The role of TypA in the virulence of enteropathogenic *Escherichia coli*

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

by

Stuart Charles Clarke MSc FIBMS
Department of Microbiology and Immunology
University of Leicester

May 1999

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

Abstract

The role of TypA in the virulence of enteropathogenic *Escherichia coli*

Stuart C. Clarke

Enteropathogenic *Escherichia coli* (EPEC) remain an important cause of diarrhoeal disease in developing countries. EPEC adhere to cultured epithelial cells in a localised pattern resulting in the effacement of microvilli and subversion of the host cell signalling pathways leading to rearrangement of the cytoskeleton. This results in the 'attaching and effacing (AE) lesion'. This project is concerned with the characterisation of a novel gene, *typA* (tyrosine phosphoprotein A), and its role in virulence.

TypA and its flanking regions were cloned from EPEC strain E2348/69 and sequenced to determine differences between EPEC *typA* and *E. coli* K-12 *typA*. Few differences were found between the TypA predicted amino acid sequences. The sequence upstream of *typA* in EPEC was identical to that in *E. coli* K-12 but the first 1.3 kb of the downstream sequence showed no homology with *E. coli* K-12. The remaining 422 bp showed homology with ORF o300 in *E. coli* K-12. ORF o300 is similar to members of the phosphofructokinase B family of carbohydrate kinases.

A frameshift mutation was constructed in *typA* of EPEC E2348/69, which was also introduced into other EPEC strains. Phenotypic assays were performed on the E2348/69 *typA* mutant and its parent. Growth rate and Bfp production of the *typA* mutant were very similar. However, adherence and invasion ability was reduced 3.5-fold. FAS tests indicated that the E2348/69 *typA* mutant was still able to cause AE lesions. FAS tests were also performed on the EAF plasmid-negative JPN15 *typA* mutant and its parent; the mutant did not cause AE lesion formation. The EAF plasmid or the *per* genes were introduced into the JPN15 *typA* mutant resulting in restoration of the AE lesion phenotype. This indicated that *typA* acted as a regulator in the absence of the *per* genes and may therefore be considered a virulence gene.

Acknowledgements

Firstly I would like to thank my supervisor, Professor Peter Williams, for his faith in taking on a person with a different scientific background and providing the opportunity to study for a PhD in his laboratory. Appreciation must also be noted for the opportunities he provided throughout my PhD and for his general support and guidance. My sincere thanks go to Richard Haigh for his continuous advice and direction, particularly when I was new to the laboratory, and for finding time to discuss problems whenever I needed it. I would also like to express my gratitude to my friend Primrose who started work on the TypA protein when in Vic Norris's group; I thank her for all the interesting discussions on TypA and other areas and for the weekly issue of *New Scientist*! Thanks must also go to Martin Goldberg without whom life in the lab would have been much less interesting.

Many of the people I now wish to thank have, like myself, moved on in their careers but these people provided much friendship and discussion throughout my PhD work. Firstly, Stuart Bayliss, formerly of the PNACL, whom, after first meeting at Exeter University some years ago, continued to be a good friend and provided much of the support required for the sequencing performed during my PhD. Thanks to Rob Kingsley for introducing me to the wonders of Leicester and for stimulating discussions. I would also like to acknowledge the willingness of my colleagues in the Genetics Department laboratory 145, particularly Julian Ketley, Arnoud Van Vliet, John Henderson and Anne Wood, to discuss technical problems.

Finally, my sincere gratitude goes to my family, friends, old colleagues, and more recently my wife Tina, for their continuous support and encouragement over the many years I have now spent as a microbiologist. Without these people I would not have got this far.

For my

parents

Contents

Figures and Tables

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Abbreviations

AE	attaching and effacing
AEEC	attaching and effacing *E. coli*
bp	base pair
COSHH	Control of Substances Hazardous to Health
CTAB	hexadecyltrimethyl ammonium bromide
DMEM	Dulbecco's Modified Eagle Medium
DMF	dimethylformamide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
EAF	EPEC adherence factor
EPEC	enteropathogenic *Escherichia coli*
FAS test	fluorescein actin staining test
g	centrifugal force
HCl	hydrochloric acid
IAP	increased autoaggregation phenotype
kb	kilo-base
kDa	kilo-daltons
LA	localised adherence
LB	Luria-Bertani broth
LBA	Luria-Bertani agar
min	minute
OD	optical density
OMP	outer membrane protein

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFGE	pulsed-field gel electrophoresis
RDEC-1	rabbit diarrhoeal *Escherichia coli*
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SSC	sodium chloride/sodium citrate
Tn	transposon
μFD	micro Faradays
μg	micrograms
μl	microlitres
v/v	volume per volume
w/v	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
X-P	5-bromo-4-chloro-3-indolyl phosphate

Chapter 1

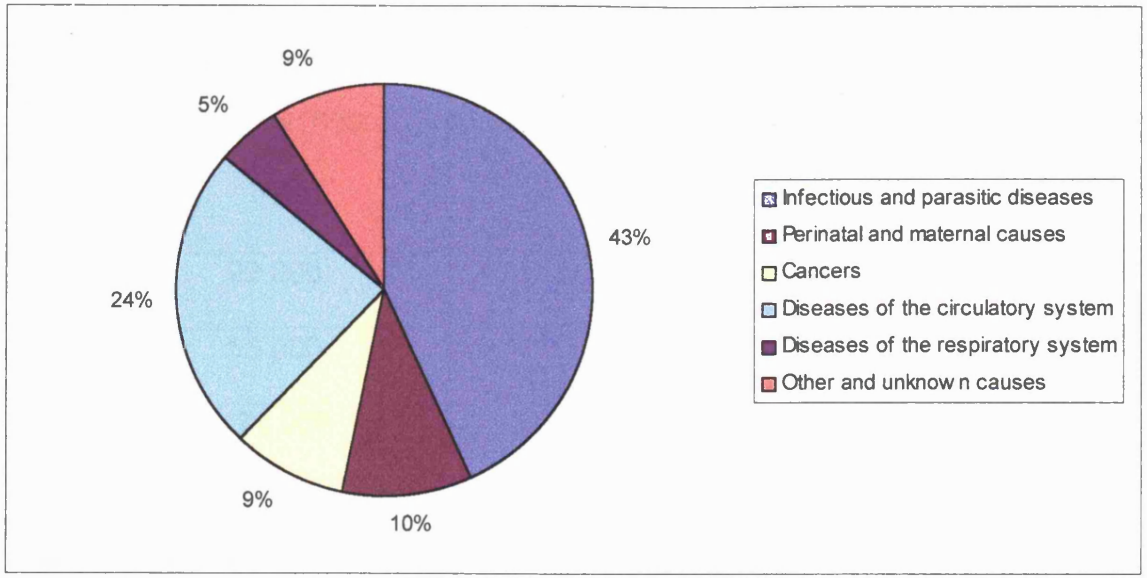
Introduction

1.1 Global health problems

According to the World Health Organisation, over 10 million children die each year before their fifth birthday, almost one half of the total number of those who die before the age of fifty (WHO 1998). This is despite most children now being immunised against all of the major childhood infectious diseases. However, the figures were once higher; the total mortality has been reduced successfully from 21 million in 1995. Even so, more than 50 million deaths occurred in total in 1997, of which 40 million were in developing countries and approximately one-third were due to infectious and parasitic diseases (**figure 1.1**). These numbers, despite continuing research and vaccine development, have not changed significantly since 1985. A large proportion of these deaths are due to diarrhoeal diseases which will be discussed in more detail in further sections. It is therefore crucial to global health that basic research into diarrhoeal and other infectious diseases continues.

