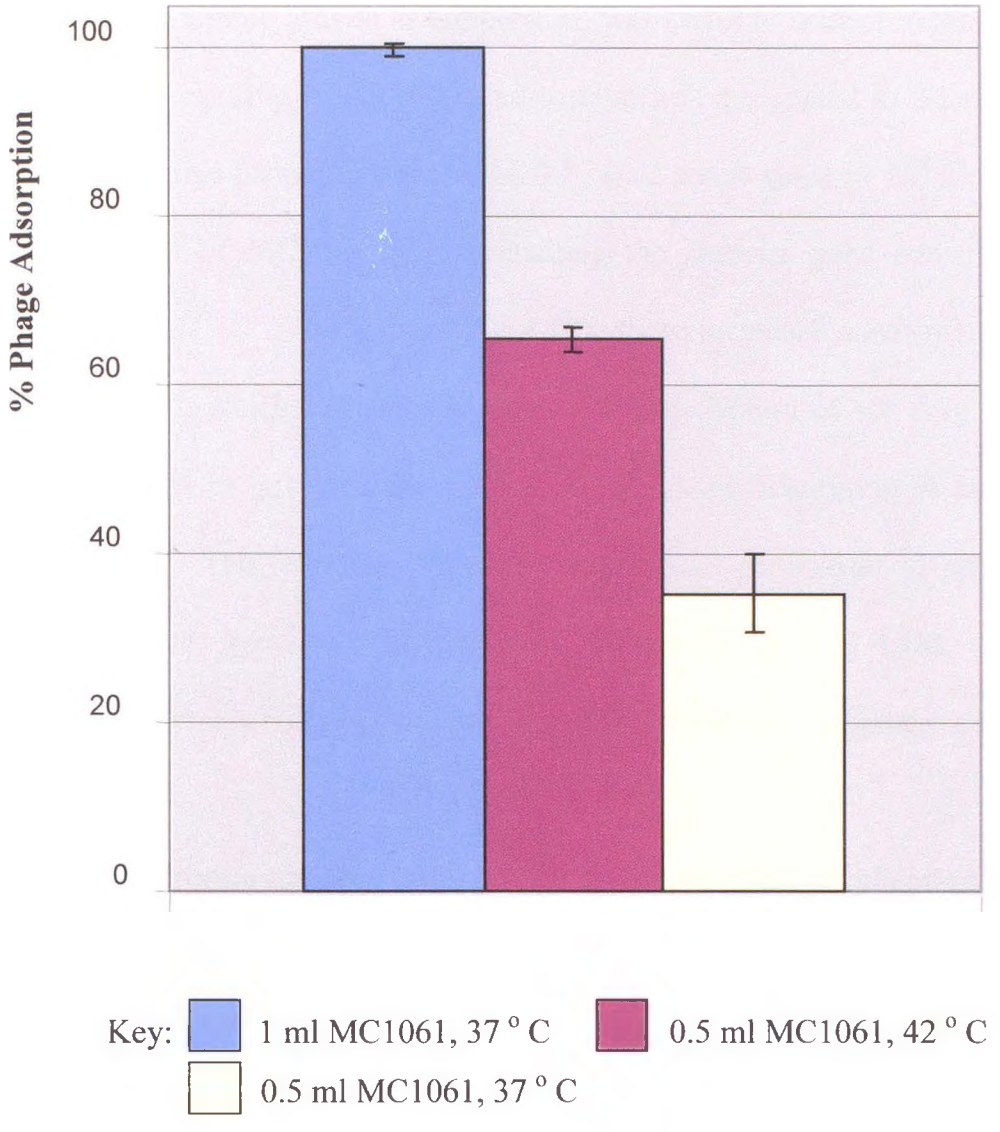


Fig 4.20 Adsorption of $\Phi 24_B$ to MC1061 cells grown at 42° C. By decreasing the number of cells present it was possible to show increased adsorption of $\Phi 24_B$ to an MC1061 cells that had been grown aerobically at 42° C when compared to the same OD_{600} of MC1061 cells grown at 37° C. 100 % adsorption was set as the adsorption assay data using 1.0 ml of MC1061 cells. n = 9, error bars = SEM

of being able to adsorb and infect its host. Growth of the *E. coli* at 42 ° C was chosen over anaerobic growth as temperature was easier to control in the laboratory. In this experiment (Fig 4.20), 100 % adsorption was designated as the reduction in the number of free bacteriophages when 0.1 ml of phage stock (~ 10^8 pfu. ml^{-1}) was added to 1 ml of MC1061 cells containing the reporter gene construct, grown aerobically at 37° C. These data (Fig 4.20), show increased adsorption to *E. coli* grown at 42° C, which correlates to increased transcription of *vpr* (Fig 4.19). Fig 4.22 shows that by growing the cells at 42° C phage adsorption is increased by almost 2-fold. This correlates with the increase in expression of *vpr* observed between MC1061 grown aerobically at 37 ° C and 42 ° C (Fig 4.18). These data suggest with increase in *vpr* expression there is increased bacteriophage adsorption.

4.5 Laser scanning confocal microscopy to demonstrate localisation of Vpr at the cell surface.

Whilst we know now that Vpr is localised at the cell surface the aim of this experiment was to further confirm localisation and to try and determine distribution across the bacterial cell surface. In brief, *E. coli* MC1061 cells were grown to mid-exponential growth phase (OD_{600} 0.55) and incubated with polyclonal anti-Vpr. The culture was then incubated with a secondary anti-rabbit IgG antibody conjugated with Alexafluor 433 (Molecular Probes). Initially a FITC-conjugated secondary antibody was used and there were distinct problems with background bleaching as evidenced by the FITC label diminishing within seconds of excitation. The Alexaflour labelled antibody is more stable than the FITC label under excitation and has a longer shelf life (4 ° C up to 1 year). Fig 4.21a (I) and 4.21b (I) show fluorescent signals at certain positions on the image.

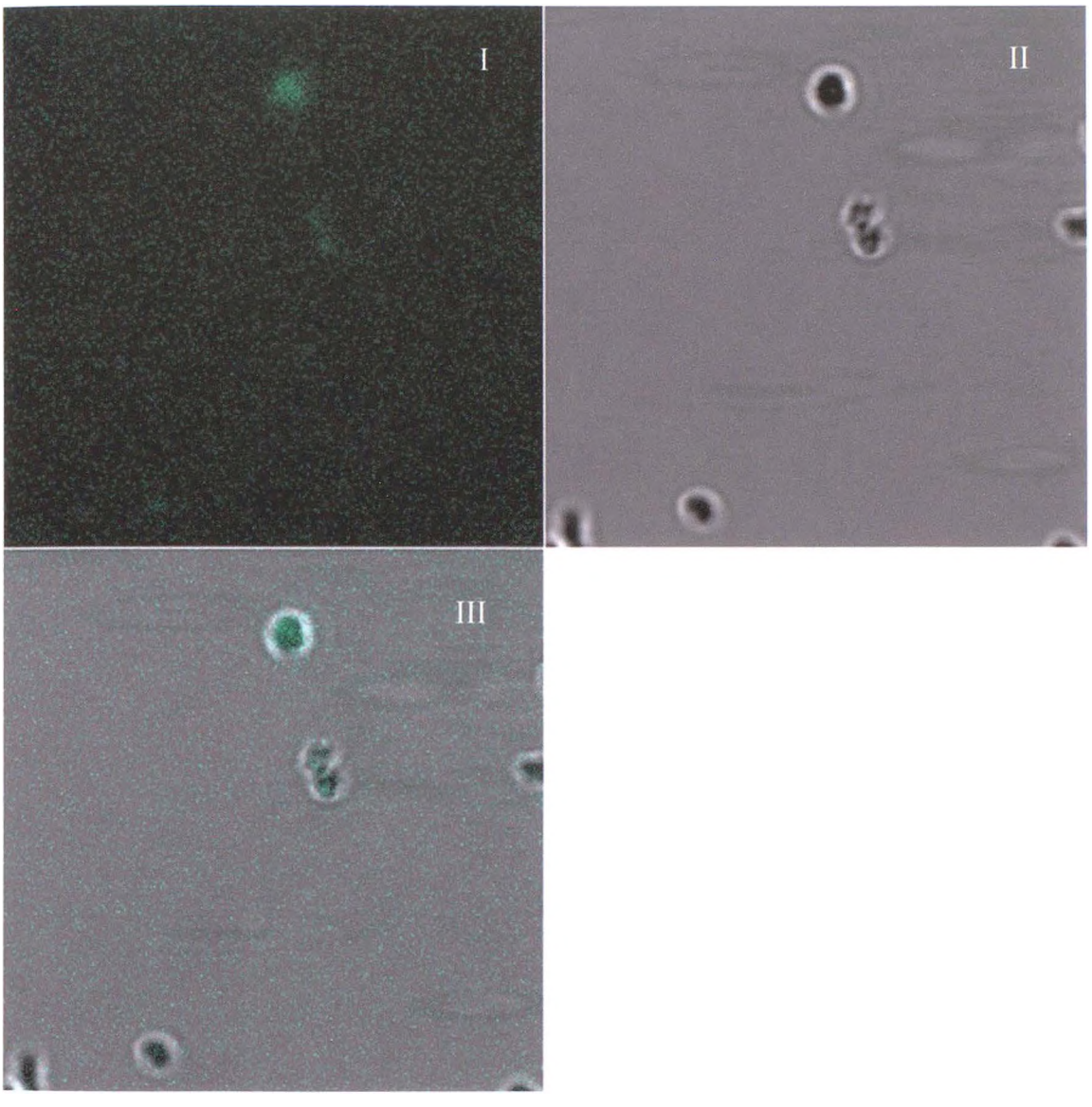


Fig 4.21a Confirmation of Vpr localisation on the cell surface of *E. coli* MC1061. Cells were labelled using a primary polyclonal rabbit anti-Vpr antibody, and the secondary antibody was anti-rabbit IgG conjugated with Alexaflour 433 (Molecular Probes). Image (I) shows immunofluorescence, image (II) shows the image by laser scanning confocal microscopy only and image (III) is an overlay of the immunofluorescence and the confocal microscopy image. This shows the prescence of immunofluorescent signal in the same areas that the cells occupy.

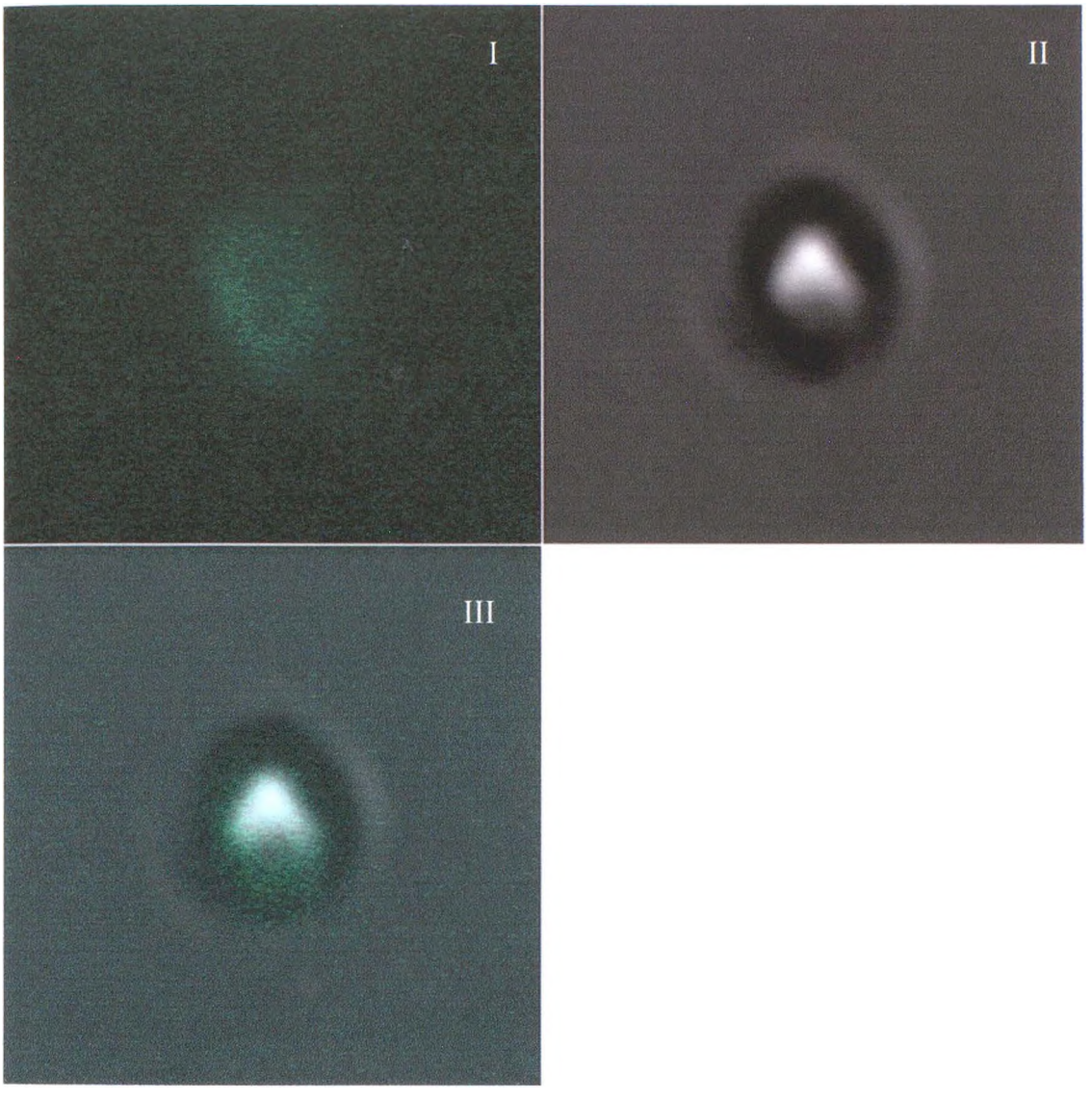


Fig 4.21b Confirmation of Vpr localisation on the cell surface of *E. coli* MC1061. Legend as Fig 4.21a. Image (I) shows classical halo ring effect of bound secondary antibody indicative of a protein that is located on the cell surface.

Fig 4.21a (II) and 4.21b (II) show the confocal microscope image, whereas 4.21a (III) and 4.23b (III) overlay the fluorescent and microscopy images showing that the immunofluorescence corresponds to the position of the bacterial cells. In Fig 4.21, the image analysis is concentrated on a single cell, and the immunofluorescence produced an almost halo effect which is indicative of the antibody adsorbing to an outer membrane protein (personal communication Dr. Dave Spiller). These images were very difficult to obtain as visualising an actively growing culture in mid-exponential growth phase was challenging due to the *E. coli* motility. This motility was subdued by placing the sample on a cover slip. Previous studies seem to show immunofluorescent detection of proteins that are in large amounts in the cell or specifically targeting DNA, such as rearrangements in the cytoskeleton (Fig. 1.12) during pedestal formation in EPEC/EHEC infection. Other studies, for example Tielker *et al.* (2005) have studied lectin (LecB) involvement in *Pseudomonas aeruginosa* biofilm formation by fluorescently labelling LecB. No previous study could be found with respect to determining distribution of an OMP at the bacterial cell surface.

4.6 Conclusions

- The *vpr* gene is essential and involved in the biogenesis of the *E. coli* outer membrane as it was not possible to inactivate the gene without bacterial cell death.
- Vpr is the bacterial ligand responsible for $\Phi 24_B$ adsorption and subsequent infection, as demonstrated by:
    - Inhibition of phage adsorption using an anti-Vpr polyclonal antibody.

- Transferring phage adsorption ability to a resistant host (*Erwinia carotovora* sbsp. *atroseptica*) by complementing the strain with the *E. coli vpr* gene.
- Increasing levels of Vpr production correlates with adsorption of 24_B.

- *vpr* is known to be regulated as part of the σ^E regulon that is responsive to environmental changes. The following were established here;
  - Increased Vpr production at 42 ^{o} C and under anaerobic conditions at 37 ^{o} C.
  - Vpr levels attenuated when grown in the presence of ox-bile.
- Vpr is localised at the cell surface, as demonstrated by laser scanning confocal microscopy.

Chapter 5 Bacteriophage-encoded factors that influence infection and host range

5.1 Other factors involved in bacteriophage infection.

In Chapter 4 infection was analysed from a bacterial standpoint, namely the identity of the host ligand Vpr in phage binding. $\Phi 24_B$ was used in chapter 4 mostly as a tool to monitor the expression of Vpr. Here, the focus is on factors that have a bearing on the infection process. Therefore, what do bacteriophages need to infect their designated hosts? In basic terms, what is required, the ability to adsorb to the cell surface, to inject its genome into the host and, with respect to temperate phage, the ability to integrate into the bacterial chromosome and subsequently to excise itself to produce viable virions that can infect other cells.

5.2 Tail Fibres

Phage tail fibres are assembled as the product of a single gene or cascade of gene products (see section 1.6). In long tailed bacteriophage, such as λ, the host specificity protein is encoded by a single gene product gpJ, although the construction of the tail itself is sequential using a cascade of gene products. These gene products not only help form the tail, but act as precursors for the next stage of production and assembly (Katsura, 1976). The tail spike or host specificity protein of long and short tailed phage must encode for both structure and host specificity. Hagard-Ljundquist *et al.* (1992) identified a high level of similarity at the carboxy-terminal of the protein between some phage (Mu, λ, P2) and hypothesised that they may therefore exhibit a similar host range. Sandmeier (1994) postulates that at some point bacteriophage such as λ, Mu, P1, P2 and T4 must have inhabited the same gene pool,

as regions of their tail fibres show distinct similarity in small regions. This theory concurs with the general view that genetic exchange occurs between phages in a gene pool. Rearrangements have been shown to occur in some phage tail genes where an inversion leads to a completely different distal end of the tail protein (Glasgow, 1989). Sandmeier *et al.* (1992) identified possible invertible regions in the tail fibre gene of bacteriophage P1 that could explain changes in the tail fibre conformation and thus contribute to the evolution of tail fibre diversity.

Originally it was thought that Stx-bacteriophage could only infect rough strains of Gram negative bacteria as they were thought not to be able to infect through the smooth lipopolysaccharide layer on the outer envelope (Smith *et al.*, 1983). For this reason it could be speculated that genetic exchange is not as prominent *in vivo e.g.* the gut as originally thought. However, James *et al.* (2001) showed that the bacteriophage used in this study ($\Phi24_B$) has a broad host range and is able to lytically infect both rough and smooth *E. coli* strains.

So why should we be interested in the tail spikes of toxin-encoding bacteriophage? Due to the high level of genetic exchange that could occur within the host, genetic recombination, including inversion could extend the host range of Stx phage. New host strains may be commensal or pathogenic, or even a different bacterial species *e.g.* within the *Enterobacteriaceae*. This not only has a bearing on phage evolution, but even more so on the evolution of bacteria.

5.3 Characterisation of the $\Phi24_B$ tail spike

$\Phi24_B$ has been classed as a member of the *Podoviridae* as it has a small stub-like tail spike, which in some transmission electron microscopy (TEM) images, would be hard to identify. A TEM image of $\Phi24_B$ is presented in Fig 5.1, and it

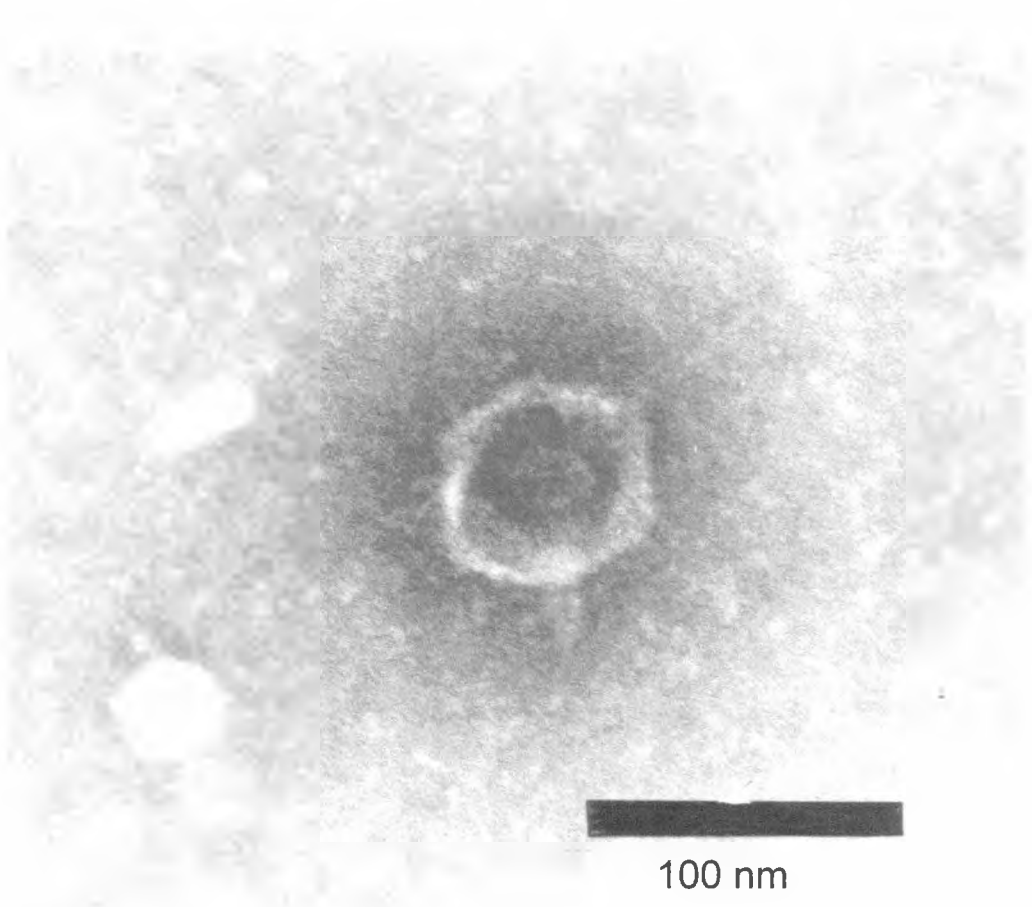


Figure 5.1 Transmission electron microscope image of $\Phi 24_B$ virion. ~ 1x 10^9 pfu ml $^{-1}$ of $\Phi 24_B$ were used to obtain this image of a single virion by Mr Brian Getty, University of Liverpool. The icosahedral capsid and short tail spike are clearly visible.