1.2 The global problem of diarrhoeal illness

Diarrhoeal disease has been recognised since the beginning of civilisation, perhaps as early as 3300 BC (Kumate and Isibasi 1986), and remains one of the most important health problems of today. Approximately two-thirds of the world's population lives in areas regarded as under-developed (Levine *et al* 1986). In these areas it is estimated that over 1.3 billion cases of diarrhoeal illness occur each year (**table 1.1**) resulting in 5 million deaths (Kumate and Isibasi 1986). Of these, over 2.7 million deaths occur in children under five years of age (World Health Organisation 1994). Under-developed areas do not have adequate sanitary provisions, housing is poor and overcrowded and personal hygiene is of a low standard. Disease is therefore readily transmitted between persons and the mortality



Data source: WHO (1998)

Figure 1.1: Main causes of death in developing countries in 1997. Diarrhoeal disease remains an important problem in developing countries. Over 40 million deaths occur annually in developing countries and over one-third are due to infectious and parasitic diseases.

	Population <5 years of age x1000	Est. diarrhoea episodes per child <5 years of age per year	Est. total diarrhoea episodes in children <5 years of age x1000
Africa	99 000	4.9	485 100
America	55 000	3.5	192 500
South east Asia	176 000	1.8	316 800
East Mediteranean region	68 500	3.7	253 450
West Pacific region (ex. China)	24 500	2.5	61 250
Global total (ex. China)	**423 000**	**3.1**	**1 311 300**

Table 1.1 Diarrhoeal episodes in children under the age of five years in developing countries excluding China 1993

Adapted from WHO Ninth Programme Report for Diarrhoeal Diseases 1994

rate is much higher than in developed countries due to the lack of supportive medicine. Diarrhoeal disease is largely a problem in children and is the second most common illness in this age group after respiratory disease (DuPont 1995); it therefore remains a problem in both developed and developing countries although it is a greater problem in the latter. The rates of diarrhoeal disease range from an average of 1.7 episodes per person per year in the United States to an average of 17 episodes per person per year in some developing countries (Guerrant and Bobak 1991). Diarrhoeal disease may be caused by a number of bacterial, parasitic or viral aetiologies which will be discussed further below. Certain species within each group are very important causes of diarrhoea resulting in millions of deaths worldwide each year. The final outcome of diarrhoeal illness depends on numerous factors including age, breast feeding, the pathogen involved, complications such as dehydration and renal failure, and the availability of medical treatment (DuPont 1995).

1.3 Gastrointestinal pathogens

There are numerous causes of gastrointestinal illness in humans. These include various genera and species of the viruses, protozoa and bacteria. The search for aetiological agents of diarrhoea began in 1875 with the discovery of *Entamoeba histolytica* by Lösch in Russia (Kumate and Isibasi 1986). Since then, new species of organisms have been continually discovered as 'new' aetiological agents of diarrhoeal disease. Even so, in some developing countries the known causes of diarrhoea account for only 60% of all episodes with the remaining 40% due to unknown aetiologies (Dr Sankaran, Anna University, Madras, India; personal communication). Additionally, the costs associated with the diagnosis of gastrointestinal pathogens is high and is estimated to be between $900 and $1000 for a positive case in the US (Guerrant *et al* 1985). There have been clear advances in recent years in the fields of research related to gastrointestinal infectious disease (Guerrant and Bobak 1991). Newly-identified pathogens have been described such as *Helicobacter pylori* and various pathovars of *Escherichia coli*. However, much more is to be learnt before effective therapies or vaccines can be produced.

1.3.1 Viral gastrointestinal pathogens.

Viruses are an important cause of diarrhoeal disease worldwide and account for approximately 10% of cases of traveller's diarrhoea (Farthing 1994). Rotavirus and Norwalk virus are the most common causes of viral diarrhoea (Kumate and Isibasi 1986; DuPont and Ericsson 1993) although other viruses such as the small round structured viruses, adenoviruses and astroviruses are also important (Crowley *et al* 1997). Symptoms of viral diarrhoea usually have an abrupt onset but are usually short-lived and include diarrhoea, vomiting and possibly mild fever (Lewis and Matsui 1993). Diarrhoeal viruses are primarily spread via contaminated water, particularly during rainy seasons in developing countries (DuPont and Ericsson 1993).

1.3.2 Protozoan and nematode gastrointestinal pathogens.

Diarrhoeal disease due to protozoan and nematode infections are numerous and include organisms as diverse as the flagellates, amoebae, coccidia, helminths and worms (Cook 1996). Not all genera within these groups are pathogenic but those that are, are responsible for a whole range of diseases in humans that may be acute or chronic (Cook 1996). Many of those that cause diarrhoeal disease are water-borne and are responsible for either sporadic cases or outbreaks. Two important protozoan diseases are caused by *Giardia* and *Cryptosporidium* and these two organisms are estimated to infect more than 600 million people worldwide (Smith H, personal communication). Another, *Entamoeba histolytica* (now known as *Entamoeba dispar*), is one of the most common infections in the world today accounting for 500 million infections with an estimated 100 000 deaths each year (WHO 1992). In addition, nematode infections account for a further 3.5 billion infections, the majority of which are due to just four organisms, namely *Ascaris lumbricoides* (roundworm), *Trichuris trichiuria* (whipworm), *Necator americans* and *Ancylostoma duodenale* (hookworms) (Chan 1997).

1.3.3 Bacterial gastrointestinal pathogens.

There are three types of bacterial gastrointestinal infection according to the site of pathogenesis and, therefore, to the causative organism (Guerrant and Bobak 1991).

Most gastrointestinal infections are non-inflammatory occurring in the small bowel, and are caused by the action of an enterotoxin such as cholera toxin, or by a pathogenic process such as that of enteropathogenic *E. coli*. Inflammatory infection occurs in the colon and is caused by an invasive pathogen such as *Shigella*. The third type, characterised by enteric fever, is caused by pathogens such as *Salmonella typhi* which enter the Peyer's patches and regional lymph nodes. Such pathogens cause systemic infection before returning to the gastrointestinal tract via the biliary tract.

Throughout the world the species of bacterial agents of diarrhoeal illness are very similar. In developing countries the most common bacterial causes of diarrhoea are enterotoxigenic *Escherichia coli*, *Vibrio cholerae*, *Salmonella typhi*, *Shigella flexneri*, *Shigella dysenteriae* and enteropathogenic *Escherichia coli*, whilst in developed countries campylobacter, non-typhoid salmonellae and *Shigella sonnei* are most common. *Campylobacter spp.*, *Yersinia spp.*, non-cholera vibrios, *Staphylococcus aureus* and *Clostridium perfringens*, occur in both developing and developed countries (Kumate and Isibasi 1986).