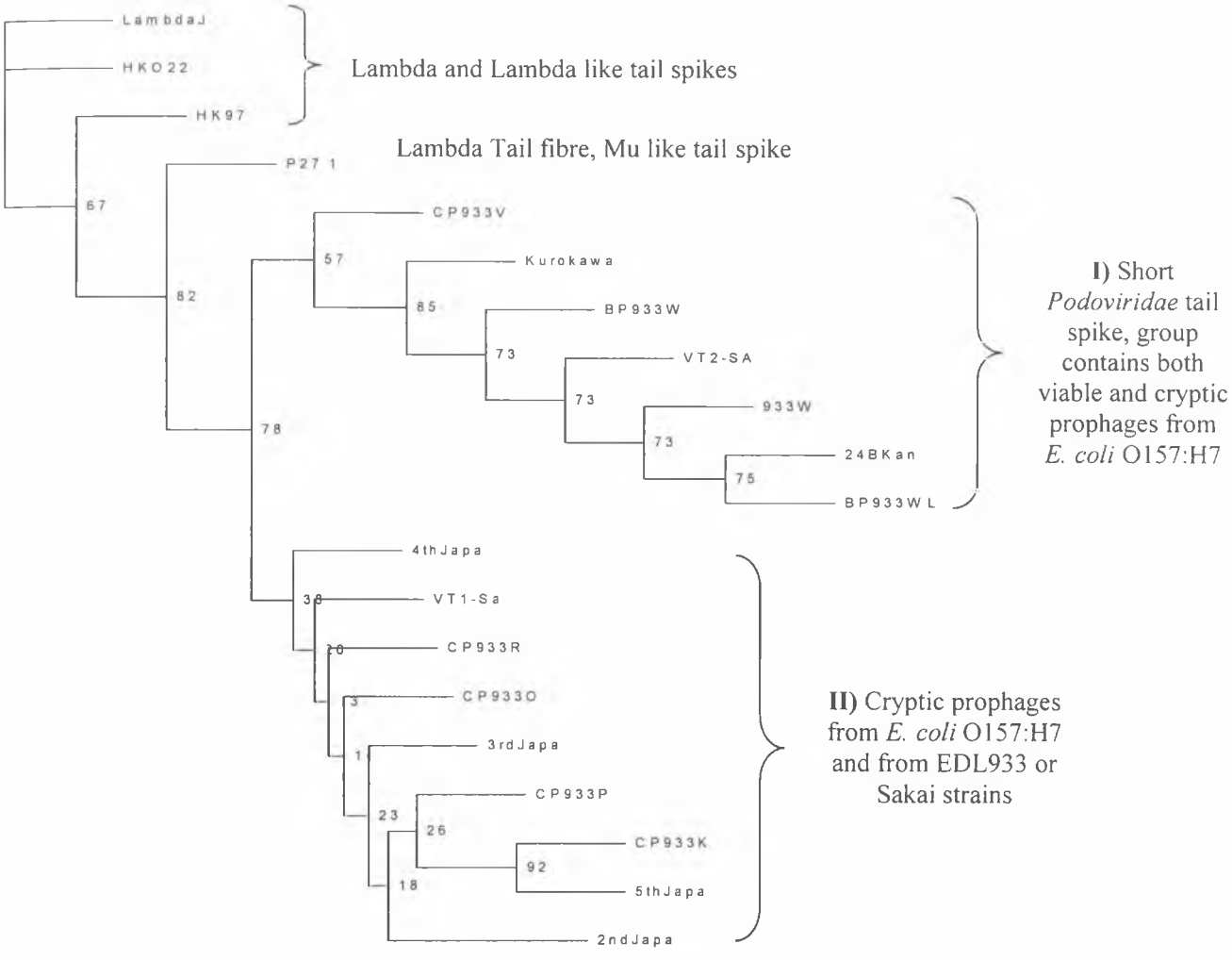


Figure 5.2 Representative protein cladogram of tail spikes from short and long tailed Stx-phage. Host recognition protein genes of different sizes were aligned using ClustalX and clades weighted using parsimony analysis in bioinformatic software package Phylip, derived from 100 bootstrapped analyses. This cladogram cannot be used to infer phylogeny due to the alignment of amino acid sequences of different lengths. Clade **I** includes tail spikes from cryptic prophages and viable short tailed phage *e.g.* 933W, whereas clade **II** possibly contains only cryptic tail spike genes.

shows a virion with an icosahedral head with a short tail. The bacteriophage tail spikes included in this study are generated from a single gene product (*e.g.* lambdoid-like Stx phage 933W (Plunkett *et al.*, 1999)). This differs from lambda phage where the construction of the tail fibre results from a cascade of genes in which each gene stimulates the production of the next so that the tail is formed in the correct manner. The *Podoviridae* tail spike must encode both the short tail structure and the host specificity region of the tail. Short tailed Stx-phages such as 933W and cryptic prophages from *E. coli* O157:H7 strains EDL 933 (933W lysogen) (Plunkett *et al.*, 1999) were aligned and a cladogram (Fig 5.2) constructed using parsimonius analysis using the Phylip bioinformatic software. Nucleotide sequence alignment of the fragments was achieved using ClustalX. At this point in time (2001), there were limited bacteriophage sequence data, so searches for Stx-phage tail spike genes were limited to 933W (Plunkett *et* al., 1999) lysogen EDL933 and *E. coli* O157:H7 (Sakai) sequences. These host recognition proteins were aligned along with the tail spikes of the archetypal λ, lambda like phage HK620 and Stx-phage P27.

For true phylogenetic analysis from gene alignment, regions of an identical size are required to correctly infer phylogeny. Due to the heterogeneous nature of the tail spikes between long tailed and short tailed phage it is therefore difficult to determine the phylogenetic relationship between genes. The cladogram presented in Fig 5.2 is therefore more a representation not to show phylogeny, but the sequence differences between the tail spike proteins found within sequenced O157:H7 strains EDL933 (Plunkett *et al.*, 1999) and the Sakai strain (Makino *et al.*, 1998) when compared to Lambda and Stx-phage P27 (Schmidt *et al.*, 2002). The genetic difference between these phages can be initially identified by gene size as Lambda's host recognition gene (~ 3.4 kbp) is significantly larger than that of the short tailed

Stx phage 933W (~1.9 kbp). The tail spike genes entered into this alignment are of different sizes so cannot be used to infer relationship or evolution of tail spikes. The cladogram (Fig 5.2) does not show the true phylogeny between the sequences aligned, but enables us to identify and group the tail spikes into two clades to which subsequently oligonucleotide primers were designed from the nucleotide sequence alignment to distinguish between the two clades. The two clades **I** and **II,** seem to separate the true cryptic phage tail spikes from those that, if incorporated into a viable virion are actually functional, although this would be speculation without further characterisation.

5.4 Amplification and nucleotide sequence analysis of $\Phi24_B$ tail spike

The alignment in Appendix 4 shows that the C terminus of the tail spike protein has high levels of amino acid sequence identity. The extent of this conservation enabled the design of a single reverse primer that would amplify short tail spikes from either clade I or II (figure 5.2). The oligonucleotide primer was named 3'-VTUTF (**V**ero **T**oxin **U**biquitous **T**ail **F**ibre), an 18-mer binding internally at the stop codon of the tail spike region. As discussed above clade I and clade II could be distinguished by designing primers at the 5' end of the alignment where significant heterogeneity is apparent. These oligonucleotides were detailed as VTTF(I) (clade I) and VTTF(II) (clade II) and by using them it was possible to amplify a region of approximately 1.9 Kb, using 2 µl of ~ 10^{8} pfu ml^{-1} $\Phi24_B::_{\Delta}cat$ lysate as template for the PCR (Fig 5.3). PCR amplification used an annealing temperature of 55 ^{o} C and extension time of 2 min, using recombinant non-proof reading DNA polymerase (MBI Fermentas). Primer pair 5' VTTF(I) and 3' VTUTF amplified a region of approximately 1.9 Kb from the $\Phi24_B$ lysate (Fig5.3). The amplification was repeated using a proof reading

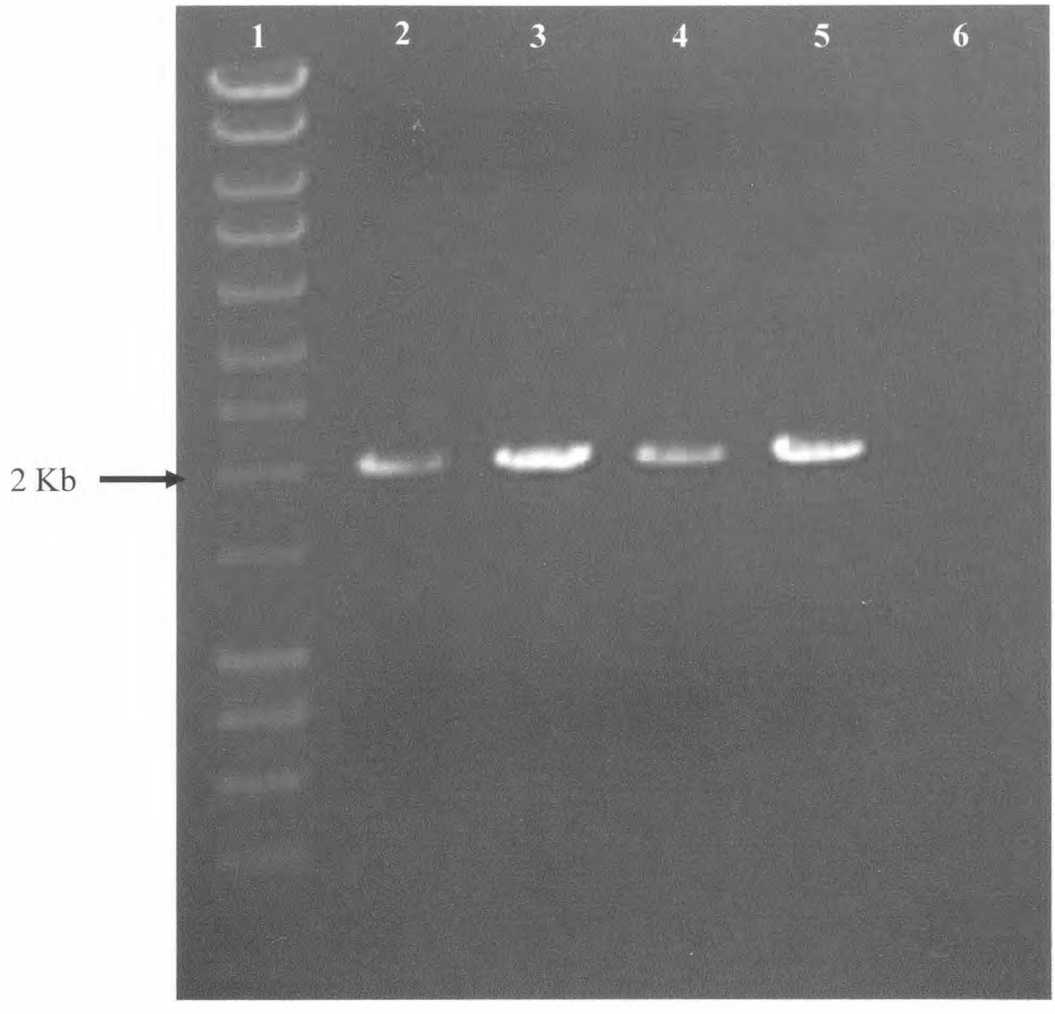


Figure 5.3 Agarose gel electrophoresis of PCR amplification of the $\Phi 24_B$ tail spike using primers 5' VTTF1 and 3' VTUTF. Lanes 2 and 4, 1 µl of ~10^9 pfu ml^{-1} of $\Phi 24_B ::_\Delta$cat and $\Phi 24_B ::_\Delta$kan respectively; Lane 3 and 5, 2 µl of ~10^9 pfu ml^{-1} of $\Phi 24_B ::_\Delta$cat and $\Phi 24_B ::_\Delta$kan respectively; Lane 1 was loaded with 5 µl of Hyperladder I (Bioline) and lane 6 was loaded with the negative control using (2 µl of phage buffer).

DNA-polymerase (*pfx* – Invitrogen) according to the manufacturer's instructions (100 µl reaction). The complete PCR product was loaded on a 0.75 % (w/v) agarose gel and electrophoresed at 80 v cm^{-1}. The gel was visualised under short wave ultra-violet light, and the PCR product band at ~ 2 Kb was excised and the DNA extracted using the Qiagen mini-prep DNA extraction kit (Qiagen). Fig 5.3 shows amplification of putative tail spike gene of $\Phi 24_B$ in lanes 2 – 5 yielding ~ 2 Kb PCR products.

The DNA extracted was ligated into the plasmid vector (~ 10:1 insert to vector ratio) as detailed in the Zero-blunt (Invitrogen) user manual. The manufacturer's calculation stipulated the addition of 140 ng of the 2 Kb blunt-ended DNA insert to 35 ng of linearised Zeroblunt plasmid (Invitrogen). 2 µl of this ligation mixture was transformed into chemically competent *E. coli* Top10 cells (Invitrogen), and incubated overnight on LB agar plus kanamycin (50 µg μl^{-1}). Putative clones were cultured to mid-exponential growth phase and plasmid extracted using Qiagen mini-prep kit, according to the manufacturer's instructions. The insert was then sequenced from the plasmid using M13 forward and reverse primers provided by the manufacturer (Invitrogen) and sequenced by MWG Biotech. The sequences were orientated into the correct reading frames and nested primers designed to the initial sequence data so that the complete gene could be sequenced. Both primers anneal approximately 500 bp downstream of the start codon (5' VTTF3) of the tail spike and 600 bp upstream of the stop codon (3' VTTF3), seen in table 2. 3. This internal region of the tail spike gene was amplified and yielded a PCR product of ~ 800 bp (Fig 5.4). These blunt-ended PCR products was cloned into sequencing vector Zeroblunt (Invitrogen) and sequenced using M13 forward and

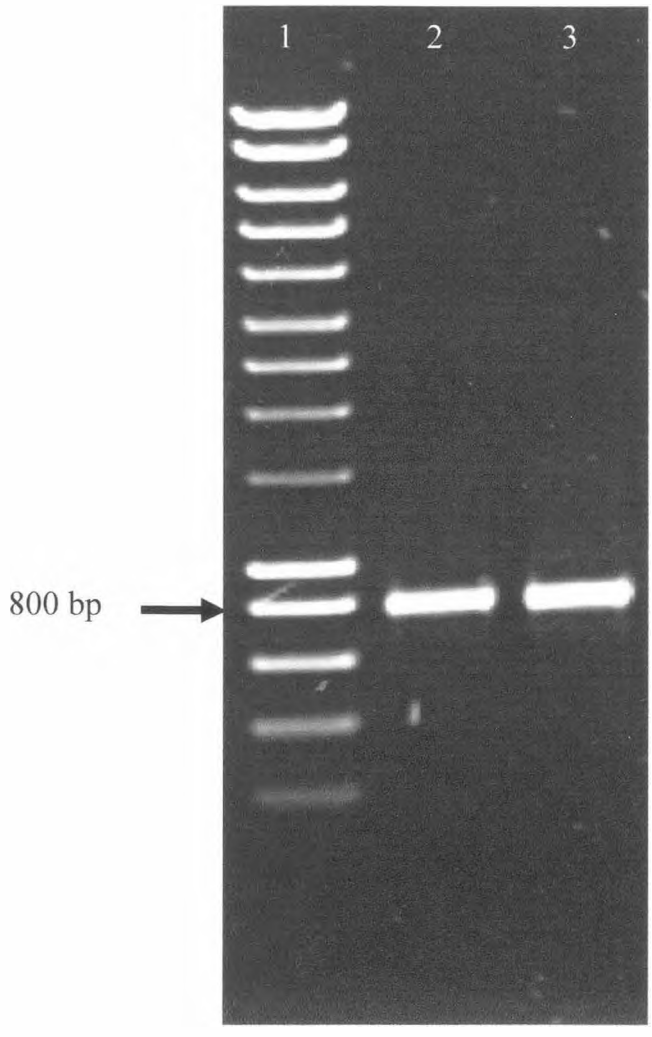


Figure 5.4 PCR amplification of nested region of $\Phi 24_B$ for sequence analysis. Oligonulceotide primers 3' VTTF3 and 5'VTTF3 (Table 2.3) were used to amplify, by PCR, the inner region of the $\Phi 24_B$ tail spike gene (~800 bp) using a proof reading enzyme *pfx* (Invitrogen). Lane 1; hyperladder I, Lane 2: PCR product, template 1 µl of $\Phi 24_B$ lysate, Lane 3: PCR product, 2 µl of $\Phi 24_B$ lysate.

reverse primers which anneal either side of the multiple cloning site (Invotrogen). Reverse sequences were orientated into the correct reading frame (*i.e.* reverse complemented) and all 4 sequences were aligned by eye to form a consensus sequence. The sequence of the $\Phi 24_B$ tail spike was then blastn searched through the NCBI website and found to be identical at the nucleotide level to short tailed Stx-phage 933W host recognition gene (Plunkett *et al.*, 1999). As $\Phi 24_B$ has an identical host recognition protein to Stx-phage 933W it would be fair to assume that it would be able to infect the same host range. $\Phi 24_B$ has been shown to have a broad host range and (James *et al.*, 2001), which may correlate with the host range of 933W. This broad host range, due to the host recognition of the phage, seems quite well conserved in the short tailed Stx-phage characterised thus far. This could lead to transfer of Stx-genes to sensitive bacterial strains, which, when in conjunction with a broad host range, could have a significant impact on toxin production during infection as demonstrated by Gamage *et al.* (2003) who showed that non-pathogenic bacteria have an influence on toxin production in the gut.

5.5 Further analysis of tail spike gene alignments

The availability of phage gene sequences in general has increased dramatically over the past 3-4 years. Now in the almost post-genomic era, increasing numbers of bacterial genomes have been sequenced and shown to contain many bacteriophage genes of both viable and cryptic phage.

Annotation problems within sequence databases such as Genbank have led to difficulties in searching for phage genes using specific gene names. Different annotations of tail fibre genes includes "tail gene", "tail spike", "putative tail gene", "tail" and "prophage related sequence". It is often unclear which of these genes are

for host recognition and which are involved in the construction and assembly of the tail fibre itself. Some of the longer tailed *Siphoviridae* are easier to annotate due to the larger number of genes used for construction of the tail, and therefore the correct gene can be selected from the bacteriophage genome. Due to the heterogeneous nature of the host recognition protein of lambdoid-like phages it was necessary to draw 2 phylogenetic trees, separating the long and short tail spike genes from blastp search of Genbank, and identifying Stx phage or cryptic phage in STEC strains using the Lambda and 933W host recognition protein sequences as blastp entries (Fig 5.4 and 5.5). Fig 5.5 and 5.6 show phylogenetic analysis of both long- and short-tailed host specificity proteins that were determined using the protpars parsimony algorithm from bioinformatic software ARB. They cannot be used to infer phylogeny due to the variation in sequence lengths caused by the presence of cryptic phage genes. The cladograms give an indication of the level of diversity amongst sequences that are found when blastp searching for the tail spikes of λ and 933W. To align the tail spike proteins of all members of the *Podoviridae* or *Siphoviridae* would be difficult and incorrect as each of these families comprise bacteriophage that transverse a range of tail morphologies. The blastp search for the host specificity proteins of λ and 933W is a partial answer to this problem as they are the best studied lambdoid phages that are found in genomes of *E. coli* O157:H7 strains characterised thus far. To infer the phylogeny of phage tail spike genes a larger sample number could make it possible to select a region of the tail spike genes that is conserved. Masking or selecting conserved regions would offer a tool to align only the conserved regions of the protein. Cryptic phage tail genes appear to heavily influence the results of any phage gene search and are important to bolster the available data. The disadvantage of using genes identified from cryptic prophage is

that it is unclear whether the tail spike is functional. All sequences used in Fig 5.5 and 5.6 contain both a stop and start codons, denoting the open reading frame of the tail spike gene. The cladogram in 5.4 shows protein sequences that were identified by blastp using the λ host recognition protein as the template for the search. The dendrogram representing the long tailed (Fig 5.5) shows 2 clades (I and II), clade I groups a number of viable phage tail spikes, whereas clade II is possibly all cryptic phage genes. The cladogram in Fig 5.6 shows protein sequences that were identified by blastp searching the Stx phage 933W host recognition protein. This also can be subdivided into two clades which could show the difference between viable and non-viable tail spike genes. Clade I mainly contains phage genes that are located on cryptic prophage, whereas clade II identifies the only tail spike genes apart from the out-group that have been identified on viable phages.

There are 284 sequenced bacteriophages and partial sequences for many more. As more sequence data becomes available we will be able to better understand tail spike and definition of conserved and variable regions and their bearing on infection.

5.6 PCR Amplification of tail spike genes isolated from phages induced from wild-type STEC strains.

From an established study site 23 STEC isolates were offered as part of a DEFRA horizontal study of cattle shedding of STEC in a well described study site on the Wirral (Robinson *et al.*, 2004)

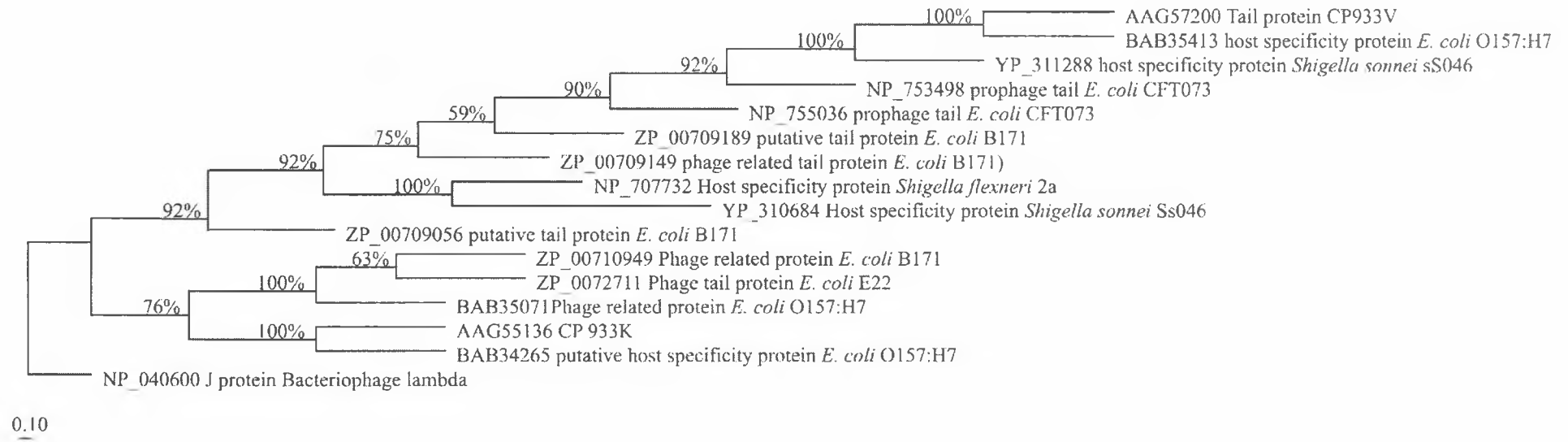


Figure 5.5 Parsimony cladogram of all phage sequences identified by blastp analysis of the λ host specificity protein. The tree was derived protpars algorithm of Phylip with branches bootstrapped from 100 analyses. Scale bar equal to 1 amino acid change per 100 amino acids.

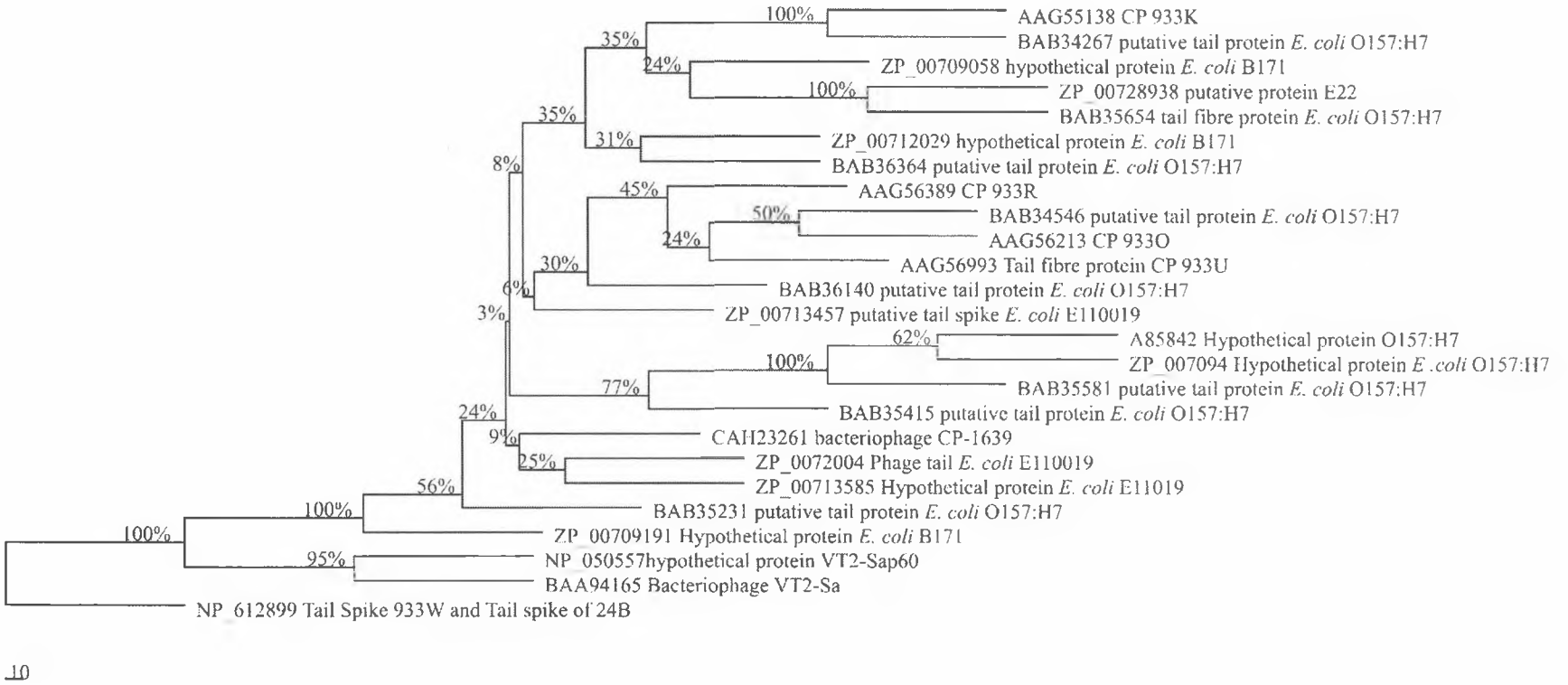


Fig 5.6 Parsimony Cladogram of all phage sequences identified by blastp analysis of the 933W host specificity protein. The tree was derived by prtotpars algorithm of Phylip with branches bootstrapped from 100 analyses. Scale bar equal to 1 amino acid change per 100 amino acids.

2005). Phages were induced using norfloxacin (Section 2.2.4) and propagated in *E. coli* DM1187, the *recA* mutant (section 2.1.2). The STEC strains were selected as a representative of the strains on offer using toxin type, serotype, sampling data to choose the representative isolates. Table 5.1 shows that a range of serotypes and pathogenicity profiles are represented. Table 5.1 also shows that the only Stx-phage that was isolated, using *E. coli* DM1187 as host, was ΦDS15. Of the 23 STEC strains that were induced using NFLX, phages that could infect DM1187 were r*EcoR*ded from 13 strains. From these 13 phages isolated only 1 encoded the Stx toxin. The oligonucleotide primers designed to identify and sequence the tail spike gene of $\Phi 24_B$ (Table 2.3) were used to screen these induced phages for related tail spike genes. Using the outermost primers, 5' VTTF1 and 3' VTUTF, it was demonstrated that tail spike genes could be amplified for bacteriophages ΦDS4, ΦDS5, ΦDS6 and ΦDS7. The ΦDS5 had a tail spike gene that was ~ 400 bp smaller than the other tail spike genes identified with this primer pair (Fig 5.7). Whether this is actually a viable tail spike gene or a remnant that is harboured by the phage is unclear.

As phage genes can have a divergent modular nature, the internal sequencing oligonucleotide primers were used to determine if they could identify any other regions of the 13 phages that were isolated. Using these primer sets, it was also possible to identify sequence divergence by using different combinations of the primers *e.g.* 5' VTTF1 and internal oligonucleotide primer 3' VTTF3. The other 10 bacteriophages that could not be identified as containing tail spike genes related to $\Phi 24_B$ remained negative although 5' VTTF – 3' VTTF3 amplified a region of the tail spike of ΦDS15 to produce a PCR product of expected size (~1.2 kb). This shows

that even though PCR based approaches are rapid, there is a possibility of missing target sequences due to heterogeneity at the oligonucleotide annealing site.

5.7 Immunity and super-infection

Integration of the phage genome into the bacterial chromosome and its regulation is described in section 1.2. When bacteriophage λ is integrated into its *E. coli* host genome it is held in its integrated state by the presence of cI, the phage lysis repressor (Ptashne, 1992). As there are approximately 100 active molecules of this repressor present at any one time, they actively bind the regulatory regions of the incoming λ immediately turning off the genes responsible for lysis (Ptashne, 1992). According to the lambda model, the lysogen is thus resistant to infection by another genetically identical λ phage. There are examples of bacteriophages such as λ that integrate in tandem into a single site on the bacterial chromosome (Kholodii *et al.*, 1985). It was previously thought that all lambdoid-like phages conform to this λ immunity model. Allison *et al.* (2003) was able to infect single *E. coli* cells with two identical phages ($\Phi 24_B$) marked with two separate antibiotic resistance markers. The first requirement was to prove that complete phage genomes were indeed integrated into separate insertion sites on the bacterial chromosome, and this is described in figure 5.8. Figure 5.8 shows the analysis of both the single and double lysogens restricted with *EcoRI* and *AflIII*. By probing the restricted lysogens with DIG-labelled DNA probes for *aph3* (Kanamycin resistance gene), *cat* (chloramphenicol acetyl transferase gene), *Q* and the stx_2B genes it was possible to show that $\Phi 24_B::_{\Delta}$cat and $\Phi 24_B::_{\Delta}$kan are inserted into different positions in the bacterial chromosome. In Fig 5.8 the Southern blots A – D lanes 1, 3 and 5 contain *EcoRI* digested DNA, lanes 2, 4 and 6 contain *AflIII* restricted DNA. Using these restriction

patterns and Southern hybridisation analysis, it was possible to determine the presence of *E. coli* that had been infected by 2 isogenic bacteriophages. Southern analysis using probes specific for the antibiotic resistance genes carried by the two phage show that they each identify their respective integrated hosts. The Southern analysis also demonstrates that by blotting with *Q* and stx_2B (C and D) it is possible to overlay the blotting pattern of the single infection over the double infection to show two distinct phage genome copies integrated into the bacterial chromosome. Because the Southern blotting patterns between the single and double infected hosts are so different, and as the bands do not overlap, it demonstrates that the phage has not integrated in tandem, but has integrated into a different position in the bacterial chromosome. The ability of multiple phage infection has been identified in *Salmonella*-specific bacteriophage P22; this has an *immI* gene region that encodes an anti-repressor (Ant) that has the ability to allow infection of a secondary phage (Susskind and Botstein. 1975). Apart from the integrase gene, all the genes sequenced thus far from $\Phi 24_B$ have shown high levels of sequence identity with Stx phage 933W. Communication with Friedman, University of Michigan (personal communication) located remnants of the P22 immunity region located in the capsid genes of 933W and that 933W was unable to form double lysogens. $\Phi 24_B$ DNA has been sent to the Wellcome Trust, Sanger Institute for genome sequencing and this will determine if $\Phi 24_B$ has an intact *immI* region. So what is the impact of isogenic phage being able to infect a single host? With respect to virulence, if the copy number of *stx* present in a single host is increased the level of toxin production increases which has serious implications on the sequelae of STEC infection.

Table 5.1 Characteristics of Wild-type STEC strains used for phage induction and tail spike gene characterisation

Strain Name	CCN	*eae*	*Stx1*	*Stx2*	VLA no.	O-Serotype	Origin	Phages Isolated	Tail Fibre Regions Identified	Viable Phage encoding Stx-toxin
DS1	1815	+		+	1860	Rough	Water			
DS2	2912			+	1861	UT	Cattle			
DS3	3016			+	1862	O15	Cattle			
DS4	3171	+	+		1863	O98	Cattle	ΦDS4	+	
DS5	3268			+	1864	O104	Cattle	ΦDS5	+	
DS6	3311	+		+	1865	UT	Cattle	ΦDS6	+	
DS7	3327			+	1866	UT	Soil	ΦDS7	+	
DS8	3328	+		+	1867	O145	Soil			
DS9	3660			+	940	O135	Cattle	ΦDS9		
DS10	3661	+		+	1869	O46	Cattle			
DS11	3788			+	1870	O2	Cattle			
DS12	4006			+	1871	O71	Cattle	ΦDS12		
DS13	4007	+		+	1872	O43	Cattle			
DS14	4159			+	1874	UT	Cattle	ΦDS14		
DS15	4228			+	1875	O73	Cattle	ΦDS15	+	+
DS16	4546			+	1879	UT	Rabbit	ΦDS16		
DS17	4672	+		+	1882	O88	Soil	ΦDS17		
DS18	4844			+	1883	O168	Cattle			
DS19	5018			+	1884	O8	Soil	ΦDS19		
DS20	5223			+	949	UT	Cattle	ΦDS20		
DS21	5573	+	+		954	Rough	Cattle			
DS22	5662	+	+		956	O84	Cattle			
DS23	7267		+		1894	UT	Cattle	ΦDS23		

CCN = culture collection number

VLA = Veterinary Laboratory Agency number

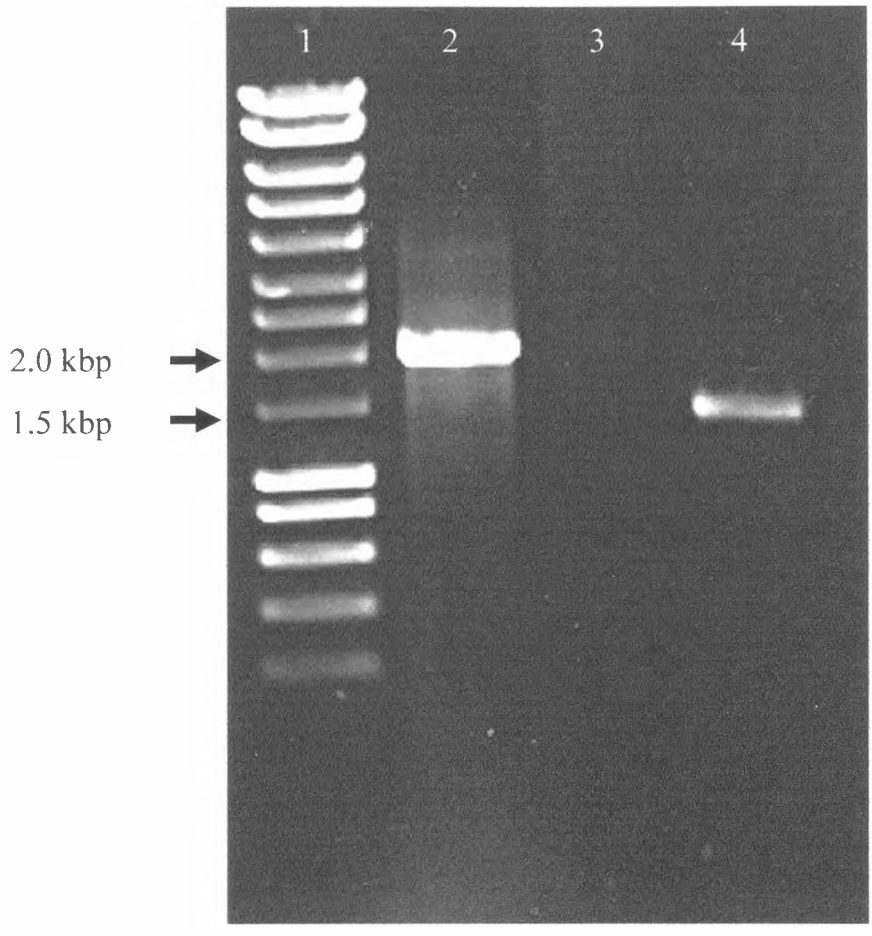


Fig 5.7 PCR amplification of the tail spike gene of bacteriophages ΦDS4 and ΦDS5 using oligonucleotide primers 5' VTTF1 and 3' VTUTF. Lane 1; Hyperladder I (Bioline), Lane 2; PCR product using 2 µl of ΦDS4 phage diffusion preparation (see section 2.2.5) as template, Lane 3: negative control, template 10 ng of DM1187 genomic DNA, Lane 4; PCR product using 2 µl of ΦDS5 phage diffusion preparation as template. Image shows ~ 400 bp difference between the sizes of tail spike genes of ΦDS4 and ΦDS5.

5.8 Identification of the $\Phi 24_B$ Integrase gene

As $\Phi 24_B$ has been shown to have similarity to Stx phage 933W in regions such as the tail spike (section 5.3), immunity region (personal communication Allison, James and Gossage, unpublished data). Primers specific for the 933W integrase gene were designed to determine if they could be used to amplify the integrase from $\Phi 24_B$. None of these oligonucleotide primers specific for the 933W integrase gene would amplify any product using $\Phi 24_B$ as template. All $\Phi 24_B$ genes sequenced thus far have shown high levels of sequence identity to 933W (Allison, James Gossage, Smith unpublished data), but this suggests that $\Phi 24_B$ has an unrelated integrase gene to 933W. Integrase is a tyrosine recombinase that is involved in the site-specific integration of the phage genome into the bacterial chromosome (section 1.5). It is not a good marker to use for bacteriophage classification due to its heterogeneous nature (Balding *et al.*, 2005). Alignment of sequenced short tailed Stx-bacteriophage integrase genes (in 2002) using ClustalX made it possible to design primers (5' 933Wint and 3' 933Wint, see Table 2.3) to amplify the integrase region, based on anticipated similarity to 933W. A more accurate approach would have been to create a DIG-labelled DNA probe by PCR using from the 933W integrase and probe against $\Phi 24_B$ by Southern hybridisation. Unfortunately, at that time propagation of 933W was not possible due to the dismantling and moving of the containment level III facility.

Subsequent PCR amplifications using the 933W-specific integrase primers (Table 2.3) failed to yield PCR products from high titre lysate of $\Phi 24_B$ (~ 10^8 pfu ml^{-1}) as template. In this period Balding *et al.*, (2005) designed a bank of primers to amplify all sequenced bacteriophage integrase genes identified thus far. These integrase identification primers (Balding *et* al., 2005) were designed to amplify the

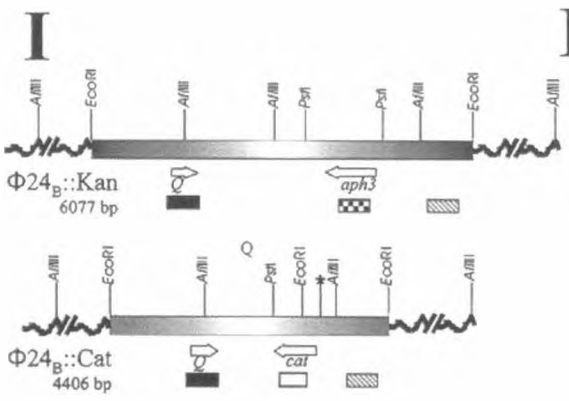
conserved regions of the integrase that govern function and thus be able to group phage on integrase diversity. Bacterial chromosomal contamination of the phage PCR template was negated by the failure to amplify the *E. coli* house keeping gene GAPDH (Table 2.3) (Fig 5.6). The bank of primers yielded a PCR fragment of expected size for integrase group 6 (~1300 bp for group 6a and ~ 250-300 bp for group 6b). Oligonucleotides specific for integrase groups 6a and 6b, offer coverage of the integrase genes aligned in that group, hence the difference in expected PCR product sizes. Amplification of the $\Phi 24_B$ integrase gene using these group-specific oligonucleotides and subsequent sequencing, related this Stx-phage integrase to that found in bacteriophage HK620 a short tailed λ-like phage and the integrase gene found in P4 and *Shigella flexneri* phage Sf6. This is the first Stx-phage to harbour a group 6 integrase gene. As more bacteriophage sequence data becomes available, it reveals the high level of recombination that occurs between phage to form chimeric virions. Therefore, finding an integrase gene that has not been previously identified in Stx-phages is not surprising. Current research in this laboratory has shown that $\Phi 24_B$ may integrate into the same point in the bacterial chromosome that HK620 integrates into. Conflicting data about grouping of the $\Phi 24_B$ integrase will be only clarified on completion of the phage genome sequencing. Characterisation of other wild type bacteriophage integrase genes has been found to be troublesome as bacterial chromosomal background in the PCR template preparation can hamper the true identity of the phage genes in question.

5.9 In summary

- $\Phi 24_B$ has an identical tail fibre to virulent wild-type Stx phage 933W.
- There are high levels of tail spike heterogeneity at both nucleotide and protein sequence level between phage families, although a conserved carboxy terminus has been identified in the short tailed Stx-phage.
- A high level of sequence divergence and difference in tail spike gene sizes makes it impossible to infer phylogeny due to low levels of homogeny between phage families.
- The levels of phage genes annotated in Genbank are rapidly increasing although the quality of annotation can hinder gene searches on the NCBI website, and if the tail is part of a multi-gene assembly cascade it becomes difficult to determine the host specificity protein.
- Using the primer sets designed to amplify the tail spike of $\Phi 24_B$ it was possible to identify related tail spikes in 5 phages isolated from 23 environmentally isolated STEC strains.
- PCR approaches for identifying specific genes from phage is inherently difficult due to the level of cryptic phage and cryptic genes being present in PCR template preparations.
- Identification by PCR can be problematic as oligonucleotides only anneal to a short length of DNA, meaning that significant alterations in this region can limit the sensitivity of identification.
- $\Phi 24_B$ does not conform to the λ immunity model as it is possible to infect a single bacterial host with 2 isogenic bacteriophages that have been labelled with different antibiotic resistance genes. This has great implications in

STEC infection as multiple copies of *stx* can lead to increased production of the toxin.

- $\Phi 24_B$ has an integrase gene that is related to bacteriophages HK620, P4 and Sf6, although this will be clarified subsequent to completion of the $\Phi 24_B$ genome sequencing. This integrase difference between $\Phi 24_B$ and 933W demonstrates that even though other gene loci show high levels of sequence identity, modular recombinational events can occur and the identification of a completely different integrase gene has been found in $\Phi 24_B$.