1.4 *Escherichia coli* as gastrointestinal pathogens

Escherichia coli was first described by the German paediatrician, Theodore Escherich (Escherich 1885). It is a Gram negative rod-shaped bacterium of the family *Enterobacteriaceae*. Strains of *E. coli* occur as gut commensals and are normally excreted in faeces (Hart *et al* 1993) and for many years it was thought that *E. coli* were not pathogenic. However, it has been found that *E. coli* may cause a number of diseases in humans, such as diarrhoeal disease, urinary tract infection, meningitis and various wound infections. These diverse diseases reflect their ability as facultative anaerobes to survive under vastly different environmental conditions.

Certain *E. coli* are a cause of diarrhoeal illness around the world. Except for traveller's diarrhoea and haemorrhagic colitis, the latter of which may be epidemic or sporadic, they are a rare cause of gastroenteritis in developed countries (Hart *et al*

1993). Initially, all diarrhoeagenic strains of *E. coli* were termed enteropathogenic *E. coli* (EPEC) but as more has been learnt about the pathogenic mechanisms of these *E. coli*, they have been grouped accordingly (Kaper 1994). There are now seven classes of diarrhoeagenic *E. coli* with only the classical serotypes still being termed EPEC. Strictly, the term 'EPEC' should be reserved for those EPEC which do not belong to any of the other diarrhoeagenic *E. coli* classes and do not produce any enterotoxin but do cause characteristic attaching and effacing (AE) lesions on the intestinal brush border (Edelman and Levine 1983). However, these organisms may generally be known as attaching and effacing *E. coli* (AEEC). The AE phenotype is described in detail in later sections. Other diarrhoeagenic *E. coli* continue to be described for which the virulence mechanisms are unknown (Monteiro-Neto *et al* 1997); such *E. coli* do not possess certain genes associated with other *E. coli* classes and therefore other, as yet undetermined, virulence genes may be involved.

The seven classes of diarrhoeagenic *E. coli* are enteropathogenic *E. coli* (EPEC), enterohaemorrhagic *E. coli* (EHEC), enteroinvasive *E. coli* (EIEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAggEC), diffuse-adherent *E. coli* (DAEC) and cytolethal distending toxin (CDT)- producing *E. coli*. These will be described in detail on the following pages. EPEC and ETEC are the most important of these in terms of total diarrhoeal episodes on a global scale. For example, in one study in Bangladesh, ETEC and EPEC were the only pathogenic *E. coli* associated with diarrhoea in children under five years of age (Albert *et al* 1995). EHEC is becoming more significant in developed countries because of numerous outbreaks in recent years (Schmidt *et al* 1997).

The evolution of *E. coli* as a gastrointestinal pathogen is probably comparatively recent. Although *E. coli* is found in the environment and occurs as a commensal in the mammalian gut, horizontal gene transfer events have allowed the transition of some *E. coli* strains from commensals to pathogens (Bäumler *et al* 1997), usually through the acquisition of a pathogenicity island (Lee 1996). These transfer events occurred independently of each other with different sets of virulence genes.

Therefore, except for enterohaemorrhagic *E. coli*, all pathovars of *E. coli* possess different virulence mechanisms when compared with EPEC. EHEC, in contrast, is a clonal group derived from enteropathogenic *E. coli* (Whittam *et al* 1993). These events are proven by experiments on non-pathogenic *E. coli*. Many commensal enteric bacteria, including *E. coli*, already possess many of the genes required for host cell interaction but lack certain genes to make them fully pathogenic (Groisman and Ochman 1997). However, if particular gene clusters, such as the AE genes of EPEC, are transferred from pathogenic bacteria to commensal *E. coli*, then the *E. coli* become fully pathogenic (McDaniel and Kaper 1997). This is confirmed by a recent study (Pupo *et al* 1997) which suggests that pathogenic strains of *E. coli* do not originate from a single ancestor but, instead, have arisen several times from several ancestors. Furthermore, it was recently reported that an enteroaggregative *E. coli* strain of serotype O111:H2 produced the Shiga toxin associated with *Shigella dysenteriae* (Morabito *et al* 1998). The strain was the cause of an outbreak of haemolytic-uraemic syndrome and may represent a potential source of future EHEC-like disease. In addition to gene transfer, genomic instability is a phenomenon which has gained considerable interest recently. Genomic rearrangements have been reported in various bacteria, most recently *Campylobacter jejuni* (Wassenaar *et al* 1998) and *Yersinia pestis* (Buchrieser *et al* 1998). The role of genomic instability in virulence is unknown but, as organisms need to adjust to their environment, then such instability is probably important.

1.4.1 Enterohaemorrhagic *Escherichia coli.*

EHEC belong to a number of O serogroups but those of serogroup O157 are the most important in human disease (Bolton and Aird 1998; Paton and Paton 1998). The organism produces a toxin which has cytotoxic activity on Vero cells and is often termed Vero cytotoxin (VT) but may be also be known as Shiga-like toxin. EHEC are therefore also termed verocytotoxigenic *E. coli* (VTEC) and Shiga toxin-producing *E. coli* (STEC). EHEC also possesses a number of other mechanisms, such as adhesins, which are involved in virulence. EHEC was first identified as a human pathogen in 1982 after two outbreaks occurred in the US (Wall *et al* 1996).

Since then outbreaks have been reported from various parts of the world including North America, Western Europe, Australia, Asia and South Africa.

Transmission of EHEC may occur by the ingestion of contaminated food such as meats, unpasteurised milk or water, by person-to-person contact or by zoonotic transmission (Wall *et al* 1996). Cattle appear to be the main reservoir of EHEC and the organism has been isolated from apparently healthy animals. Human infection with EHEC is associated with a range of symptoms from non-bloody diarrhoea, fever and vomiting through to haemorrhagic colitis and haemolytic uraemic syndrome (HUS) (Coia 1998). HUS is characterised by acute renal failure, haemolytic anaemia and thrombocytopaenia. Between 2% and 7% of infected individuals usually develop HUS and it is more common in young children. The infectious dose has been estimated to be less than 100 organisms as determined from outbreak investigation (Nataro and Kaper 1998). The incubation period is usually one to six days but may be up to 14 days and the infection is usually self-limiting, often resolving within eight days. There is no specific treatment for EHEC infection; therapy is symptomatic only and antibiotic therapy, as in many gastrointestinal infections, may not be beneficial. Since the early 1990s there has been a general global increase in the number of EHEC infections and recent large outbreaks have led to an increased interest in the pathogen (Anon 1996a; Anon 1996b; Anon 1996c; Anon 1997, Clark *et al* 1997, McDonnell *et al* 1998). This is thought to be mostly due to an increase in the consumption of fast foods. The use of such food preparation has increased significantly in recent years and holds a greater risk of foods being improperly cooked.