II

Recombinant Phage	Restriction Enzyme	Fragment Size (bp)
$\Phi24_B$::Kan	*Eco*RI	6072
	*Afl*III	6163
		1414
		2336
		7529
$\Phi24_B$::Cat	*Eco*RI	1375
		3026
	*Afl*III	6163
		2079
		7529

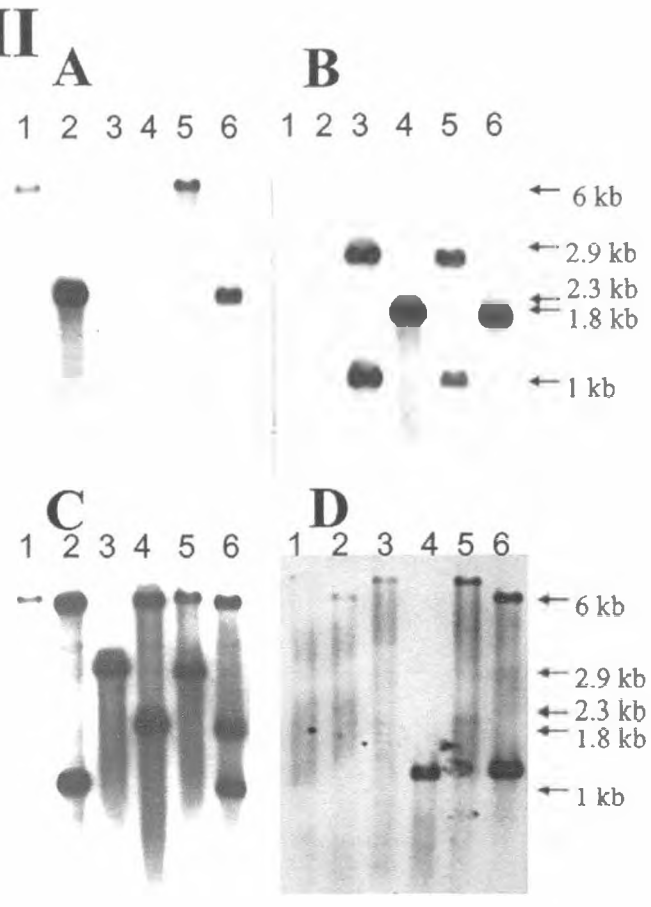


Figure 5.8 Confirmation of the residence of two complete toxin operons within the double lysogen genome. I, Genetic maps showing the location of relevant restriction endonuclease sites and areas recognised by the relevant probes: *aph3*, ▩ ; *cat*, □; *Q*; ■; *subvt*$_2$*B*, ▨. The asterisk, *, indicates the position of a *Pst*I site that was lost following a single nucleotide substitution, as confirmed by DNA sequencing. II, Tabulated data of the RFLP lengths of the relevant *Afl*III and *Eco*RI fragments from each isogenic recombinant phage. III, Southern blots using probes specific to: A, *aph3*; B, *cat*; C, *Q*; D, *subvt*$_2$*B*. Lanes 1, 3 and 5 on blots A-D contain *EcoR* I-digested DNA and lanes 2, 4 and 6 contain *Afl*III digested DNA. The source of the DNA is *E. coli* MC1061 lysogens containing the following phage: lanes 1 & 2, $\Phi24_B$::Kan; 3 & 4, $\Phi24_B$::Cat; 5 & 6, $\Phi24_B$::Kan & $\Phi24_B$::Cat. from Allison *et al.*, 2003

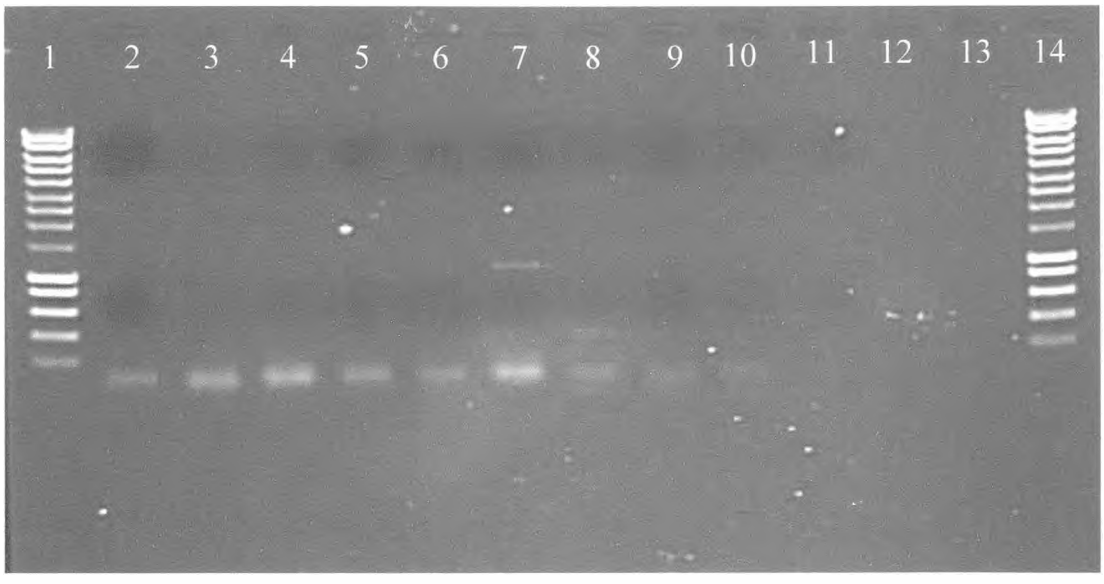


Figure 5.9 Identification of $\Phi 24_B$ integrase gene by PCR

Amplification with integrase primer sets for groups 1-9 (Balding *et al.*, 2005) to dentify the $\Phi 24_B$ integrase gene using bank of integrase specific primers Table 2.3. Lane 1 and 14; Hyper Ladder I, Lane 2; group 1, Lane 3; group 2, Lane 4; group 3, lane 4; group 4, Lane 5; group 4, lane 6; group 5, lane 7; group 6a, lane 8; group 6b, lane 9; group 7, lane 9; group 8, lane 10; group 9, lane 11; GAPDH, lane 12 and 13 blank. Template for PCR was 2 µl of ~10^8 pfu ml^{-1} lysate.

Chapter 6: General Discussion

STEC infection is still a global health concern and incidence remains a problem throughout the world. The improved ability to detect and react to an incidence of STEC infection seems to have led to a limiting effect on the size of an outbreak. For example the Japanese Sakai outbreak recorded 9000 infected patients compared to recent outbreaks in South Wales were the infection was limited to 163 recorded cases. Increased regulations on slaughter, handling and the separation storage of raw and cooked meats aim to limit the opportunity for infection. Early detection and improved treatment through enhanced understanding of the downstream sequelae has led to better treatment of the life threatening symptoms.

Over the last 10 years there has been a re-emergence of the study of phage biology. This has been part due to the identification of phage genes studding the sequenced bacterial chromosomes. Sequencing of the *E. coli* K-12 genome (Blattner *et al.*, 1997) identified bacteriophage remnants offering a basal level of bacteriophage coverage of the bacterial chromosome (5 cryptic prophages). Sequencing of the 2 *E. coli* O157:H7 (Plunkett *et al.*, 1999; Hayashi *et al.*, 1999) strains has identified that this level is probably quite low, as these pathogenic wild type strains show an increased level of remnant bacteriophage genes (~ 18 cryptic prophages per isolate).

We are only beginning to understand the level of influence bacteriophages, whether cryptic or viable, have on the lysogenised host. As discussed above, bacteriophage genes have been shown to aid colonisation of both the bovine and porcine gut in some STEC isolates (Dziva *et al.*, 2004). These are all nuances that are being found amongst different bacteriophages and will constantly change due to the ability of these viruses to either accumulate genes from bacteria by

recombination, or swap a gene or set of genes between phages by the modular switching. Most of the phage genes that encode a positive selection advantage for the lysogen, have thus far offered a macro effect on the lysogen *i.e.* have an increased ability to colonise the gut. There are still a great proportion of phage genes that as yet, have no designated function. This may be that their influence is cumulative or that they increase sensitivity of a certain host response in bacteria that aid both the phage and bacterial survival in any given environment. It has also been shown that Stx may have an effect on the sensitivity of the lysogenic to lytic life cycle switch (Livny *et al.*, 2004), which may aid its propagation by inducing and reinfecting before a competitive phage. This may be a reason why the *stx* genes are conserved with respect to location in the bacteriophage genome. Stx-toxins do increase the pathogenicity of the bacterial host, but must hold some other positive selection otherwise their location would probably vary in the late genes of the phage.

As yet, we do not fully understand the stimulus in the environment and in gut of humans and animals that drives whole communities to the lytic life cycle. There are a number of reasons why this could occur. It is generally thought that changes in environmental conditions may have an imposing effect on stimulus of the lytic life cycle. It has been shown that the addition of phosphate to a *Synechococcus* bloom, leads to bloom collapse by phage-mediated lysis (Wilson *et al.*, 1998). It may be that on induction the amplified production of some phage-encoded proteins may act as a global stimulus for the lytic life cycle of the lysogen population.

The model phage $\Phi 24_B$ used in this study has a distinct advantage over bacteriophage described thus far, as it has been shown to infect through an essential outer membrane protein (Vpr). The findings in this study support the evidence that this is an essential protein as we were unable to make a complete knockout. Vpr has

been designated as integral to the biogenesis of the outer membrane, and in the transfer of outer membrane proteins to the cell surface (Wu *et al.*, 2005). The data in this thesis also concurs with studies that have designated Vpr as regulatable by a host response regulon (σ^E) (Dartigalongue *et al.*, 2001) and that expression can be altered by changing certain growth parameters. In this study, we tried to impose a range of single environmental changes that simulate key environmental factors in the mammalian gut. Anaerobiosis and increasing the growth temperature gave a marked increase in the expression of Vpr. Not only was Vpr expressed in higher amounts when *E. coli* was grown at 42 ^{o} C, it was localised at an increased level demonstrated by raised phage adsorption. As the bovine core body temp can increase above 40 ^{o} C in the summer, if there is a higher presence of Vpr on the gut surface this would also make it easier for any incoming bacteriophage, with the appropriate tail spike to bind and therefore increase lysogen formation or the lytic response.

Gut environmental changes may also play a role in the increase of O157:H7 shedding; cattle tend to shed STEC in higher frequencies in the summer months (Bonardi *et al.*, 1999; McEvoy *et al.* 2003). Change in diet, increasing temperature, increase in certain gut chemicals under certain feeding regimes may all play a role in the variation of the gut environment. If infection rates increase in the summer due to increased levels of bacteriophage, this could mean that free bacteriophages are responsible for the increased shedding. From animal to animal, there must be gut environment differences, some of which result in increased stimulus of the lytic response; thus this diversity could differentiate between high and low STEC shedders. *E. coli* O157:H7 are thought to colonise the rectoanal junction (Naylor *et al.*, 2003). Seasonal STEC shedding levels may be due to high amounts of bacteriophages that infect and lyse key bacteria that are pivotal in adherence of the

STEC bacterial community to its preferred colonisation site thus effectively resulting in an increased level of STEC shedding.

From sequence analysis of the $\Phi 24_B$ tail spike gene, 100 % amino acid sequence identity with the archetypal well characterised Stx2-phage 933W was established. Watarai *et al.* (1998) proposed that bacteriophage 933W infected via the outer membrane protein FadL, although neutralisation studies using FadL were unable to reduce the numbers of short tailed bacteriophage desorbed or unattached in the assay. Deletion of *fadL* did have a marked effect on the infection rate of 933W, although this may have been due to the fluidity of the cell surface being altered. Watarai *et al.* (1998) also showed that with increasing osmolaric pressure, phage infection decreased; this was also identified for $\Phi 24_B$ in the reporter gene assays when Vpr expression was lowered by the addition of bile salts. From this data it would be pertinent to designate Vpr as the bacterial ligand for bacteriophages $\Phi 24_B$ and 933W. Vpr and homologue have been identified throughout the *Enterobacteriaceae* which correlates with the findings of James *et al.* (2001) who showed that $\Phi 24_B$ has a relatively broad host range. Even though phage adsorption assays are a good tool to characterise the tail spike-bacterial OMP association / regulation there is often variation from experiment to experiment. The trend between experiments is constant for analysis of differences between cultures, and it may be due to slight differences in the stage of growth or physiological state of the indicator strain culture.

Producing definite proof that Vpr is the receptor for $\Phi 24_B$ was time consuming and labour-intensive. Consequently studies directed at identification of the short tail spike gene region responsible for binding the phage to the bacterial

receptor, and the degree of conservation/heterogeneity amongst these genes, was curtailed.

Due to the similarity between $\Phi24_B$ and 933W, it was initially pertinent to think that they had similar integrase genes, although PCR analysis proved that this was not true. Casjens (2005) recently reviewed the genomics of tailed phages and how classifying phages as related through certain genes is seriously misguided due to their modular heterogeneity. The integrase of $\Phi24_B$ was found to be similar to phages HK620, P4 and Sf6, although the difficulties with the bacterial chromosomal background of sample preparations has meant that the true identity of the $\Phi24_B$ integrase will only be truly classified once the phage genome is sequenced. This still does not explain why it is possible for $\Phi24_B$ to integrate into the bacteriophage in two separate positions in the bacterial chromosome. Upon induction of these double lysogens, it was possible to identify that the primary infecting phage was always induced in greater numbers than the secondary infecting phage. Paul Fogg (personal communication) is investigating double integration and that it is thought that the secondary integration is mediated by an integrase that is located on the bacterial chromosome and seems to be constitutively expressed.

In conclusion, the data described in this thesis have further characterised the association between $\Phi24_B$ and *E. coli* by unequivocally proving that Vpr is the ligand required in adsorption of the phage. Vpr, as a member of a bacterial host response, is regulatable by stimulating this host response to changes in the bacterial growth environment. This protein, when up-regulated, is also localised at the cell surface at an increased level and thus shows increased binding of the phage. $\Phi24_B$ has been shown to form double lysogens upon infection, which integrate at different positions in the bacterial genome. Bacteriophage $\Phi24_B$ has many similarities to Stx-phage

933W in the genes thus far characterised, including the tail spike gene. It has a distinct integrase, and it may be that this in conjunction with slight differences in the immunity region accounts for the unusual phenomenon of multiple lysogen formation.

Characterisation of the immunity region and the genes that may modulate its function is ongoing. Identification of both integration sites and identification of the phage tail gene epitope that that recognises the binds to the bacterial receptor Vpr and identifying the bacterial epitope required for adsorption are the current research priorities.

Chapter 7 References

Acheson, D. W., Moore, R., De Breucker, S., Lincicome, L., Jacewicz, M., Skutelsky, E. & Keusch, G. T. (1996). Translocation of Shiga toxin across polarized intestinal cells in tissue culture. *Infect Immun* **64**, 3294-3300.

Aertsen, A., Faster, D. & Michiels, C. W. (2005). Induction of Shiga Toxin-Converting Prophage in *Escherichia coli* by High Hydrostatic Pressure. *Appl Environ Microbiol* **71**, 1155-1162.

Alberts, B., Bray, D., Hopkin, K., Johnson, H., Lewis, J., Raff, M., Roberts, K., Walter, P. (2003). *Essential Cell Biology*, 2nd edn: Garland Science.

Allison, H. E., Sergeant, M. J., James, C. E., Saunders, J. R., Smith, D. L., Sharp, R. J., Marks, T. S. & McCarthy, A. J. (2003). Immunity profiles of wild-type and recombinant shiga-like toxin-encoding bacteriophages and characterization of novel double lysogens. *Infect Immun* **71**, 3409-3418.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.

Andrewes, C. H. E., Elford W. J. (1932). The Killing of bacteria by bacteriophage. *Brit J Exp Path* **13**, 13.

Ang, D., Richardson, A., Mayer, M. P., Keppel, F., Krisch, H. & Georgopoulos, C. (2001). Pseudo-T-even bacteriophage RB49 encodes CocO, a cochaperonin for GroEL, which can substitute for *Escherichia coli*'s GroES and Bacteriophage T4's Gp31. *J Biol Chem* **276**, 8720-8726.

Ausabel, F. M., Brent, R., Kingston, R.E., Moore, D.D., Seldmann, J.G. and Struhl, K. (1992). *Current protocols in molecular biology*. Brooklyn: John Wiley and Sons.

Azaro, M. A. a. L., A. (2002). *Lambda integrase and the lambda Int family*: ASM Press Washington D.C.

Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F. C., Bagdasarian, M. M., Frey, J. & Timmis, K. N. (1981). Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**, 237-247.

Bailey, S., Clokie, M. R., Millard, A. & Mann, N. H. (2004). Cyanophage infection and photoinhibition in marine cyanobacteria. *Res Microbiol* **155**, 720-725.

Balding, C., Bromley, S. A., Pickup, R. W. & Saunders, J. R. (2005). Diversity of phage integrases in *Enterobacteriaceae*: development of markers for environmental analysis of temperate phages. *Environ Microbiol* 7, 1558-1567.

Barondess, J. J. & Beckwith, J. (1995). bor gene of phage lambda, involved in serum resistance, encodes a widely conserved outer membrane lipoprotein. *J Bacteriol* **177**, 1247-1253.

Batchelor, M., Prasannan, S., Daniell, S., Reece, S., Connerton, I., Bloomberg, G., Dougan, G., Frankel, G. & Matthews, S. (2000). Structural basis for recognition of the translocated intimin receptor (Tir) by intimin from enteropathogenic *Escherichia coli*. *Embo J* **19**, 2452-2464.

Bell, B. P., Goldoft, M., Griffin, P. M. & other authors (1994). A multistate outbreak of *Escherichia coli* O157:H7-associated bloody diarrhea and hemolytic uremic syndrome from hamburgers. The Washington experience. *Jama* **272**, 1349-1353.

Bettelheim, K. A. (2000). Role of non-O157 VTEC. *Symp Ser Soc Appl Microbiol*, 38S-50S.

Beutin, L. & Muller, W. (1998). Cattle and verotoxigenic *Escherichia coli* (VTEC), an old relationship? *Vet Rec* **142**, 283-284.

Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A. & Ellenberger, T. (2005). A structural basis for allosteric control of DNA recombination by lambda integrase. *Nature* **435**, 1059-1066.

Black, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annu Rev Microbiol* **43**, 267-292.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A. & other authors (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453-1474.

Bonardi, S., Maggi, E., Bottarelli, A., Pacciarini, M. L., Ansuini, A., Vellini, G., Morabito, S. & Caprioli, A. (1999). Isolation of Verocytotoxin-producing *Escherichia coli* O157:H7 from cattle at slaughter in Italy. *Vet Microbiol* **67**, 203-211.

Bonardi, S., Brindani, F., Pizzin, G., Lucidi, L., D'Incau, M., Liebana, E. & Morabito, S. (2003). Detection of *Salmonella* spp., *Yersinia enterocolitica* and verocytotoxin-producing *Escherichia coli* O157 in pigs at slaughter in Italy. *Int J Food Microbiol* **85**, 101-110.

Brunder, W., Schmidt, H. & Karch, H. (1996). KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* **142 (Pt 11)**, 3305-3315.

Brussow, H., Canchaya, C. & Hardt, W. D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**, 560-602, table of contents.

Burland, V., Shao, Y., Perna, N. T., Plunkett, G., Sofia, H. J. & Blattner, F. R. (1998). The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucl Acids Res* **26**, 4196-4204.

Calderwood, S. B., Auclair, F., Donohue-Rolfe, A., Keusch, G. T. & Mekalanos, J. J. (1987). Nucleotide sequence of the Shiga-like toxin genes of *Escherichia coli*. *Proc Natl Acad Sci U S A* **84**, 4364-4368.

Campellone, K. G., Robbins, D. & Leong, J. M. (2004). EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell* **7**, 217-228.

Campellone, K. G., Leong, J. M. (2005). Nck-independent actin assembly is mediated by two phosphorylated tyrosines within enteropathogenic *Escherichia coli* Tir. *Mol Microbiol* **56**, 416-432.

Canchaya, C., Fournous, G. & Brussow, H. (2004). The impact of prophages on bacterial chromosomes. *Mol Microbiol* **53**, 9-18.

Casjens, S. & Hendrix, R. (1974). Comments on the arrangement of the morphogenetic genes of bacteriophage lambda. *J Mol Biol* **90**, 20-25.

Casjens, S. & Adams, M. B. (1985). Posttranscriptional modulation of bacteriophage P22 scaffolding protein gene expression. *J Virol* **53**, 185-191.

Casjens, S., Winn-Stapley, D. A., Gilcrease, E. B., Morona, R., Kuhlewein, C., Chua, J. E., Manning, P. A., Inwood, W. & Clark, A. J. (2004). The chromosome of Shigella flexneri bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. *J Mol Biol* **339**, 379-394.

Casjens, S., and R. Hendrix, R. (1988). *Control mechanisms in dsDNA bacteriophage assembly*: Plenum Publishing Corp., New York, N.Y.

Casjens, S. R. (2005). Comparative genomics and evolution of the tailed-bacteriophages. *Curr Opin Microbiol* **8**, 451-458.

Casjens, S. R., Gilcrease, E. B., Winn-Stapley, D. A. & other authors (2005). The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. *J Bacteriol* **187**, 1091-1104.

Caspar, D. L. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* **27**, 1-24.

Cerritelli, M. E., Cheng, N., Rosenberg, A. H., McPherson, C. E., Booy, F. P. & Steven, A. C. (1997). Encapsidated conformation of bacteriophage T7 DNA. *Cell* **91**, 271-280.

Chapman, P. A., Cerdan Malo, A. T., Ellin, M., Ashton, R. & Harkin (2001). *Escherichia coli* O157 in cattle and sheep at slaughter, on beef and lamb carcasses

and in raw beef and lamb products in South Yorkshire, UK. *Int J Food Microbiol* **64**. 139-150.

Chen, R. & Henning, U. (1996). A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol Microbiol* **19**, 1287-1294.

Chen, Y., Narendra, U., Iype, L. E., Cox, M. M. & Rice, P. A. (2000). Crystal structure of a Flp recombinase-Holliday junction complex: assembly of an active oligomer by helix swapping. *Mol Cell* **6**, 885-897.

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O. & Hanawalt, P. C. (2001). Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* **158**, 41-64.

Courcelle, J. & Hanawalt, P. C. (2003). RecA-dependent recovery of arrested DNA replication forks. *Annu Rev Genet* **37**, 611-646.

Cowden, J. M., Ahmed, S., Donaghy, M. & Riley, A. (2001). Epidemiological investigation of the central Scotland outbreak of *Escherichia coli* O157 infection, November to December 1996. *Epidemiol Infect* **126**, 335-341.

Craig, N. L. & Roberts, J. W. (1980). E. coli recA protein-directed cleavage of phage lambda repressor requires polynucleotide. *Nature* **283**, 26-30.

Dartigalongue, C., Missiakas, D. & Raina, S. (2001). Characterization of the *Escherichia coli* sigma E regulon. *J Biol Chem* **276**, 20866-20875.

Datz, M., Janetzki-Mittmann, C., Franke, S., Gunzer, F., Schmidt, H. & Karch, H. (1996). Analysis of the enterohemorrhagic *Escherichia coli* O157 DNA region containing lambdoid phage gene p and Shiga-like toxin structural genes. *Appl Environ Microbiol* **62**, 791-797.

Dhar, A. & Feiss, M. (2005). Bacteriophage lambda terminase: alterations of the high-affinity ATPase affect viral DNA packaging. *J Mol Biol* **347**, 71-80.

Djordjevic, S. P., Hornitzky, M. A., Bailey, G., Gill, P., Vanselow, B., Walker, K. & Bettelheim, K. A. (2001). Virulence properties and serotypes of Shiga toxin-producing *Escherichia coli* from healthy Australian slaughter-age sheep. *J Clin Microbiol* **39**, 2017-2021.

Dobbins, A. T., George, M., Jr., Basham, D. A., Ford, M. E., Houtz, J. M., Pedulla, M. L., Lawrence, J. G., Hatfull, G. F. & Hendrix, R. W. (2004). Complete genomic sequence of the virulent *Salmonella* bacteriophage SP6. *J Bacteriol* **186**, 1933-1944.

Dobrindt, U. (2005). (Patho-)Genomics of *Escherichia coli*. *International Journal of Medical Microbiology* **295**, 357.

Doerrler, W. T. & Raetz, C. R. (2005). Loss of outer membrane proteins without inhibition of lipid export in an *Escherichia coli* YaeT mutant. *J Biol Chem* **280**, 27679-27687.

Doughty, S., Sloan, J., Bennett-Wood, V., Robertson, M., Robins-Browne, R. M. & Hartland, E. L. (2002). Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic *Escherichia coli*. *Infect Immun* **70**, 6761-6769.

Dundas, S., Todd, W. T., Stewart, A. I., Murdoch, P. S., Chaudhuri, A. K. & Hutchinson, S. J. (2001). The central Scotland *Escherichia coli* O157:H7 outbreak: risk factors for the hemolytic uremic syndrome and death among hospitalized patients. *Clin Infect Dis* **33**, 923-931.

Dupont, K., Janzen, T., Vogensen, F. K., Josephsen, J. & Stuer-Lauridsen, B. (2004). Identification of *Lactococcus lactis* genes required for bacteriophage adsorption. *Appl Environ Microbiol* **70**, 5825-5832.

Dziva, F., van Diemen, P. M., Stevens, M. P., Smith, A. J. & Wallis, T. S. (2004). Identification of *Escherichia coli* O157: H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. *Microbiology* **150**, 3631-3645.

Earnshaw, W. C. & Harrison, S. C. (1977). DNA arrangement in isometric phage heads. *Nature* **268**, 598-602.

Ehrbar, K. & Hardt, W. D. (2005). Bacteriophage-encoded type III effectors in Salmonella enterica subspecies 1 serovar Typhimurium. *Infect Genet Evol* **5**, 1-9.

Evilevitch, A., Lavelle, L., Knobler, C. M., Raspaud, E. & Gelbart, W. M. (2003). Osmotic pressure inhibition of DNA ejection from phage. *Proc Natl Acad Sci U S A* **100**, 9292-9295.

Feiss, M., Widner, W., Miller, G., Johnson, G. & Christiansen, S. (1983). Structure of the bacteriophage lambda cohesive end site: location of the sites of terminase binding (cosB) and nicking (cosN). *Gene* **24**, 207-218.

Feiss, M., Sippy, J. & Miller, G. (1985). Processive action of terminase during sequential packaging of bacteriophage lambda chromosomes. *J Mol Biol* **186**, 759-771.

Foster, J. W. (2004). *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat Rev Microbiol* **2**, 898-907.

Frankel, G., Candy, D. C., Everest, P. & 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.

Gabig, M., Herman-Antosiewicz, A., Kwiatkowska, M., Los, M., Thomas, M. S. & Wegrzyn, G. (2002). The cell surface protein Ag43 facilitates phage infection of *Escherichia coli* in the presence of bile salts and carbohydrates. *Microbiology* **148**, 1533-1542.

Gamage, S. D., Strasser, J. E., Chalk, C. L. & Weiss, A. A. (2003). Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect Immun* **71**, 3107-3115.

Genevrois, S., Steeghs, L., Roholl, P., Letesson, J.-J. & van der Ley, P. (2003). The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *EMBO J* **22**, 1780-1789.

Gerdes, S. Y., Scholle, M. D., Campbell, J. W. & other authors (2003). Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol* **185**, 5673-5684.

Glasgow, A. C. (1989). *Mobile DNA*. Washington DC: American Society for Microbiology.

Hacker, J. & Kaper, J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**, 641-679.

Haggard-Ljungquist, E., Halling, C. & Calendar, R. (1992). DNA sequences of the tail fiber genes of bacteriophage P2: evidence for horizontal transfer of tail fiber genes among unrelated bacteriophages. *J Bacteriol* **174**, 1462-1477.

Hambly, E. & Suttle, C. A. (2005). The viriosphere, diversity, and genetic exchange within phage communities. *Curr Opin Microbiol* **8**, 444-450.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.

Hancock, D., Besser, T., Lejeune, J., Davis, M. & Rice, D. (2001). The control of VTEC in the animal reservoir. *Int J Food Microbiol* **66**, 71-78.

Hartland, E. L., Batchelor, M., Delahay, R. M., Hale, C., Matthews, S., Dougan, G., Knutton, S., Connerton, I. & Frankel, G. (1999). Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Mol Microbiol* **32**, 151-158.

Hashemolhosseini, S., Stierhof, Y. D., Hindennach, I. & Henning, U. (1996). Characterization of the helper proteins for the assembly of tail fibers of coliphages T4 and lambda. *J Bacteriol* **178**, 6258-6265.

Hayashi, T., Makino, K., Ohnishi, M. & other authors (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**, 11-22.

Helgstrand, C., Wikoff, W. R., Duda, R. L., Hendrix, R. W., Johnson, J. E. & Liljas, L. (2003). The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 A resolution. *J Mol Biol* **334**, 885-899.

Hendrix, R. W. & Casjens, S. R. (1974). Protein cleavage in bacteriophage lambda tail assembly. *Virology* **61**, 156-159.

Hendrix, R. W. (2005a). Bacteriophage HK97: assembly of the capsid and evolutionary connections. *Adv Virus Res* **64**, 1-14.

Hendrix, R. W. C., S. (2005b). *Bacteriophage lambda and its genetic neighbourhood.*, 2 edn: Oxford Press.

Herold, S., Siebert, J., Huber, A. & Schmidt, H. (2005). Global expression of prophage genes in *Escherichia coli* O157:H7 strain EDL933 in response to norfloxacin. *Antimicrob Agents Chemother* **49**, 931-944.

Herskowitz, I. & Hagen, D. (1980). The lysis-lysogeny decision of phage lambda: explicit programming and responsiveness. *Annu Rev Genet* **14**, 399-445.

Hicks, S., Frankel, G., Kaper, J. B., Dougan, G. & 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.

Hirai, K., Aoyama, H., Suzue, S., Irikura, T., Iyobe, S. & Mitsuhashi, S. (1986). Isolation and characterization of norfloxacin-resistant mutants of *Escherichia coli* K-12. *Antimicrob Agents Chemother* **30**, 248-253.

Hohn, B. (1983). DNA sequences necessary for packaging of bacteriophage lambda DNA. *Proc Natl Acad Sci U S A* **80**, 7456-7460.

Hsiao, C. L. & Black, L. W. (1977). DNA packaging and the pathway of bacteriophage T4 head assembly. *Proc Natl Acad Sci U S A* **74**, 3652-3656.

Jackson, E. N., Jackson, D. A. & Deans, R. J. (1978). EcoRI analysis of bacteriophage P22 DNA packaging. *J Mol Biol* **118**, 365-388.

James, C. E., Stanley, K. N., Allison, H. E., Flint, H. J., Stewart, C. S., Sharp, R. J., Saunders, J. R. & McCarthy, A. J. (2001). Lytic and lysogenic infection of diverse *Escherichia coli* and *Shigella* strains with a verocytotoxigenic bacteriophage. *Appl Environ Microbiol* **67**, 4335-4337.

James, C. E. (2002). Use of a recombinant verocytotoxigenic bacteriophage to investigate infection of bacterial hosts.: Ph.D Thesis, University of Liverpool.

Janisiewicz, W. J., Conway, W. S., Brown, M. W., Sapers, G. M., Fratamico, P. & Buchanan, R. L. (1999). Fate of *Escherichia coli* O157:H7 on fresh-cut apple tissue and its potential for transmission by fruit flies. *Appl Environ Microbiol* **65**, 1-5.

Jarvis, K. G., Giron, J. A., Jerse, A. E., McDaniel, T. K., Donnenberg, M. S. & 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 U S A* **92**, 7996-8000.

Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige, P. E., Jr. & Chiu, W. (2003). Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**, 131-135.

Kaper, J. B. (1998). Enterohemorrhagic *Escherichia coli*. *Curr Opin Microbiol* **1**, 103-108.

Kaper, J. B., Nataro, J. P. & Mobley, H. L. (2004). Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**, 123-140.

Karaolis, D. K., Somara, S., Maneval, D. R., Jr., Johnson, J. A. & Kaper, J. B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* **399**, 375-379.

Karch, H. H., J. Laufs, R. O'Brien, A. D. Tacket, C.O. Levine, M.M. (1987). A plasmid of enterohemorrhagic *Escherichia coli* O157:H7 is required for expression of a new fimbrial antigen and for adhesion to epithelial cells. *Infect Immun* **55**, 455-461.

Keene, W. E., Sazie, E., Kok, J., Rice, D. H., Hancock, D. D., Balan, V. K., Zhao, T. & Doyle, M. P. (1997). An outbreak of *Escherichia coli* O157:H7 infections traced to jerky made from deer meat. *Jama* **277**, 1229-1231.

Kenny, B. & Finlay, B. B. (1995). Protein secretion by enteropathogenic *Escherichia coli* is essential for transducing signals to epithelial cells. *Proc Natl Acad Sci U S A* **92**, 7991-7995.

Kenny, B., Lai, L. C., Finlay, B. B. & 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.

Keppel, F., Rychner, M. & Georgopoulos, C. (2002). Bacteriophage-encoded cochaperonins can substitute for *Escherichia coli*'s essential GroES protein. *EMBO Rep* **3**, 893-898.

Kholodii, G. Y. & Mindlin, S. Z. (1985). Integration of bacteriophages lambda and phi 80 in wild-type *Escherichia coli* at secondary attachment sites. II. Genetic structure and mechanism of polylysogen formation for lambda, phi 80 and the lambda att80 hybrid. *Mol Gen Genet* **198**, 491-496.

Kiino, D. R. & Rothman-Denes, L. B. (1989). Genetic analysis of bacteriophage N4 adsorption. *J Bacteriol* **171**, 4595-4602.

Kimura, T., Tani, S., Motoki, M. & Matsumoto, Y. (2003). Role of Shiga toxin 2 (Stx2)-binding protein, human serum amyloid P component (HuSAP), in Shiga toxin-producing *Escherichia coli* infections: assumption from in vitro and in vivo study using HuSAP and anti-Stx2 humanized monoclonal antibody TMA-15. *Biochem Biophys Res Commun* **305**, 1057-1060.

King, J. & Casjens, S. (1974). Catalytic head assembling protein in virus morphogenesis. *Nature* **251**, 112-119.

Koudelka, A. P., Hufnagel, L. A. & Koudelka, G. B. (2004). Purification and characterization of the repressor of the shiga toxin-encoding bacteriophage 933W: DNA binding, gene regulation, and autocleavage. *J Bacteriol* **186**, 7659-7669.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Landy, A. (1989). Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem* **58**, 913-949.

Law, D. (2000). Virulence factors of *Escherichia coli* O157 and other Shiga toxin-producing E. coli. *J Appl Microbiol* **88**, 729-745.

Lee, C. A. (1999). *Vibrio cholerae* TCP: a trifunctional virulence factor? *Trends Microbiol* **7**, 391-392; discussion 393.

Lee, K. K., Gan, L., Tsuruta, H., Hendrix, R. W., Duda, R. L. & Johnson, J. E. (2004). Evidence that a local refolding event triggers maturation of HK97 bacteriophage capsid. *J Mol Biol* **340**, 419-433.

Leung, P. H., Peiris, J. S., Ng, W. W., Robins-Browne, R. M., Bettelheim, K. A. & Yam, W. C. (2003). A newly discovered verotoxin variant, VT2g, produced by bovine verocytotoxigenic *Escherichia coli*. *Appl Environ Microbiol* **69**, 7549-7553.

Lin, J., Smith, M. P., Chapin, K. C., Baik, H. S., Bennett, G. N. & Foster, J. W. (1996). Mechanisms of acid resistance in enterohemorrhagic *Escherichia coli*. *Appl Environ Microbiol* **62**, 3094-3100.

Livny, J. & Friedman, D. I. (2004). Characterizing spontaneous induction of Stx encoding phages using a selectable reporter system. *Mol Microbiol* **51**, 1691-1704.

Louise, C. B. & Obrig, T. G. (1995). Specific interaction of *Escherichia coli* O157:H7-derived Shiga-like toxin II with human renal endothelial cells. *J Infect Dis* **172**, 1397-1401.

Ludwig, W., Strunk, O., Westram, R. & other authors (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**, 1363-1371.

Luo, Y., Frey, E. A., Pfuetzner, R. A., Creagh, A. L., Knoechel, D. G., Haynes, C. A., Finlay, B. B. & Strynadka, N. C. (2000). Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* **405**, 1073-1077.

Lwoff, A., Siminovitch, L. & Kjeldgaard, N. (1950). Induction of the production of bacteriophages in lysogenic bacteria.. *Ann Inst Pasteur (Paris)* **79**, 815-859.

Mann, N. H., Cook, A., Millard, A., Bailey, S. & Clokie, M. (2003). Marine ecosystems: bacterial photosynthesis genes in a virus. *Nature* **424**, 741.

Marches, O., Ledger, T. N., Boury, M. & other authors (2003). Enteropathogenic and enterohaemorrhagic *Escherichia coli* deliver a novel effector called Cif, which blocks cell cycle G2/M transition. *Mol Microbiol* **50**, 1553-1567.

Matsushiro, A., Sato, K., Miyamoto, H., Yamamura, T. & Honda, T. (1999). Induction of prophages of enterohemorrhagic *Escherichia coli* O157:H7 with norfloxacin. *J Bacteriol* **181**, 2257-2260.

McDaniel, T. K. & 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.

McEvoy, J. M., Doherty, A. M., Sheridan, J. J., Thomson-Carter, F. M., Garvey, P., McGuire, L., Blair, I. S. & McDowell, D. A. (2003). The prevalence and spread of *Escherichia coli* O157:H7 at a commercial beef abattoir. *J Appl Microbiol* **95**, 256-266.

McLeod, S. M. & Waldor, M. K. (2004). Characterization of XerC- and XerD-dependent CTX phage integration in *Vibrio cholerae*. *Mol Microbiol* **54**, 935-947.

Melton-Celsa, A. R., Rogers, J. E., Schmitt, C. K., Darnell, S. C. & O'Brien, A. D. (1998). Virulence of Shiga toxin-producing *Escherichia coli* (STEC) in orally-infected mice correlates with the type of toxin produced by the infecting strain. *Jpn J Med Sci Biol* **51 Suppl**, S108-114.

Menge, C., Stamm, I., Wuhrer, M., Geyer, R., Wieler, L. H. & Baljer, G. (2001). Globotriaosylceramide (Gb3/CD77) is synthesized and surface expressed by bovine lymphocytes upon activation in vitro. *Veterinary Immunology and Immunopathology* **83**, 19.

Menge, C., Stamm, I., Blessenohl, M., Wieler, L. H. & Baljer, G. (2003). Verotoxin 1 from *Escherichia coli* affects Gb3/CD77+ bovine lymphocytes independent of interleukin-2, tumor necrosis factor-alpha, and interferon-alpha. *Exp Biol Med (Maywood)* **228**, 377-386.

Menge, C., Blessenohl, M., Eisenberg, T., Stamm, I. & Baljer, G. (2004). Bovine ileal intraepithelial lymphocytes represent target cells for Shiga toxin 1 from *Escherichia coli*. *Infect Immun* **72**, 1896-1905.

Michalowski, C. B. & Little, J. W. (2005). Positive autoregulation of cI is a dispensable feature of the phage lambda gene regulatory circuitry. *J Bacteriol* **187**, 6430-6442.

Miller, J. H. (1972). *Experiments in molecular genetics*: Cols Spring Harbour Laboratory Press.

Milne, L. M., Plom, A., Strudley, I. & other authors (1999). *Escherichia coli* O157 incident associated with a farm open to members of the public. *Commun Dis Public Health* **2**, 22-26.

Missiakas, D. & Raina, S. (1998). The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* **28**, 1059-1066.

Miwa, T. & Matsubara, K. (1983). Lambda phage DNA sequences affecting the packaging process. *Gene* **24**, 199-206.

Morabito, S., Tozzoli, R., Oswald, E. & Caprioli, A. (2003). A mosaic pathogenicity island made up of the locus of enterocyte effacement and a pathogenicity island of *Escherichia coli* O157:H7 is frequently present in attaching and effacing E. coli. *Infect Immun* **71**, 3343-3348.

Nakagawa, S., Kojio, S., Taneike, I., Iwakura, N., Tamura, Y., Kushiya, K., Gondaira, F. & Yamamoto, T. (2003). Inhibitory action of telithromycin against Shiga toxin and endotoxin. *Biochem Biophys Res Commun* **310**, 1194-1199.

Nataro, J. P. & Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**, 142-201.

Naylor, S. W., Low, J. C., Besser, T. E., Mahajan, A., Gunn, G. J., Pearce, M. C., McKendrick, I. J., Smith, D. G. & Gally, D. L. (2003). Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. *Infect Immun* **71**, 1505-1512.

Neely, M. N. & Friedman, D. I. (1998). Arrangement and functional identification of genes in the regulatory region of lambdoid phage H-19B, a carrier of a Shiga-like toxin. *Gene* **223**, 105-113.

O'Brien, A. D. & LaVeck, G. D. (1983). Purification and characterization of a Shigella dysenteriae 1-like toxin produced by *Escherichia coli*. *Infect Immun* **40**, 675-683.

O'Loughlin, E. V. & Robins-Browne, R. M. (2001). Effect of Shiga toxin and Shiga-like toxins on eukaryotic cells. *Microbes Infect* **3**, 493-507.

Obrig, T. G., Moran, T. P. & Brown, J. E. (1987). The mode of action of Shiga toxin on peptide elongation of eukaryotic protein synthesis. *Biochem J* **244**, 287-294.

Ogden, I. D., MacRae, M. & Strachan, N. J. (2004). Is the prevalence and shedding concentrations of E. coli O157 in beef cattle in Scotland seasonal? *FEMS Microbiol Lett* **233**, 297-300.

Ogierman, M. A., Paton, A. W. & Paton, J. C. (2000). Up-Regulation of Both Intimin and eae-Independent Adherence of Shiga Toxigenic *Escherichia coli* O157 by ler and Phenotypic Impact of a Naturally Occurring ler Mutation. *Infect Immun* **68**, 5344-5353.

Olivares, E. C., Hollis, R. P. & Calos, M. P. (2001). Phage R4 integrase mediates site-specific integration in human cells. *Gene* **278**, 167-176.

Olson, N. H., Gingery, M., Eiserling, F. A. & Baker, T. S. (2001). The structure of isometric capsids of bacteriophage T4. *Virology* **279**, 385-391.

Onufryk, C., Crouch, M. L., Fang, F. C. & Gross, C. A. (2005). Characterization of six lipoproteins in the sigmaE regulon. *J Bacteriol* **187**, 4552-4561.

Penfold, R. J. & Pemberton, J. M. (1992). An improved suicide vector for construction of chromosomal insertion mutations in bacteria. *Gene* **118**, 145-146.

Perna, N. T., Plunkett, G., 3rd, Burland, V. & other authors (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529-533.

Peterson, J. W., Houston, C. W. & Koo, F. C. (1981). Influence of cultural conditions on mitomycin C-mediated bacteriophage induction and release of *Salmonella* toxin. *Infect Immun* **32**, 232-242.

Piccione, G., Caola, G. & Refinetti, R. (2003). Daily and estrous rhythmicity of body temperature in domestic cattle. *BMC Physiol* **3**, 7.

Plunkett, G., 3rd, Rose, D. J., Durfee, T. J. & Blattner, F. R. (1999). Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J Bacteriol* **181**, 1767-1778.

Pradel, N., Santini, C. L., Ye, C. Y., Fevat, L., Gerard, F., Alami, M. & Wu, L. F. (2003). Influence of tat mutations on the ribose-binding protein translocation in *Escherichia coli*. *Biochem Biophys Res Commun* **306**, 786-791.

Pritchard, G. C., Willshaw, G. A., Bailey, J. R., Carson, T. & Cheasty, T. (2000). Verocytotoxin-producing *Escherichia coli* O157 on a farm open to the public: outbreak investigation and longitudinal bacteriological study. *Vet Rec* **147**, 259-264.

Ptashne, M. (1992). *A genetic switch: phage lambda and higher organisms*, 2nd edn. Oxford: Blackwell Scientific Publications.

Ptashne, M. (2004). The Genetic Swithc: Phage Lambda Revisited. Cold Spring Harbour Press, Cold Spring Harbour (NY).