1.4.2 Enteroinvasive *Escherichia coli.*

Enteroinvasive *E. coli* are an important cause of diarrhoeal disease. Infection results in a bacillary dysentery-like illness characterised by a bloody, mucoid diarrhoea. Not surprisingly, therefore, they share some of the virulence properties associated with *Shigella dysenteriae* (Hart *et al* 1993). EIEC are a significant cause of morbidity and mortality in young children in developed countries although, like most enteropathogens, they are more important in developing countries where

sanitation and hygiene levels are of a poor standard. EIEC attach to and subsequently invade the colonic enterocytes by endocytosis. The bacteria multiply within these cells eventually causing their death and the bacteria are released back into the colon after which they invade other cells. This process results in an inflammatory response accompanied by necrosis and ulceration of the large bowel leading to release of blood and mucous in the stools (Hart *et al* 1993). The ability of EIEC to invade, survive and multiply within colonic enterocytes is dependent upon the presence of a 120-140 MDa plasmid which encodes all the genes necessary for such virulence (Hart *et al* 1993) including outer membrane proteins required for invasion (Levine 1987).

1.4.3 Enterotoxigenic *Escherichia coli.*

ETEC were first identified as a cause of diarrhoea in 1970 and are now considered a major cause of *E. coli*-associated diarrhoea worldwide (Hart *et al* 1993). The symptoms of ETEC infection are due to the production of one or both of two types of enterotoxins by the bacterium, namely heat stable enterotoxins (ST) and heat labile enterotoxins (LT).

ST's are small, monomeric plasmid-mediated toxins comprising two unrelated classes, namely STa (or ST-I) and STb (Nataro and Kaper 1998). The receptor for STa is guanylate cyclase C and binding results in increased intracellular cyclic guanylate monophosphate (cGMP) levels leading to stimulation of chloride secretion or inhibition of sodium chloride absorption and subsequent intestinal fluid secretion. The receptor for STb is unknown but it causes histologic changes in the form of loss of villous epithelial cells and partial villous atrophy.

The LT toxins are oligomeric and consist of two major serogroups, namely LT-I and LT-II (Nataro and Kaper 1998). LT-I is related to the cholera toxin (Hart *et al* 1993); it is composed of one A subunit and five identical B subunits (Nataro and Kaper 1998). The latter bind to the ganglioside GM_1. The A subunit is responsible for the enzymatic activity of the toxin by activating adenylate cyclase resulting in an intracellular increase in cyclic adenosine monophosphate (cAMP) concentrations.

This leads to a decrease in sodium absorption by villous cells and subsequent active chloride secretion by crypt cells thereby resulting in osmotic diarrhoea toxin (Hart *et al* 1993; Nataro and Kaper 1998).

Like EIEC, ETEC are associated with poor hygiene and sanitation and are therefore of most importance in developing countries. In such countries the number of diarrhoeal episodes due to ETEC is high and is equivalent to that caused by EPEC. Many individuals visiting developing countries suffer from ETEC infection during their stay or upon their return. Most of these cases probably occur due to the ingestion of water or food contaminated with ETEC and outbreaks can occur in hotels after person-to-person spread. The illness is known as traveller's diarrhoea and is usually an acute illness with symptoms consisting of loose stools, nausea, vomiting and abdominal cramps. In developing countries the most severe form of acute infection is a cholera-like illness which occurs in areas where *Vibrio cholerae* is already endemic. It may occur in all age groups and is difficult to distinguish clinically from true cholera. ETEC is not usually a problem in developed countries although it has been linked with outbreaks, some of which have been large with approximately 2000 individuals being infected. ETEC has been associated with a few outbreaks of infantile enteritis in hospitals and also with larger outbreaks due to contaminated water or food (Gross 1990).

1.4.4 Enteroaggregative *Escherichia coli*

EAggEC are a more recent addition to the diarrhoeagenic classes of *E. coli* and are so named due to their adherence pattern to cultured epithelial cells. Such an adherence pattern provides the gold standard for the definition of EAggEC and is seen as bacteria adhering in a "stacked brick-like" formation on the cell surface (Elliot and Nataro 1995). EAggEC are a significant cause of diarrhoea in developing countries and are epidemiologically associated with acute or persistent diarrhoea, the latter of which is of particular significance in children (Nataro *et al* 1998; Law and Chart 1998). Furthermore, EAggEC are a cause of diarrhoea in travellers who have visited countries endemic for EAggEC (Gascon *et al* 1998). However, a high rate of asymptomatic carriage often occurs with EAggEC infection and therefore the

presence of other pathogens should be excluded (Law and Chart 1998). EAggEC do not induce attaching/effacing lesions characteristic of enteropathogenic *E. coli* or produce cholera-like or Shiga-like toxins characteristic of *Vibrio cholerae* or *Shigella dysenteriae* respectively. Little is known of their pathogenicity except that they produce a toxin known as the enteroaggregative heat-stable enterotoxin (EAST1) (Elliot and Nataro 1995) and that their aggregative adherence phenotype is mediated by at least two fimbriae (Nataro *et al* 1985; Czeczulin *et al* 1997). The EAST1 enterotoxin is partly homologous and functionally similar to, but distinct from ST, and is thought to be responsible for the symptoms of infection. The production of a heat labile toxin which is antigenically related to *E. coli* haemolysin has also been reported from EAggEC (Baldwin *et al* 1992). More recently, a cytotoxic 108 kDa heat-labile enterotoxin has been described (Eslava *et al* 1998) although its effects *in vitro* or *in vivo* have not yet been fully described (Navarro-García *et al* 1998). There are currently two fimbriae which have been characterised from EAggEC. They are termed aggregative adherence fimbriae I and II (AAFI and AAFII) and are plasmid-encoded (Nataro *et al* 1985; Czeczulin *et al* 1997) but are genetically, phenotypically and morphologically distinct (Czeczulin *et al* 1997).

1.4.5 Diffuse-adherent *Escherichia coli*

DAEC, also known as cell-detaching *E. coli*, have been described as a cause of diarrhoea and are so named because of their diffuse adherence pattern to cultured epithelial cells (Elliot and Nataro 1995, Gunzburg *et al* 1993). Their aetiological role in diarrhoea has been a controversial issue because some studies have shown an association of DAEC with diarrhoea (Girón *et al* 1991; Gunzburg *et al* 1993), whilst others have not (Albert *et al* 1995). Individuals infected with DAEC may experience mucoid watery stools and suffer from fever and vomiting with a mean duration of illness of about 8 days. DAEC have been reported as a cause of diarrhoeal illness in both developed and developing countries. Their pathogenic mechanisms are poorly understood although it is thought that DAEC may produce one or more toxins, one of which is a haemolysin (Elliot and Nataro 1995). Recently DAEC have been studied further by Elliot and colleagues (1998a). Their preferred name is now diarrhoea-associated haemolytic *E. coli* (DHEC). Data from this study strongly

support previous studies showing that DHEC are virulent strains. It was shown that their virulence is mediated by α hemolysin and cytotoxic necrotising factor 1.

1.4.6 Cytolethal distending toxin-producing *Escherichia coli.*

Cytolethal distending toxin (CDT)-producing *E. coli* are the latest class of diarrhoeagenic *E. coli* to be described. These *E. coli* have been reported in a number of studies (Bouzari *et al* 1990; Bouzari *et al* 1992; Albert *et al* 1996) and cause the distension and eventual disintegration of cells of certain lines (Pickett *et al* 1994). In one report (Albert *et al* 1996), CDT-producing *E. coli* additionally possessed the virulence factors of EPEC or EAggEC. However, the association of CDT-producing *E. coli* with diarrhoea is not clear because in one report in Bangladesh (Albert *et al* 1996) there was no statistically significant association of carriage with diarrhoea. Therefore results of further studies are required to confirm or disprove the diarrhoeagenic potential of CDT-producing *E. coli.* Some species of campylobacter also produce a cytolethal distending toxin related to that of CDT-producing *E. coli* and which also causes distension and eventual disintegration of cells of certain lines (Pickett *et al* 1996). Additionally, CDT-producing *E. coli* have recently been reported as a cause of mortality in birds (Foster *et al* 1998).

1.4.7 Enteropathogenic *Escherichia coli.*

EPEC remain a major cause of infantile diarrhoea worldwide and possess a number of virulence factors that enable them to cause disease (Kaper *et al* 1997). Before the explosion of molecular biology in research, relatively little was known about the virulence mechanisms of EPEC although large amounts of epidemiological data were available (Frant and Abramson 1938; Ewing *et al* 1963; Moffet *et al* 1968). However, there has been a huge surge of interest in EPEC research in the past decade leading to a significant increase in knowledge of the virulence mechanisms of EPEC. It is now known that they possess a virulence plasmid and numerous chromosomally-encoded genes which together are responsible for virulence factors such as bundle-forming pili, secreted proteins and intimin (Nataro and Kaper 1998). EPEC will be discussed in more detail in the next section.

1.5 Enteropathogenic *Escherichia coli* - a historical overview

The disease described here as infantile diarrhoea has been noted for a number of centuries. Infantile diarrhoea has had a number of synonyms during the last four centuries including the terms 'griping in the guts', *cholera infantum*, 'summer diarrhoea' and 'gastro-enteritis' (Thomson 1956). Up to and including the early parts of this century, infantile diarrhoea was a major problem worldwide with high morbidity and mortality. From 1920, diarrhoea-associated mortality in infants decreased in developed countries resulting in a decline in interest to determine the aetiology of infantile diarrhoea (Robins-Browne 1987), but during the 1930s a number of severe nosocomial outbreaks of neonatal enteritis occurred in New York, each with a high mortality (Frant *et al* 1938). In 23 outbreaks occurring in 15 hospitals there was a total of 711 cases of hospitalised infants with infantile diarrhoea and the average morbidity and mortality rates were 15.5% and 7.3% respectively (although these ranged from 4% to 49% and 0% to 21.5% respectively). This led to a renewed interest in the aetiology of infantile diarrhoea.

Bray (1945) reported the isolation of diarrhoeagenic *E. coli* from cases of infantile summer diarrhoea although the intestinal pathogenicity of *E. coli* had already been noted in 1889 (Laurelle 1889) and 1897 (Lesage 1897). Lesage suggested that there were pathogenic and non-pathogenic strains of *E. coli* on the basis that convalescent serum from a patient with diarrhoea agglutinated bacteria from other patients with diarrhoea during an epidemic. The same serum did not agglutinate bacteria from healthy individuals (Lesage 1897). Early epidemiological investigations utilised carbohydrate fermentation (Bray 1945) or slide agglutination tests (Bray and Bevan 1948) to identify diarrhoeagenic *E. coli*. In 1947, Kauffman published a serotyping scheme based on somatic (O), flagellar (H) and capsular (K) antigens (Kauffman 1947) which provided a reliable method of typing diarrhoeagenic *E. coli*. This method is still used today for the serotyping of *E. coli* and other members of the *Enterobacteriaceae*. Epidemiological studies in the UK to determine the source of diarrhoeagenic *E. coli* were mostly unsuccessful in the first half of this century. However, in the mid 1950s, serogroups of *E. coli* associated with infantile diarrhoea

were isolated from cow's milk (Thomson 1956). It is unlikely that such *E. coli* could be isolated from commercially-prepared milk today because of stringent bacteriological testing and pasteurisation, although some raw milks sold on farms possess the danger of containing diarrhoeagenic *E. coli.*

The term 'enteropathogenic *E. coli*' (EPEC) was introduced in 1955 to describe strains of *E. coli* implicated epidemiologically with infantile diarrhoea (Neter *et al* 1955). EPEC are not a homogeneous group of enteropathogens (Clausen *et al* 1982) because serotype does not always correlate with pathogenicity (Ewing *et al* 1963; Goldschmidt *et al* 1976). Some serotypes of *E. coli*, which may be assigned classical EPEC serotypes, produce a heat-labile enterotoxin whilst others may be invasive. Such serotypes should not strictly be classed as EPEC; those EPEC serotypes that are most commonly associated with disease are listed in **table 1.2** below.

Table 1.2 *E. coli* serotypes commonly associated with infantile diarrhoea

O18:H7	O18:H14	O20:H-	O20:H26	O26:H-	O26:H11	O28:H-	O44:H18
O44:H34	O55:H-	O55:H6	O55:H7	O86:H34	O111:H-	O111:H2	O111:H12
O111:H5	O111:H7	O112:H-	O114:H2	O119:H6	O124:H-	O124:H30	O125:H21
O126:H-	O126:H27	O127:H-	O127:H6	O127:H21	O128:H2	O128:H7	O128:H12
O142:H6	O158:H23						

Modified from WHO 1987 and Law 1994

Work which provided much of this information (Ewing *et al* 1963) was important because it provided a list of serotypes of diarrhoeagenic *E. coli* which has remained comprehensive with only a few serotypes being subsequently added to the list. As the early methods of identifying EPEC relied on serotyping alone, there was always some reluctance in accepting the organisms as pathogens. It could be argued that changes in the gut epithelium due, for example, to infection by an unknown bacterium or virus allowed the multiplication of the EPEC serotypes. However, the pathogenic ability of EPEC was confirmed by Levine *et al* (1978); they showed that

strains of EPEC which did not produce enterotoxins, were not invasive and were negative in the infant rabbit assay for gross fluid accumulation caused diarrhoea when given to adult volunteers. No heat-labile or heat-stable enterotoxins were detected in *E. coli* isolated from volunteers with diarrhoea and volunteers did not show a rise in LT antitoxin titre.

As little as three decades ago, EPEC was a significant cause of infantile diarrhoea in developed countries, accounting for 16% of cases in one study (Moffet 1968). However, the mortality rate of EPEC infection has fallen considerably since the beginning of this century apart from a significant rise during the First World War (**figure 1.2**) (Thomson 1956). Since then the mortality rate continued to fall from 70% to 25% in the two decades up to 1970 (Taylor 1970). EPEC infection in developed countries is now rare and is no longer regarded as a problem. The last severe outbreak of infantile diarrhoea due to EPEC in the UK occurred in 1980 (Scotland 1983) in the paediatric unit of a hospital and was initiated by the introduction onto the ward of an infant suffering from diarrhoea. Thirty infants were affected and the outbreak resulted in two deaths. Although EPEC is no longer a problem in developed countries, it is still isolated from young children as sporadic cases, as shown by recent studies (Hill *et al* 1991; Crowley *et al* 1997). In a US study, it was found that 25 of 445 children in a paediatric hospital were excreting EPEC (Bokete *et al* 1997), but it was not known whether such EPEC were actually causing disease. In another study, in the UK (Crowley *et al* 1997), the aetiology of gastroenteritis in children under the age of 5 years was investigated. *E. coli* was identified in 2% of reports and of these 44% were of classical EPEC serotypes. Due to the low incidence of EPEC infection in the UK, it has been proposed that screening for the organisms is not necessary (Morris and Rao 1992). Many sporadic EPEC infections seen in developed countries occur due to foreign travel and are probably acquired in developing countries. Such travel may lead to exposure to previously unencountered pathogens.