Purohit, P. K., Inamdar, M. M., Grayson, P. D., Squires, T. M., Kondev, J. & Phillips, R. (2005). Forces during bacteriophage DNA packaging and ejection. *Biophys J* **88**, 851-866.

Recktenwald, J. & Schmidt, H. (2002). The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage phiP27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. *Infect Immun* **70**, 1896-1908.

Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**, 1295-1306.

Richards, K. E., Williams, R. C. & Calendar, R. (1973). Mode of DNA packing within bacteriophage heads. *J Mol Biol* **78**, 255-259.

Richardson, S. E., Rotman, T. A., Jay, V., Smith, C. R., Becker, L. E., Petric, M., Olivieri, N. F. & Karmali, M. A. (1992). Experimental verocytotoxemia in rabbits. *Infect Immun* **60**, 4154-4167.

Riley, L. W., Remis, R. S., Helgerson, S. D. & other authors (1983). Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N Engl J Med* **308**, 681-685.

Robinson, S. E., Wright, E. J., Bennett, M., French, N. P. & Hart, C. A. (2004). Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. *Journal of Applied Microbiology* **97**, 1045.

Roessner, C. A. & Ihler, G. M. (1984). Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH in complexes with the lambda receptor. *J Bacteriol* **157**, 165-170.

Rohatgi, R., Ma, L., Miki, H., Lopez, M., Kirchhausen, T., Takenawa, T. & Kirschner, M. W. (1999). The Interaction between N-WASP and the Arp2/3 Complex Links Cdc42-Dependent Signals to Actin Assembly. *Cell* **97**, 221.

Ruiz, N., Falcone, B., Kahne, D. & Silhavy, T. J. (2005). Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* **121**, 307-317.

Sambrook, J. E., Fritsch, E.F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*, 2 edn. Plainview, NY: Cold Spring Harbor Laboratory Press.

Sampson, L. L., Hendrix, R. W., Huang, W. M. & Casjens, S. R. (1988). Translation initiation controls the relative rates of expression of the bacteriophage lambda late genes. *Proc Natl Acad Sci U S A* **85**, 5439-5443.

Sandmeier, H., Iida, S. & Arber, W. (1992). DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. *J Bacteriol* **174**, 3936-3944.

Sandvig, K. & van Deurs, B. (1996). Endocytosis, intracellular transport, and cytotoxic action of Shiga toxin and ricin. *Physiol Rev* **76**, 949-966.

Sasaki, T., Kobayashi, M. & Agui, N. (2000). Epidemiological potential of excretion and regurgitation by Musca domestica (Diptera: Muscidae) in the dissemination of *Escherichia coli* O157: H7 to food. *J Med Entomol* **37**, 945-949.

Schlesinger, M. (1934). Zur Frage der chemischen Zusammensetzung des Bakteriophagen. *Biochem Zschr*, 273:306.

Schmidt, H., Bielaszewska, M. & Karch, H. (1999). Transduction of Enteric *Escherichia coli* Isolates with a Derivative of Shiga Toxin 2-Encoding Bacteriophage phi 3538 Isolated from *Escherichia coli* O157:H7. *Appl Environ Microbiol* **65**, 3855-3861.

Scholl, D., Adhya, S. & Merril, C. R. (2002). Bacteriophage SP6 is closely related to phages K1-5, K5, and K1E but encodes a tail protein very similar to that of the distantly related P22. *J Bacteriol* **184**, 2833-2836.

Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M. & Scherzinger, E. (1989). Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene* **75**, 271-288.

Schwartz, M. (1976). The adsorption of coliphage lambda to its host: effect of variations in the surface density of receptor and in phage-receptor affinity. *J Mol Biol* **103**, 521-536.

Sergeant, M. J. (1998). Molecular biological characterisation of verotoxigenic bacteriophages in *E. coli*: Ph.D Thesis, University of Liverpool.

Shotland, Y., Koby, S., Teff, D. & other authors (1997). Proteolysis of the phage lambda CII regulatory protein by FtsH (HflB) of *Escherichia coli*. *Mol Microbiol* **24**, 1303-1310.

Shotland, Y., Shifrin, A., Ziv, T., Teff, D., Koby, S., Kobiler, O. & Oppenheim, A. B. (2000). Proteolysis of bacteriophage lambda CII by *Escherichia coli* FtsH (HflB). *J Bacteriol* **182**, 3111-3116.

Smith, H. W., Green, P. & 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.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* **98**, 503-517.

Spakowitz, A. J. & Wang, Z.-G. (2005). DNA Packaging in Bacteriophage: Is Twist Important? *Biophys J* **88**, 3912-3923.

Staats, J. J., Chengappa, M. M., DeBey, M. C., Fickbohm, B. & Oberst, R. D. (2003). Detection of *Escherichia coli* Shiga toxin (stx) and enterotoxin (estA and elt) genes in fecal samples from non-diarrheic and diarrheic greyhounds. *Vet Microbiol* **94**, 303-312.

Sternberg, N. & Weisberg, R. (1977). Packaging of coliphage lambda DNA. II. The role of the gene D protein. *J Mol Biol* **117**, 733-759.

Susskind, M. M. & Botstein, D. (1975). Mechanism of action of Salmonella phage P22 antirepressor. *J Mol Biol* **98**, 413-424.

Tesh, V. L. & O'Brien, A. D. (1991). The pathogenic mechanisms of Shiga toxin and the Shiga-like toxins. *Mol Microbiol* **5**, 1817-1822.

Thomas, A., Smith, H. R. & Rowe, B. (1993). Use of digoxigenin-labelled oligonucleotide DNA probes for VT2 and VT2 human variant genes to differentiate Vero cytotoxin-producing *Escherichia coli* strains of serogroup O157. *J Clin Microbiol* **31**, 1700-1703.

Thomas, A., Cheasty, T., Frost, J. A., Chart, H., Smith, H. R. & Rowe, B. (1996). Vero cytotoxin-producing *Escherichia coli*, particularly serogroup O157, associated with human infections in England and Wales: 1992-4. *Epidemiol Infect* **117**, 1-10.

Thyagarajan, B., Guimaraes, M. J., Groth, A. C. & Calos, M. P. (2000). Mammalian genomes contain active recombinase recognition sites. *Gene* **244**, 47-54.

Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S. & Calos, M. P. (2001). Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol* **21**, 3926-3934.

Tielker, D., Hacker, S., Loris, R., Strathmann, M., Wingender, J., Wilhelm, S., Rosenau, F. & Jaeger, K.-E. (2005). *Pseudomonas aeruginosa* lectin LecB is

located in the outer membrane and is involved in biofilm formation. *Microbiology* **151**, 1313-1323.

Twort, F. W. (1915). An investigation on the nature of the ultra-microscopic viruses. *Lancet*, 1241-1243.

Tyler, J. S., Mills, M. J. & Friedman, D. I. (2004). The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression. *J Bacteriol* **186**, 7670-7679.

Unkmeir, A. & Schmidt, H. (2000). Structural Analysis of Phage-Borne stx Genes and Their Flanking Sequences in Shiga Toxin-Producing *Escherichia coli* and *Shigella dysenteriae* Type 1 Strains. *Infect Immun* **68**, 4856-4864.

Van Duyne, G. D. (2001). A structural view of cre-loxp site-specific recombination. *Annu Rev Biophys Biomol Struct* **30**, 87-104.

Veltkamp, C. J., Chubb, J. C., Birch, S. P., and Eaton, J. W. (1994). A simple freeze dehydration method for studying epiphytic and epizoic communities using the scanning electron microscope. *Hydrobiologica* **288**, 33-38.

Virrankoski-Castrodeza, V., Fraser, M. J. & Parish, J. H. (1982). Condensed DNA structures derived from bacteriophage heads. *J Gen Virol* **58 Pt 1**, 181-190.

Wagner, P. L., Acheson, D. W. K. & Waldor, M. K. (1999). Isogenic Lysogens of Diverse Shiga Toxin 2-Encoding Bacteriophages Produce Markedly Different Amounts of Shiga Toxin. *Infect Immun* **67**, 6710-6714.

Waldor, M. K. & Friedman, D. I. (2005). Phage regulatory circuits and virulence gene expression. *Curr Opin Microbiol* **8**, 459-465.

Wang, G. & Doyle, M. P. (1998). Survival of enterohemorrhagic *Escherichia coli* O157:H7 in water. *J Food Prot* **61**, 662-667.

Wang, J., Michel, V., Hofnung, M. & Charbit, A. (1998). Cloning of the J gene of bacteriophage lambda, expression and solubilization of the J protein: first in vitro studies on the interactions between J and LamB, its cell surface receptor. *Research in Microbiology* **149**, 611.

Wang, J., Hofnung, M. & Charbit, A. (2000). The C-terminal portion of the tail fiber protein of bacteriophage lambda is responsible for binding to LamB, its receptor at the surface of *Escherichia coli* K-12. *J Bacteriol* **182**, 508-512.

Watanabe, Y., Ozasa, K., Mermin, J. H., Griffin, P. M., Masuda, K., Imashuku, S. & Sawada, T. (1999). Factory outbreak of *Escherichia coli* O157:H7 infection in Japan. *Emerg Infect Dis* **5**, 424-428.

Watarai, M., Sato, T., Kobayashi, M., Shimizu, T., Yamasaki, S., Tobe, T., Sasakawa, C. & Takeda, Y. (1998). Identification and characterization of a newly

isolated shiga toxin 2-converting phage from shiga toxin-producing *Escherichia coli*. *Infect Immun* **66**, 4100-4107.

Welch, M. D. (1999). The world according to Arp: regulation of actin nucleation by the Arp2/3 complex. *Trends in Cell Biology* **9**, 423.

Westerholt, S., Hartung, T., Tollens, M. & other authors (2000). Inflammatory and immunological parameters in children with haemolytic uremic syndrome (HUS) and gastroenteritis-pathophysiological and diagnostic clues. *Cytokine* **12**, 822-827.

Willshaw, G. A., Smith, H. R., Scotland, S. M., Field, A. M. & Rowe, B. (1987). Heterogeneity of *Escherichia coli* phages encoding Vero cytotoxins: comparison of cloned sequences determining VT1 and VT2 and development of specific gene probes. *J Gen Microbiol* **133**, 1309-1317.

Willshaw, G. A., Cheasty, T., Smith, H. R., O'Brien, S. J. & Adak, G. K. (2001). Verocytotoxin-producing *Escherichia coli* (VTEC) O157 and other VTEC from human infections in England and Wales: 1995-1998. *J Med Microbiol* **50**, 135-142.

Wilson, W. H. T., Turner, S. and Mann, N. H. (1998). Population dynamics of phytoplankton and viruses in a phosphate-limited mesocosm and their effect on DMSP and DMS production. *Estuarine, Coastal and Shelf Science* **46**, 49 - 59.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235-245.

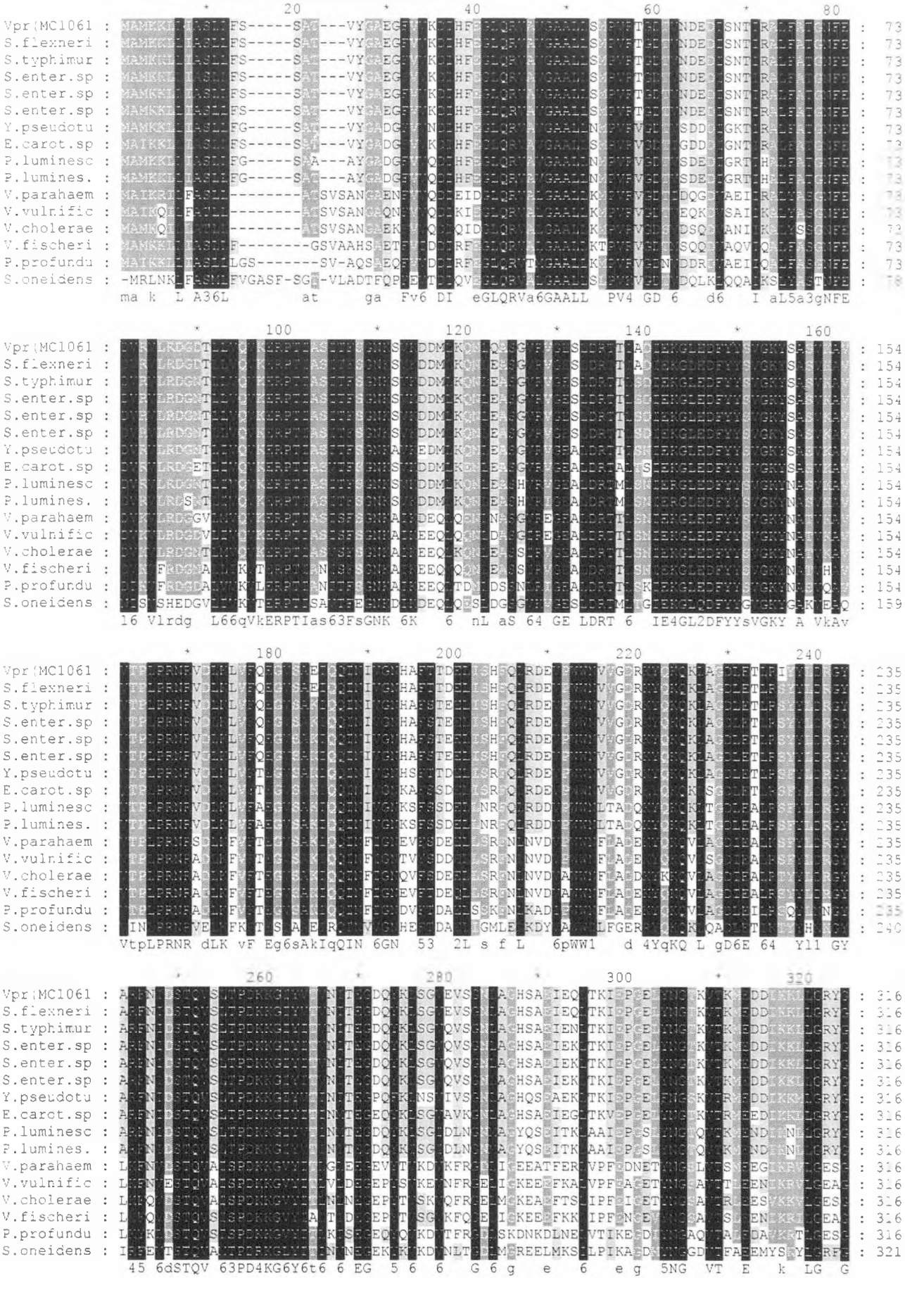
Xu, J., Hendrix, R. W. & Duda, R. L. (2004). Conserved translational frameshift in dsDNA bacteriophage tail assembly genes. *Mol Cell* **16**, 11-21.

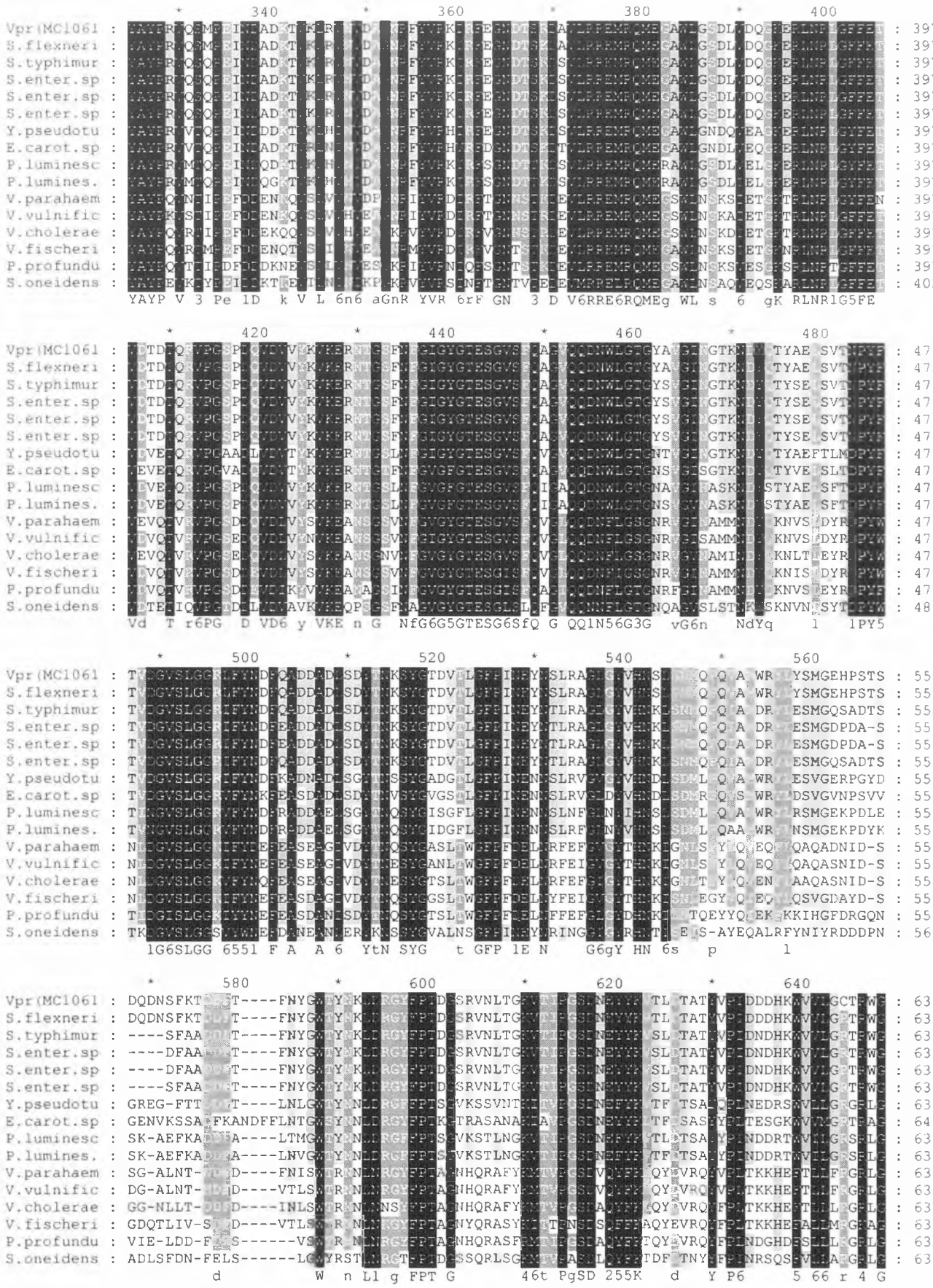
Yarnell, W. S. & Roberts, J. W. (1992). The phage lambda gene Q transcription antiterminator binds DNA in the late gene promoter as it modifies RNA polymerase. *Cell* **69**, 1181-1189.

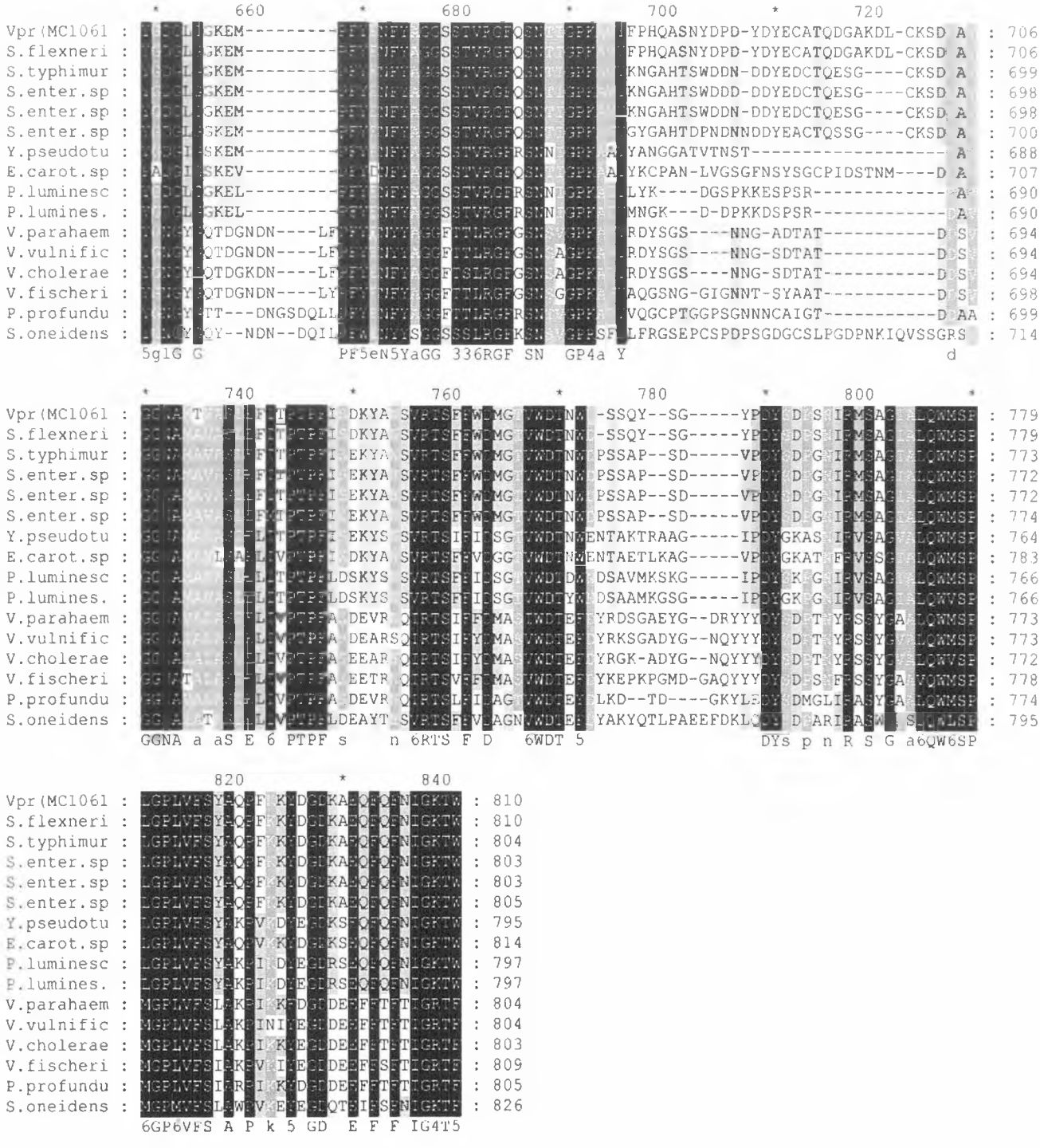
Young, I., Wang, I. & Roof, W. D. (2000). Phages will out: strategies of host cell lysis. *Trends Microbiol* **8**, 120-128.