EPEC are now generally defined as "diarrhoeagenic *E. coli* belonging to serogroups epidemiologically incriminated as pathogens but whose pathogenic mechanisms

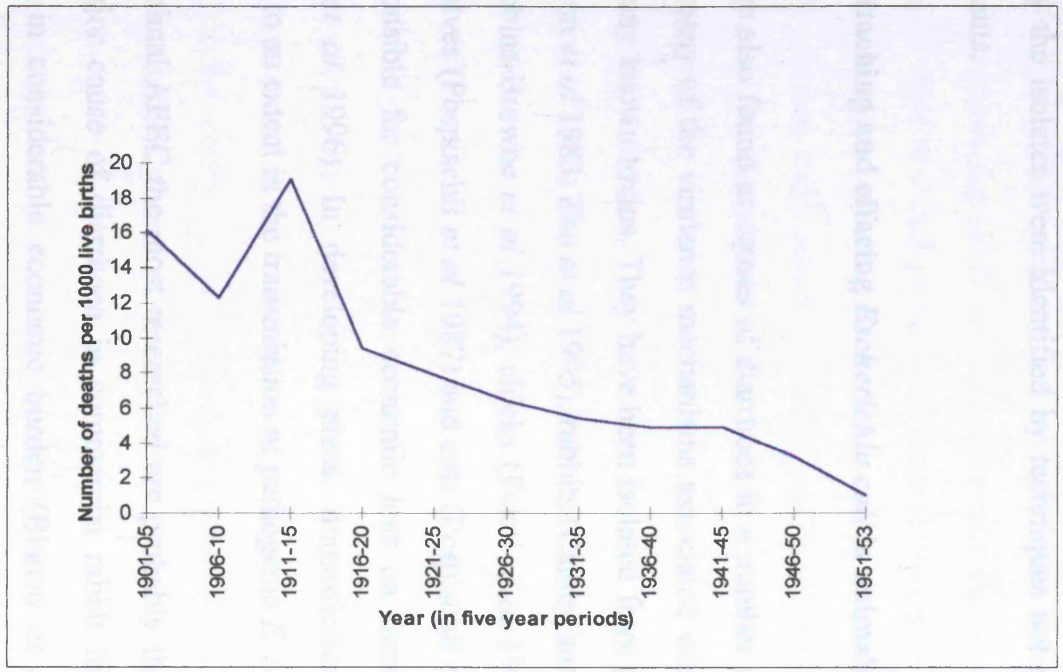


Figure 1.2 Summer diarrhoea as a cause of death in babies under one year of age in England from 1901 to 1953

Data source: Thomson (1956).

have not been proven to be related either to LT enterotoxins or ST enterotoxins or to Shigella-like invasiveness. EPEC adhere in a seemingly pathognomic way to the intestinal epithelium" (Edelman and Levine 1983). Recently, a study has shown that strains serologically identified as EPEC prior to 1960 and isolated from infants with diarrhoea are, in fact, EPEC as confirmed by molecular methods (Robins-Browne *et al* 1993). This is an important finding because the study identified these isolates on the basis of their possessing virulence determinants associated with EPEC, whereas originally the isolates were identified by techniques not associated with virulence determinants.

1.6 Attaching and effacing *Escherichia coli* in animals

AEEC are also found as causes of diarrhoea in a number of animals. Such bacteria possess many of the virulence mechanisms associated with EPEC and also do not produce any known toxins. They have been isolated from goats (Drolet *et al* 1994), pigs (Moon *et al* 1983; Zhu *et al* 1995), rabbits (Cantey and Blake 1997; Moon *et al* 1983; Robins-Browne *et al* 1994), chicks (Fukui *et al* 1995), dogs (Beaudry *et al* 1996), calves (Pospischil *et al* 1987) and cats (Pospischil *et al* 1987). These *E. coli* are responsible for considerable economic loss on farms throughout the world (Pillien *et al* 1996). In developing areas, domesticated animals are probably involved to an extent in the transmission of pathogenic *E. coli.*

Of the animal AEEC the most researched are probably those from rabbits. *E. coli* are a major cause of diarrhoea in commercial rabbit farms of Western Europe resulting in considerable economic burden (Blanco *et al* 1997). A number of serotypes are known to be pathogenic for rabbits including O15, O33 and O103. These are generally known as rabbit-specific EPEC (REPEC) but also, in the case of serotype O15:H-, RDEC-1 (Cantey and Blake 1997). At least two adhesins have been reported from rabbit EPEC. These are known as adhesive factor/rabbit 1 and 2 (AF/R1, AF/R2) (Wolf *et al* 1988; Fiederling *et al* 1997). A putative third adherence factor has also been described from an REPEC strain which is plasmid-mediated and shows homology to genes encoding the K88 and CS31A fimbriae of ETEC

(Adams *et al* 1997). Rabbit EPEC strains adhere to Peyer's patch epithelium and it has been suggested that, if human EPEC strains do likewise, then these strains may be useful as mucosal vaccine vectors (Von Moll and Cantey 1997).

1.7 Worldwide incidence of enteropathogenic *Escherichia coli* in humans

Between the 1960s and 1980s, a number of studies showed that the incidence of EPEC infection in developing countries ranged from 4% to 37% (Thoren 1983). It is unlikely that these figures have changed much since because sanitation has not significantly improved and populations have increased in such countries. It is now estimated that there are 117 million diarrhoeal episodes attributable to EPEC each year in developing countries alone (excluding China). In some areas, EPEC remains the most prevalent enteropathogen isolated from children under two years of age (Rosa *et al* 1998).

A number of studies (Mayatepek *et al* 1993; Hulan *et al* 1991) have attempted to determine the incidence of EPEC in children with and without diarrhoea. In one study (Mayatepek *et al* 1993), EPEC was isolated from 16% of children with diarrhoea and from 4% of those without. In another study (Hulan *et al* 1991), EPEC was isolated from 9% of children with diarrhoea, and in another (Lim *et al* 1992) EPEC was isolated from 2.7%. Therefore, the actual incidence of EPEC varies according to the setting and may depend on a number of factors such as the country, time of year and the quality of the study.