Zhang, Z., Greene, B., Thuman-Commike, P. A., Jakana, J., Prevelige, J. P. E., King, J. & Chiu, W. (2000). Visualization of the maturation transition in bacteriophage P22 by electron cryomicroscopy. *Journal of Molecular Biology* **297**, 615.

Appendix 1: Protein Alignment of Vpr against highest ranking blastP proteins of >50% sequence identity.

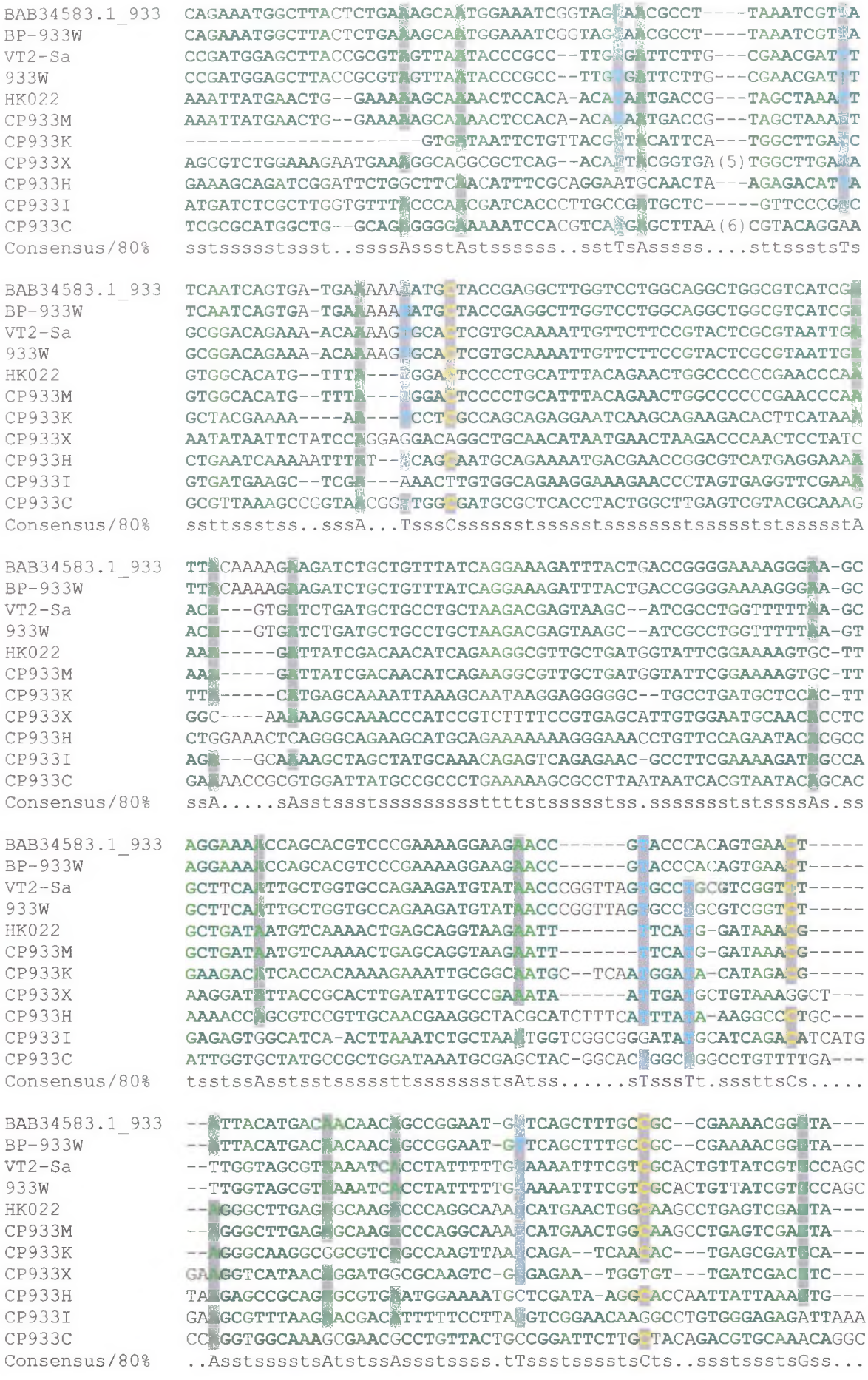


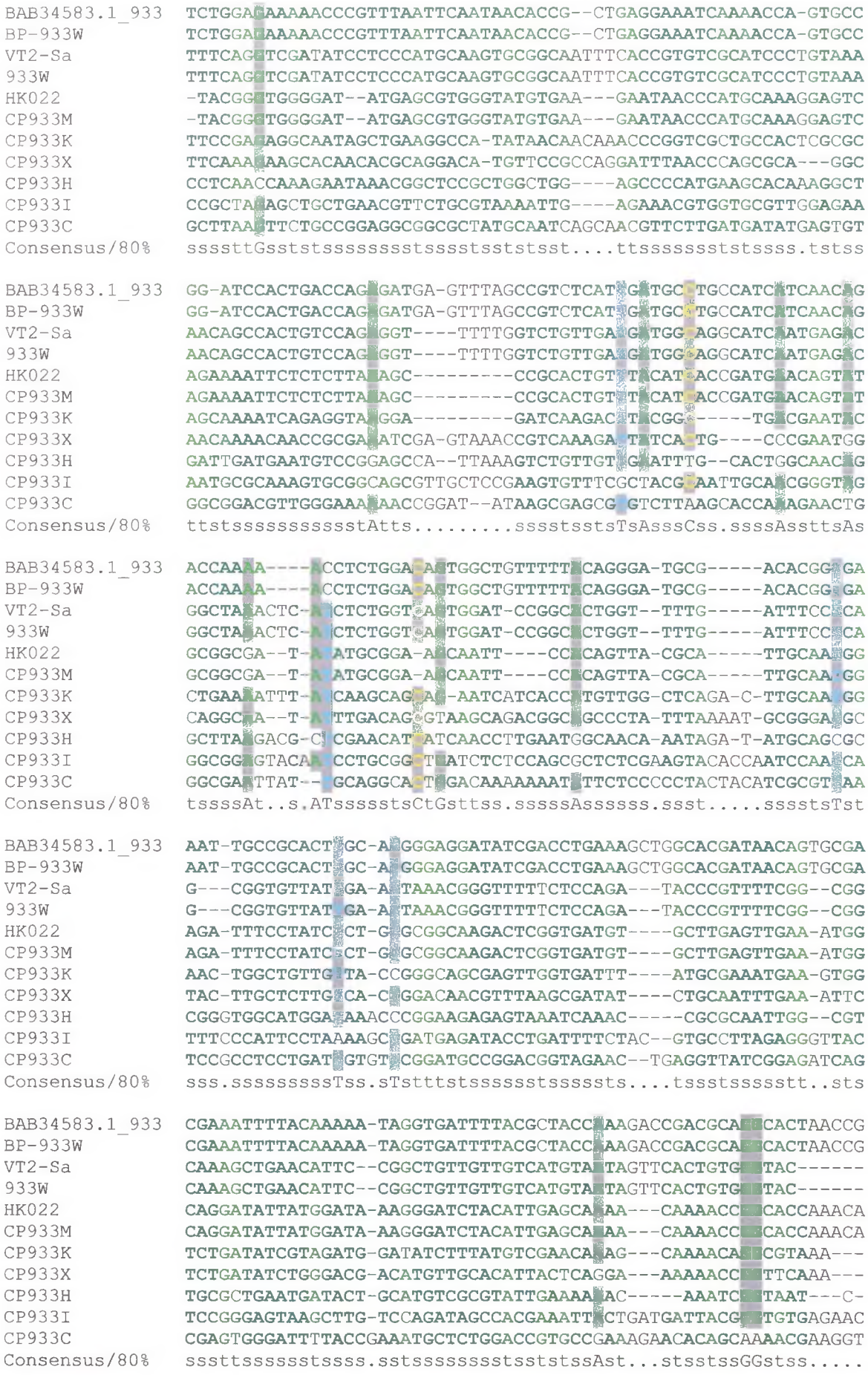


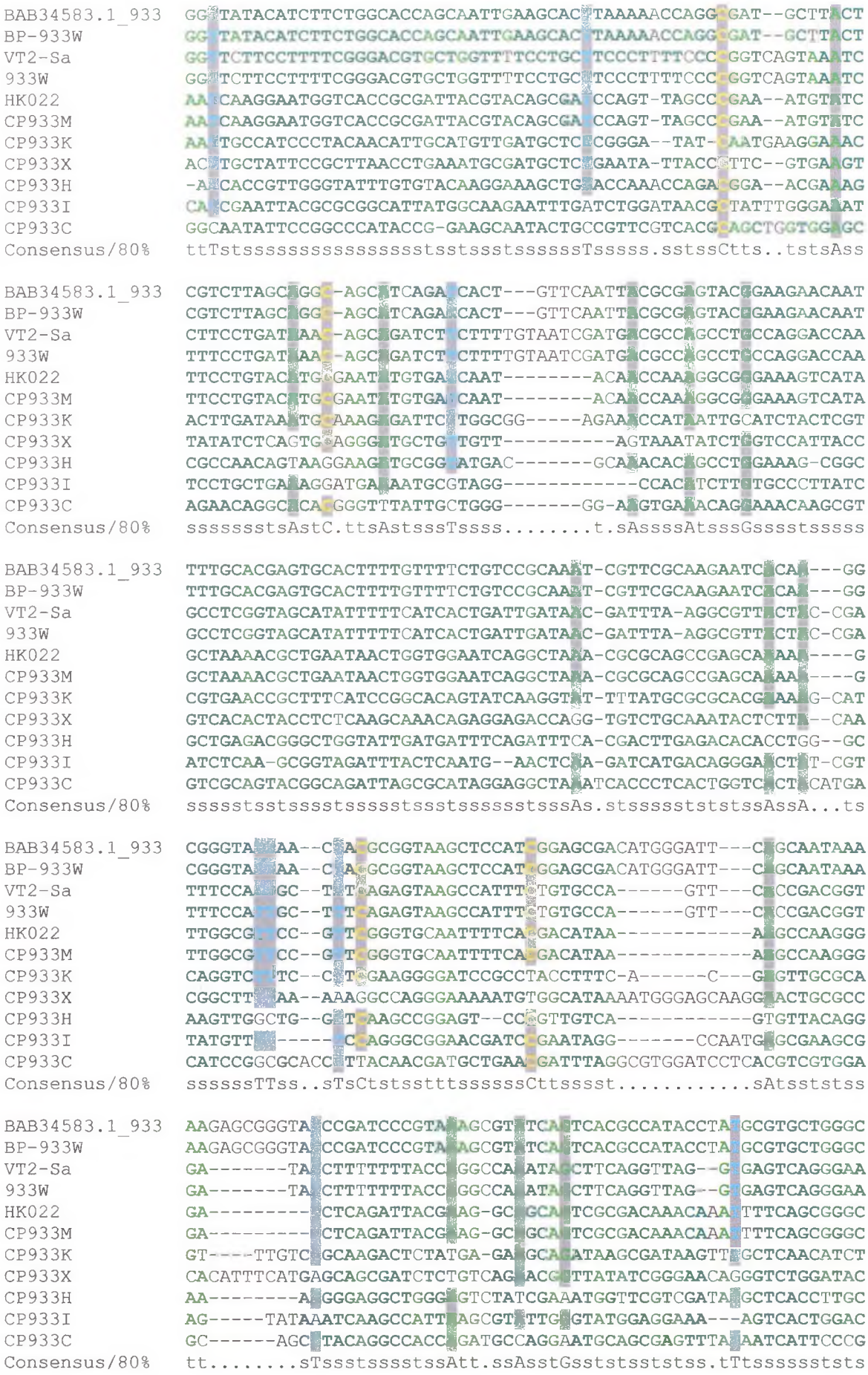


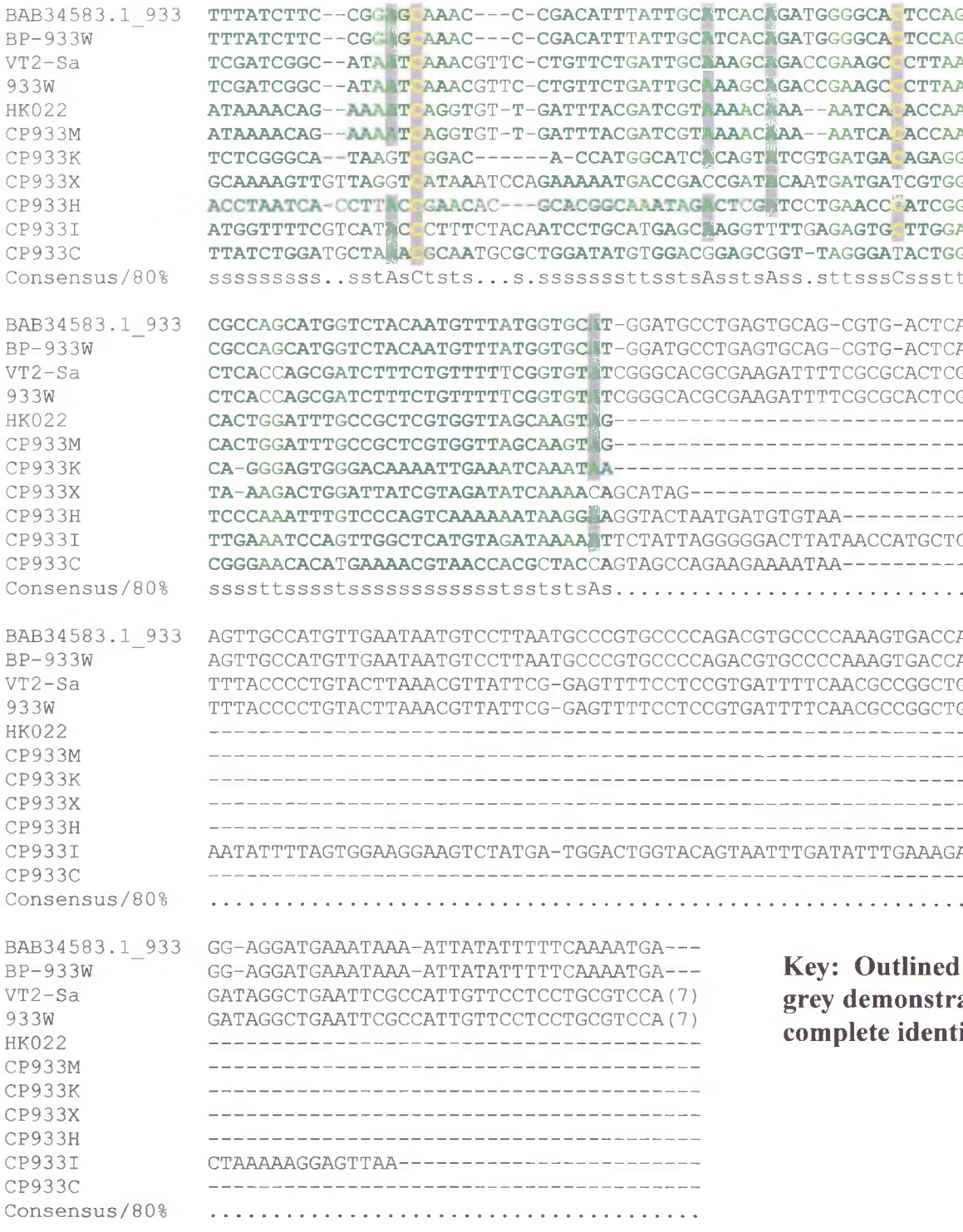
Key: Black outline demonstrates complete identity; light grey demonstrates partial identity.

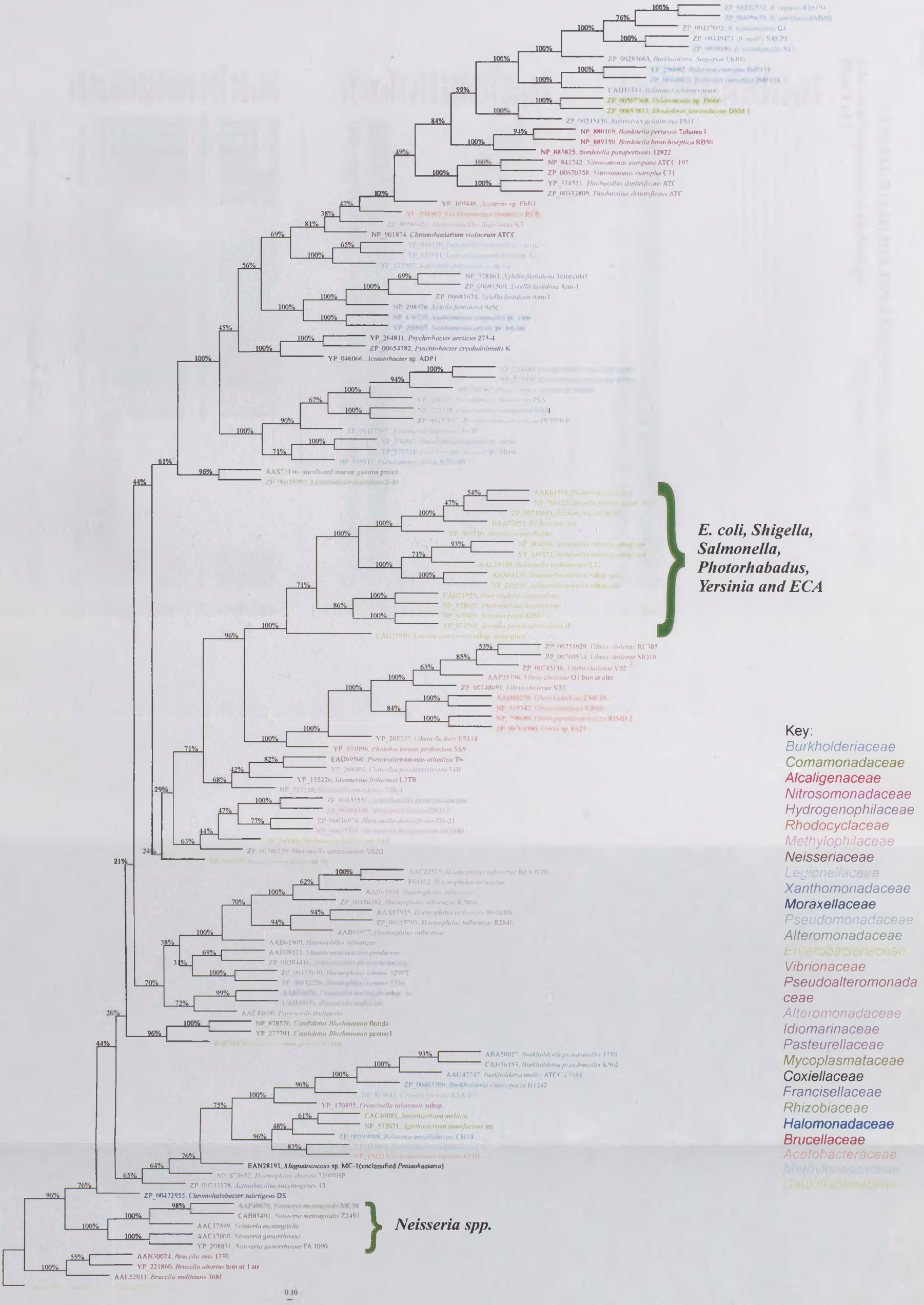
Appendix 2: ClustalX Alignment of Stx_2-phages integrase Genes


BAB34583.1_933
BP-933W
VT2-Sa
933W
HK022
CP933M
CP933K
CP933X
CP933H
CP933I
CP933C
Consensus/80%


BAB34583.1_933
BP-933W
VT2-Sa
933W
HK022
CP933M
CP933K
CP933X
CP933H
CP933I
CP933C
Consensus/80%


BAB34583.1_933
BP-933W
VT2-Sa
933W
HK022
CP933M
CP933K
CP933X
CP933H
CP933I
CP933C
Consensus/80%


BAB34583.1_933
BP-933W
VT2-Sa
933W
HK022
CP933M
CP933K
CP933X
CP933H
CP933I
CP933C
Consensus/80%
Key: Outlined in grey demonstrates complete identity

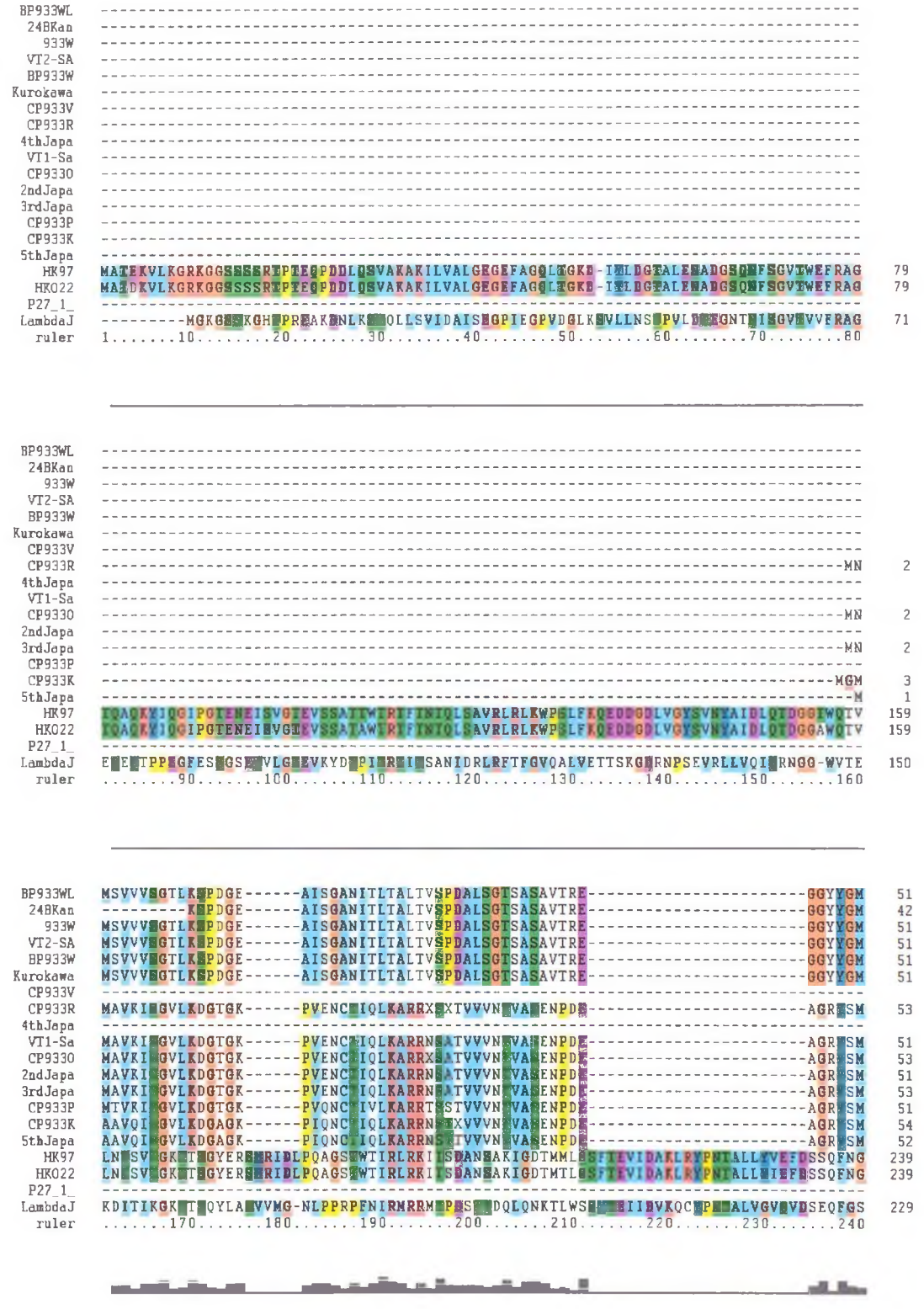


Appendix 3: Phylogenetic tree of Vpr orthologs that have > 30 % amino acid sequence identity, identified by blastp analysis of Vpr.
The tree was derived by parsimony algorithm, with branches bootstrapped from 100 analyses. Scale bar equal to 1 % amino acid change.