E. coli gastroinestinal infections, apart from those due to EHEC, are not required to be reported to public health centres in either the UK or the US, although outbreaks are reported. In view of the fact that EHEC has gained prominence as a major foodborne pathogen recently in both the UK and US, it would perhaps be useful to report all *E. coli* gastrointestinal infections to the Communicable Disease Surveillance Centre and Centers for Disease Control respectively. There are also no surveillance systems for *E. coli* infections, as well as many other infections, in developing countries where EPEC is now most common. There is therefore no

accurate method for calculating the incidence of EPEC infections in either developed or developing countries. Studies of the incidence of diarrhoeal pathogens in communities have been performed but these often omit EPEC. In such studies, up to 40% of infections are of unknown aetiology, a percentage to which EPEC may contribute. Those studies that include EPEC are still unrepresentative of the global situation because they are often performed in isolated communities (Mayatepek *et al* 1993).

In line with most other gastrointestinal pathogens, reservoirs of EPEC infection are many. They may include symptomatic or asymptomatic infants or children, and asymptomatic adult carriers or animals. However, the actual methods of transmission are poorly understood although, again like other gastrointestinal pathogens, they probably rely on the faecal-oral route of infection. Via this route, close contact between animals or people, sharing of food items and food itself can all act as sources of infection. A recent study, for example, screened 402 *E. coli* isolates from various food items for the presence of EPEC. Surprisingly, a total of 19 EPEC isolates were found with 17 of them from cooked foods (Norazaeh *et al* 1998). The simplest method of avoiding infection is therefore good hygiene but, as discussed, areas where EPEC remains a problem often do not have adequate hygiene provisions.

1.8 Clinical aspects of enteropathogenic *Escherichia coli* infection

The infective dose of EPEC required to cause disease in infants has not been established. Volunteer studies in adults indicate that a high infective dose, between $5x10^8$ and 10^{10} organisms, may be required (Levine *et al* 1978, Bieber *et al* 1998). However, the infective dose in adults cannot be used to reflect the infective dose required in infants because of differences in physiology and immune status. The incubation period for EPEC infection is not really known because infant studies have not been performed. In adult studies, one report states that the incubation period is 8 to 60 hours (Hart *et al* 1993) whilst another states that it varies from 2 to 12 days (Neter 1959). However, a recent volunteer study used 48 hour infection as

the study end-point (Bieber *et al* 1998). These figures should not be used to estimate the incubation period in infants because the adult intestine differs from the infant intestine.

Diarrhoeal disease due to EPEC is extremely age-related, usually occurring in infants less than six months of age. The reason for this age specificity is not understood; factors such as mucosal immunity and differences in gut structure between infants and adults (i.e. possession of receptors) could play a part. Lack of immunity appears to be the most probable reason due to the association of infection with bottle-fed infants. The presence of IgA in breast milk may be a protective factor although other factors in breast milk may also contribute. It would appear that contaminated feeding bottles do not account for all cases of EPEC infection, although they may play some part in transmission of the organism. It is also interesting to note that EPEC infection is not alone in its age specificity; EHEC O157:H7 is only pathogenic in calves less than 3 weeks old (Dean-Nystrom *et al* 1997). After this time the calves are not affected by EHEC infection although humans are affected at any age. A volunteer study in adults (Donnenberg *et al* 1998) showed that prior infection with EPEC can reduce disease after homologous rechallenge. This protection is thought to be due to serum IgG against the bacterial lipopolysaccharide. Therefore, prior infection with serotype O55 may not protect against future infection with serotype O127.

EPEC infection results in an acute or persistent watery, non-bloody or mucoid diarrhoea (Hart *et al* 1993) often accompanied by fever and vomiting. The disease ranges from a fulminating diarrhoea to a subclinical infection (Neter 1959), presumably according to host factors. After colonisation of the intestine with EPEC, bacteria can be isolated for about 4-7 days before the onset of symptoms (Taylor 1970). During the symptomatic stage EPEC is present in pure culture in the faeces (Taylor 1970; Moffet 1968). In most cases, if recovery occurs, then the organisms are also cleared; however, in some cases carriage may occur for some weeks (Taylor 1970).

1.9 Therapy of enteropathogenic *Escherichia coli* infection

Most infants with diarrhoea caused by EPEC recover uneventfully if water and electrolyte disturbances are corrected promptly (Lim *et al* 1992; Hill *et al* 1991). In addition, the introduction of a protein-hydrolysate, lactose-free formula in one study led to the prompt cessation of diarrhoea and nutritional recovery in two infants (Fagundes-Neto *et al* 1996). Antimicrobial therapy may also be of benefit to those in a life-threatening condition (Hill *et al* 1991). The antimicrobial susceptibility patterns of EPEC are variable as would be expected with the number of serotypes implicated in infantile diarrhoea (Senerwa *et al* 1991). In one study (Senerwa *et al* 1991) of a nosocomial outbreak in Kenya, 82% of EPEC strains belonged to two resistance patterns although there was no consistent relationship between the plasmid profile and the antimicrobial resistance pattern. Most strains are susceptible to cefotaxime, colistin and amikacin, and are resistant to ampicillin (Ruczkowska *et al* 1990; Olukoya *et al* 1988) but studies use different antibiotics and there is not much overlap between them, therefore making comparisons difficult.

1.10 Mechanisms of virulence of enteropathogenic *Escherichia coli*

Historically, EPEC were defined in terms of their negative characteristics, namely their lack of toxin production and their inability to demonstrate Shigella-like invasiveness in gut epithelium. In recent years, however, the virulence mechanisms of EPEC have become better understood due to the advent of molecular biology. As such, a number of reviews have been published throughout this decade describing the increase in our knowledge of EPEC, two of which have been published recently (Baldwin 1998; Nataro and Kaper 1998). Although there is currently no evidence to support the hypothesis that EPEC cause secretory diarrhoea (Embaye *et al* 1993) the subject remains slightly controversial. A number of studies have been performed to determine the possibility of toxin production by EPEC. In one study (Polotsky *et al* 1977), two strains of *E. coli* of serotype O26:H11 adhered to gut epithelium, penetrated the enterocyte cytoplasm, reached the lamina propria and caused epithelial hypersecretion. They were shown to produce an enterotoxin as determined

by the rabbit gut loop test performed with living cultures, supernatants and whole cell lysates. This serotype is currently classified as EPEC although other members of the same serogroup are classified as EHEC.

The elucidation of some of the virulence mechanisms of EPEC has led to a greater understanding not just of EPEC but also of other pathogenic bacteria. This is particularly true for other enteric bacteria, some of which were previously thought to be non-pathogenic. It has been found that homologous genes are present in various enteric pathogenic bacteria. For example, *Hafnia alvei* possesses a gene which hybridises with the *eae* gene of EPEC (Albert *et al* 1992). Similarly, *Citrobacter freundii*, the cause of murine colonic hyperplasia, possesses a chromosomal *eae* gene necessary for colonic colonisation (Schauer *et al* 1993). The *eae* gene will be discussed further in later sections. Furthermore, type III secretion systems are present in a number of pathogens including Salmonella and Yersinia.