Appendix 4 ClustalX Protein Alignment of Long and short tailed lambdoid phage

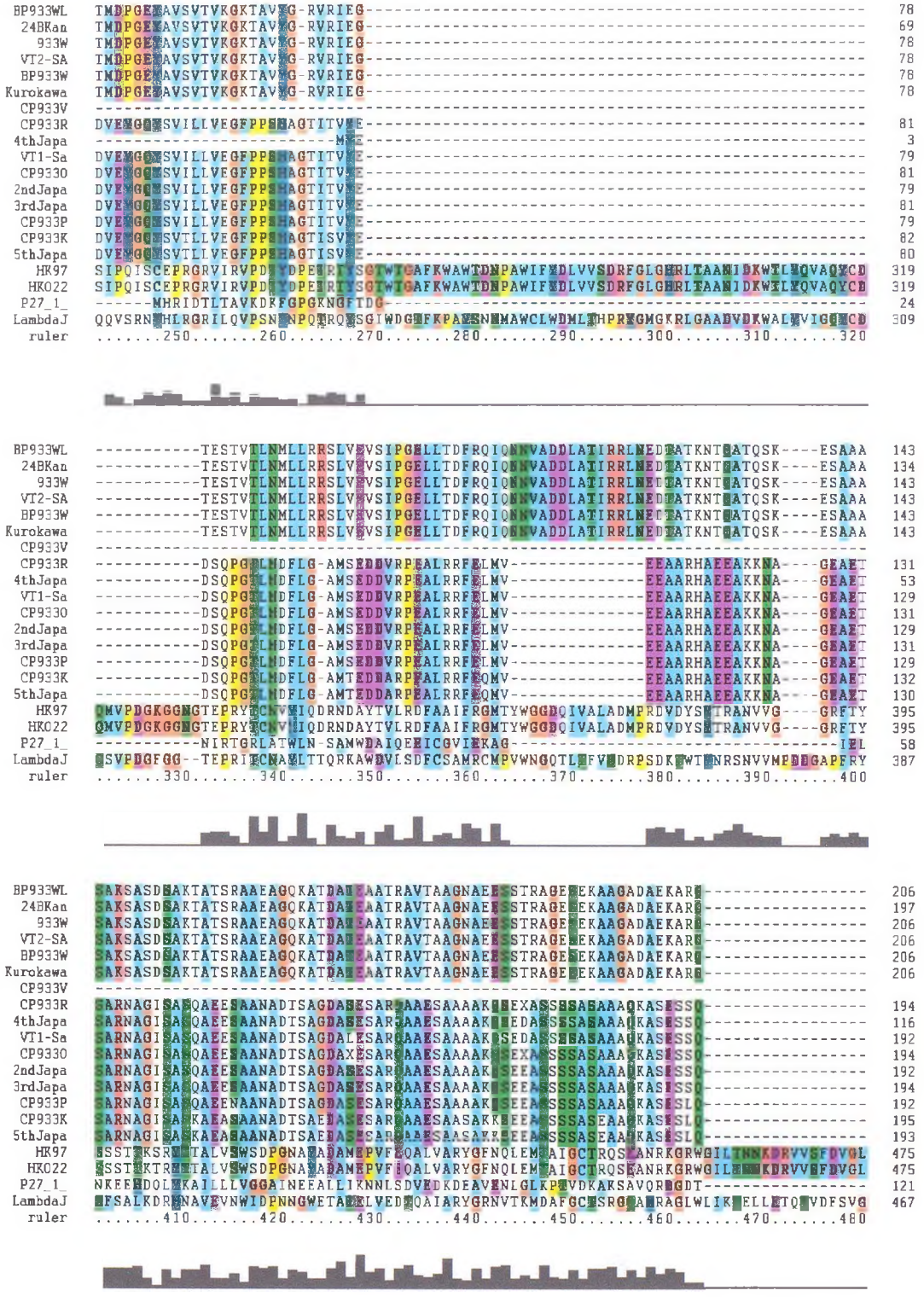
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableA.ps **Date: Mon May 27 12:37:23 2002**
Page 1 of 1



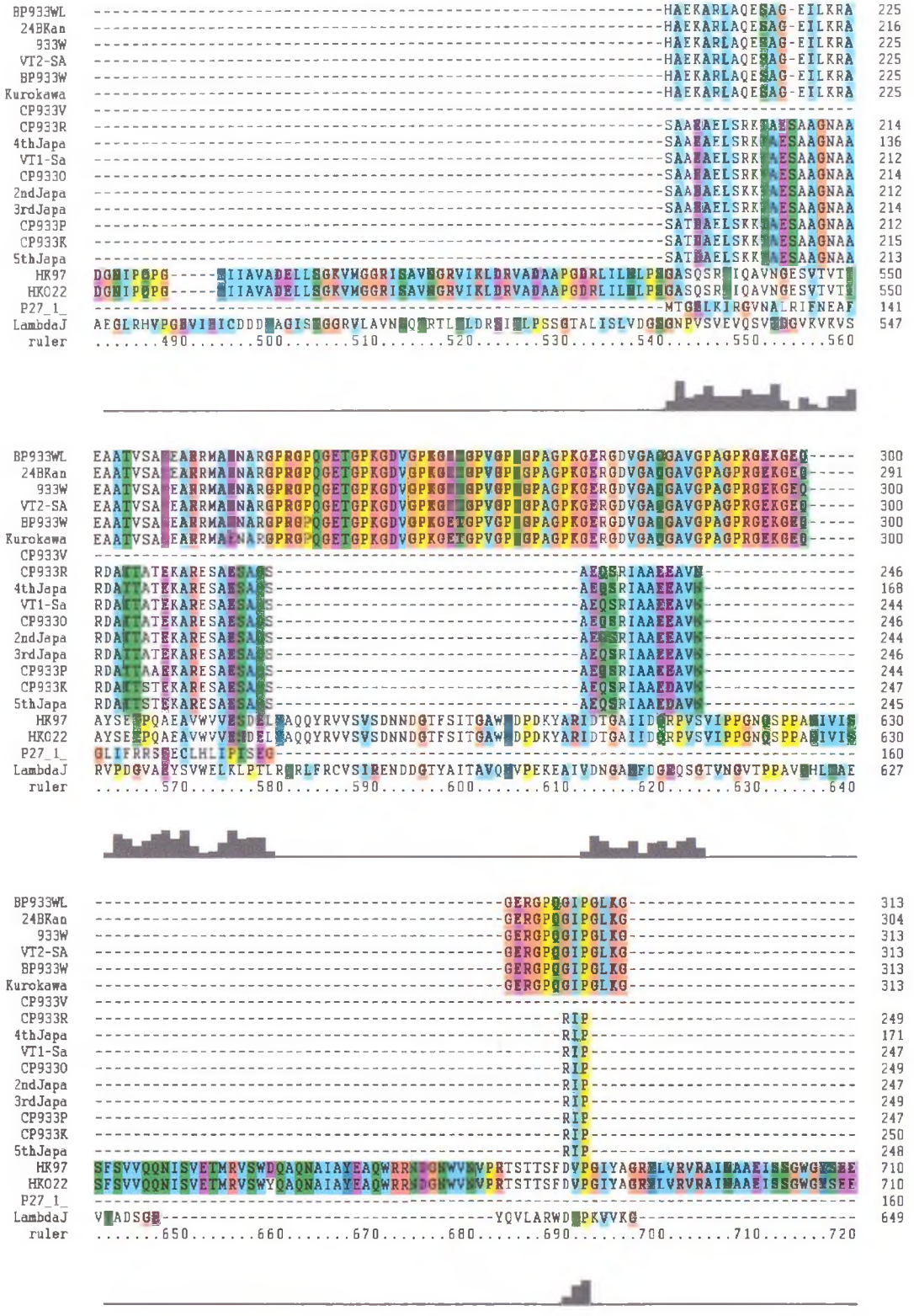
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableB.ps Date: Mon May 27 12:38:02 2002
Page 1 of 1



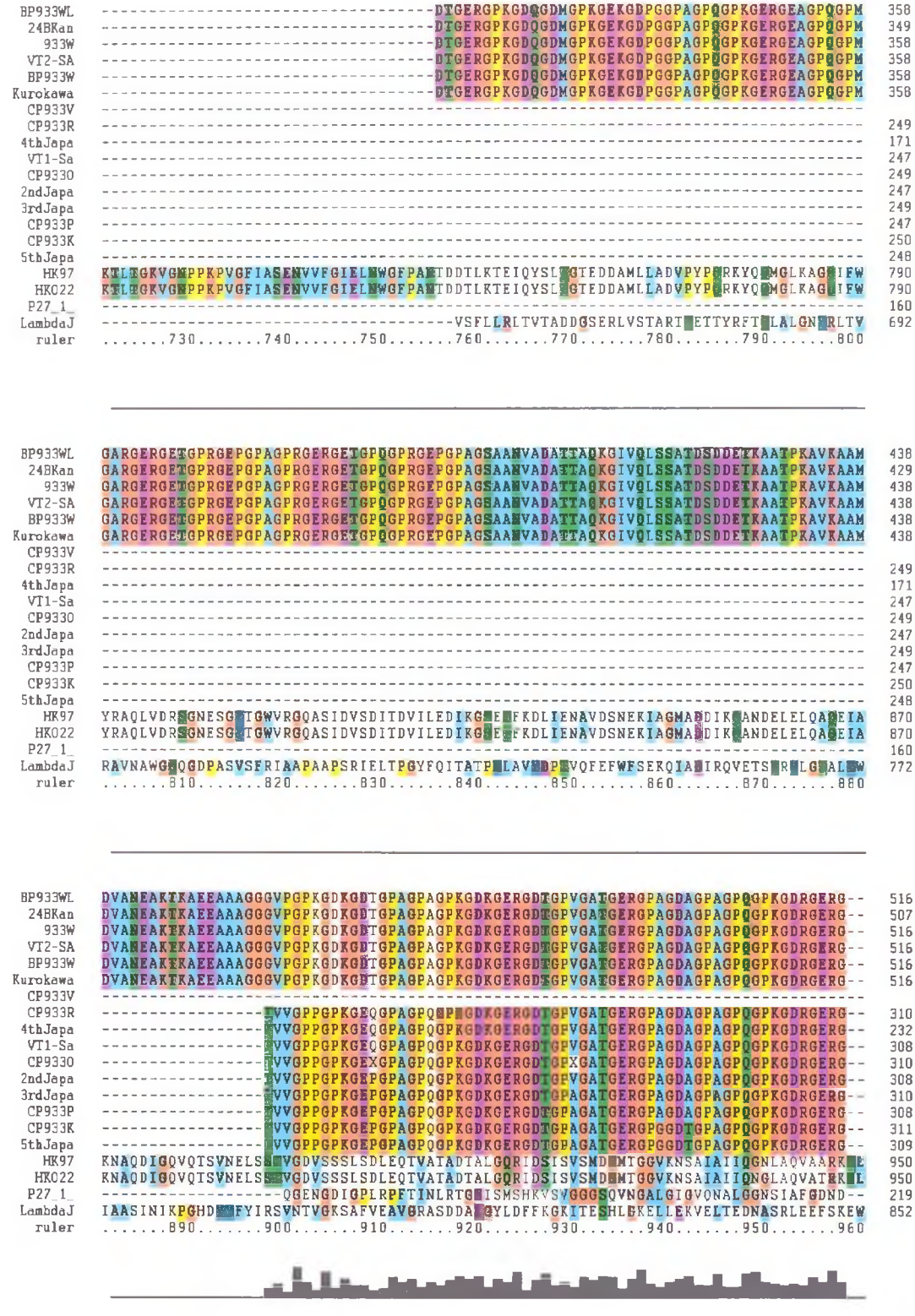
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableC.ps **Date: Mon May 27 12:38:41 2002**
Page 1 of 1



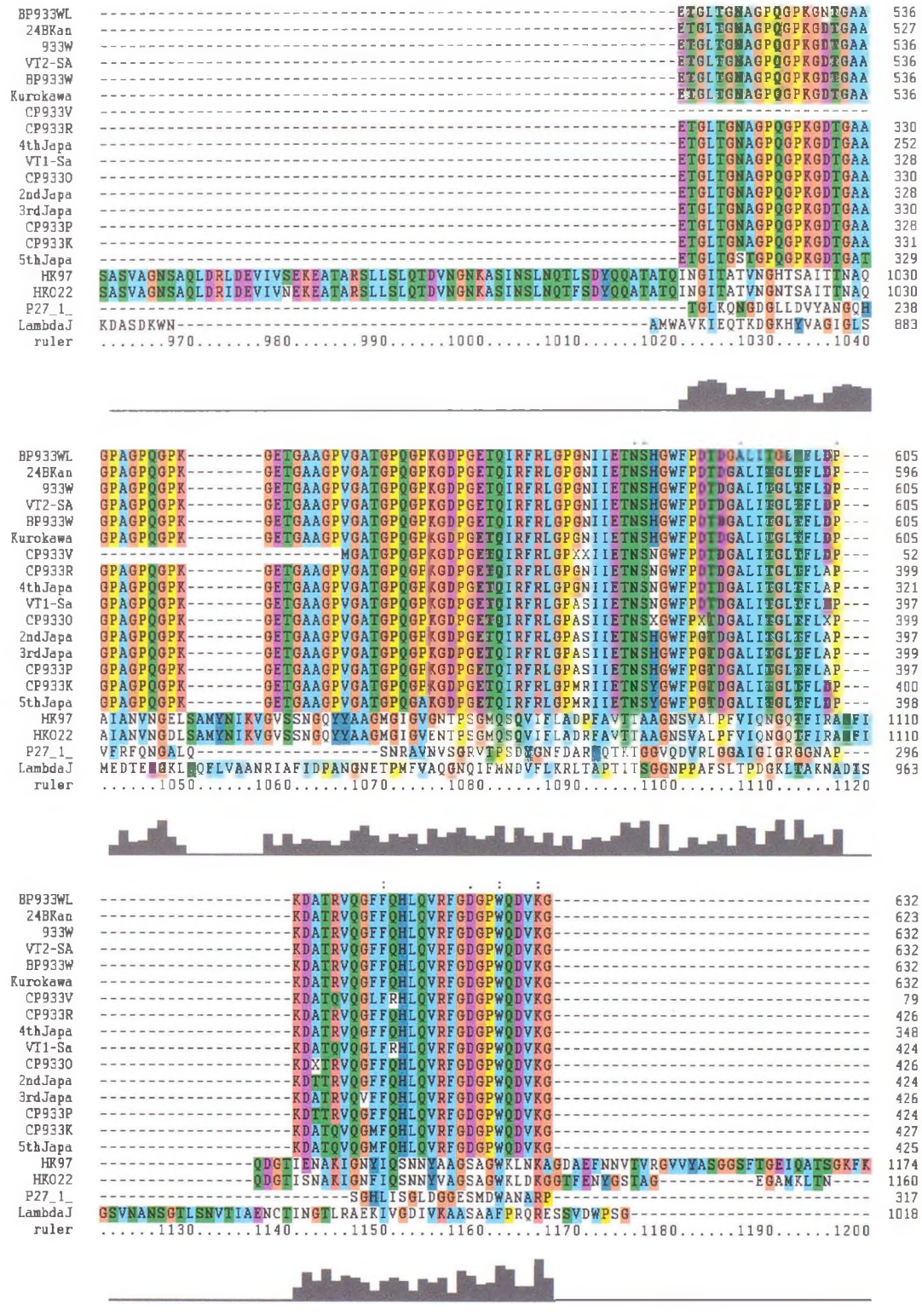
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableD.ps Date: Mon May 27 12:39:16 2002
Page 1 of 1



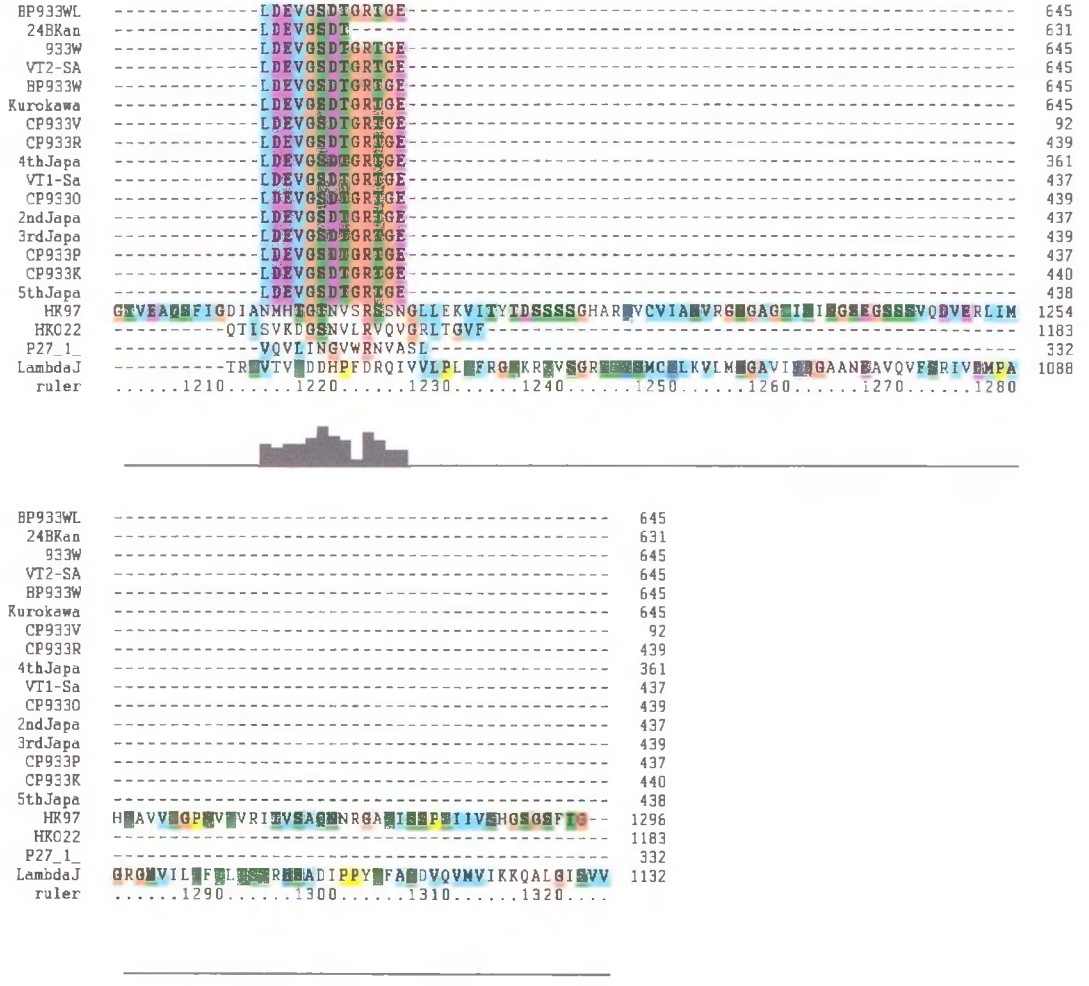
CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableE.ps Date: Mon May 27 12:39:58 2002
Page 1 of 1



CLUSTAL X (1.81) MULTIPLE SEQUENCE ALIGNMENT

File: A:all sequences availableF.ps **Date: Mon May 27 12:41:37 2002**
Page 1 of 1



Key: Shaded bar chart below the alignment details sequence identity at that amino acid position.