1.10.1 Localised adhesion of enteropathogenic *Escherichia coli.*

Strains of *E. coli* produce pili which are involved in adherence, some of which may be involved in pathogenicity (Hacker *et al* 1992; Edwards and Puente 1998). EPEC adhere to epithelial cells *in vitro* with a localised adherence (LA) pattern which is dependent upon the presence of a ~92 kb (60 MDa) plasmid (Girón *et al* 1991), initially known as pMAR2 (Baldini *et al* 1983) but later termed the EPEC adherence factor (EAF) plasmid (Levine *et al* 1985). The EAF plasmid belongs to the IncFII group of plasmids and is phenotypically negative for alpha-haemolysin, colicin and aerobactin synthesis, nor does it possess any biochemical or antibiotic resistance markers (Nataro *et al* 1987). Some strains of EPEC do not possess the EAF plasmid and therefore EPEC may be divided into two classes, those that possess the EAF plasmid being known as class I EPEC and those that do not as class II EPEC (Nataro *et al* 1985). EAF plasmids from various EPEC strains show 50% to 90% homology (Nataro *et al* 1987). However, homology between the EPEC EAF plasmid and the rabbit AE *E. coli* strain RDEC-1 adherence plasmid is only 50% suggesting that there may be uncharacterised genes on either of these plasmids which may be involved in host specificity or virulence. Furthermore, although

RDEC-1 gives rise to the characteristic EPEC AE lesion, it is EAF probe negative and does not adhere to HEp-2 cells (Nataro *et al* 1987). It has been shown (Vuopio-Varkila *et al* 1991) that LA is an inducible phenotype which occurs more rapidly *in vitro* if EPEC cells are pre-incubated with HEp-2 cells. Those cells that were non-adherent after 60 minutes were transferred to uninfected HEp-2 cells; LA occurred within 15 minutes compared with 30-60 minutes for non-induced cells. Expression of the LA phenotype is coded for by genes on both the chromosome and the EAF plasmid. These genes are located on the chromosome and in the *bfp* gene cluster on the EAF plasmid; such genes are described in further detail in later sections. Results from a study on clinical EPEC strains with and without the EAF plasmid and *eae* genes (Gabastou *et al* 1995) showed that localised adherence only occurred in EAF and *eae* positive strains cultured on differentiated enterocyte-like HT-29 cells. LA did not occur in the same strains cultured on undifferentiated HT-29 cells. This suggests that LA may be exhibited only when host cells are differentiated and therefore an unknown host cell receptor, which is only expressed after differentiation, may be required for exhibition of the EPEC LA phenotype. Furthermore, EAF-cured EPEC cells are unable to adhere to HEp-2 cells showing that the plasmid is required for the expression of the LA phenotype (Baldini *et al* 1983). Also, EAF-positive EPEC form bacterial aggregates when cultured in defined media but not complex media, while EAF-negative EPEC do not form bacterial aggregates in defined media (Vuopio-Varkila *et al* 1991). However, the bacterial aggregation seen in EAF positive EPEC is not prevented by *D*-mannose and is therefore probably not due to the expression of type 1 pili. Expression of the LA phenotype is associated with the expression of outer membrane proteins. When EPEC cells are treated with rifampicin, a bacteriostatic compound which blocks transcription, and therefore protein synthesis, the LA phenotype is not expressed. EPEC cells treated with nalidixic acid, a bacteriostatic agent that does not affect protein synthesis, still express the LA phenotype.

The LA phenotype is partly or wholly dependent on the expression of bundle-forming pili (Bfp) and were first described by Girón and colleagues (Girón *et al* 1991). These Bfp exist as filaments 50 to 500 nm in width and 14 to 20 μm long.

The pili create bundles which intertwine with those of other bacteria to create a three-dimensional network. Bfp are partially or wholly responsible for LA by the recruitment of bacteria in the surrounding environment of the host cell (Girón *et al* 1993). The amino-terminal sequence of the major structural protein of Bfp shows that it is a type IV pilus (Girón *et al* 1991); these type of pili are expressed by other pathogenic bacteria including *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae*, *Moraxella bovis*, *Dichelobacter nodosus* and *Vibrio cholerae* (Stone *et al* 1996). They were initially classified by their polar location and their association with 'twitching motility'. They possess similar structural features, share a characteristic, short peptide leader sequence, and have a highly conserved N-terminus (Hultgren *et al* 1996, Girón *et al* 1997). A PCR method has been developed for the detection of a *bfp* gene of EPEC (Gunzberg *et al* 1995). The method did not amplify DNA from any other bacterial enteropathogens tested and was 100% specific for those EPEC which exhibited characteristic LA. Monoclonal antibodies have also been produced against Bfp (Girón *et al* 1995) and may be useful in diagnosis or for studying the interaction of Bfp with host cells. The genes coding for Bfp are present on the EAF plasmid and are contained within the *bfp* gene cluster. The first gene to be cloned was *bfpA* (Donnenberg *et al* 1992). This gene was shown to encode the major structural subunit of Bfp, called bundlin, and its expression is regulated at the transcriptional level (Puente *et al* 1996). Expression occurs during the exponential growth phase and is greatest at temperatures between 35 and 37°C and also in the presence of calcium. Ammonium ions significantly reduce *bfpA* mRNA and protein expression and therefore the LA phenotype. The *bfpB* gene, located two genes downstream of *bfpA*, is required for Bfp production even when BfpA is being produced. BfpB is found in the outer membrane and is required for full pili assembly after initial synthesis (Ramer *et al* 1996). A pre-pili peptidase gene was then cloned (Zhang *et al* 1994) which was termed *bfpP*, the product of which was shown to be homologous to pre-pili peptidases of other bacteria. Several open reading frames were also found upstream of *bfpP*. The presence of fourteen genes has recently been reported independently by two laboratories (Stone *et al* 1996; Sohel *et al* 1996). The organisation of the gene cluster is similar to that of the *tcp* operon of *Vibrio cholerae*. In both the *bfp* and *tcp* systems the clusters begin with

the major structural subunit gene and are followed by a region of dyad symmetry which is predicted to reduce expression of the downstream genes (Stone *et al* 1996). Such a reduction in expression is important for pili biogenesis. The proteins encoded by the genes within the *bfp* gene cluster include several with homology to proteins involved in the biogenesis of type IV pili in other bacteria (Stone *et al* 1996) although several of the proteins are unique to EPEC. Additionally, three open-reading frames have been reported and are designated *bfpT*, *bfpV* and *bfpW*. They are located 6 kb from the *bfp* operon and are required for the transcriptional activation of *bfpA* (Tobe *et al* 1996). However, *bfpTVW* are essentially the same as *perABC* (Kaper and Gómez-Duarte 1997) (see below). As mentioned previously, additional genes external to the *bfp* gene cluster are required for full expression of Bfp. One such gene is the chromosomal locus *dsb*A that encodes disulphide isomerase which catalyses disulphide bonds. Inactivation of the *dsbA* gene leads to the loss of LA (Zhang and Donnenberg 1996). Another gene located on the EAF plasmid is *bfpT* which activates *bfpA* transcription (Sohel *et al* 1996) and is located 6.7 kb downstream of the distal gene of the *bfp* gene cluster. It has also been suggested that *lspA*, a chromosomal gene that codes for lipoprotein signal peptidase, is also probably necessary for Bfp biogenesis (Sohel *et al* 1996).

Observations made with Bfp have led to the conclusion that the pili are responsible for the LA phenotype. However, it has recently been shown using a number of different EPEC strains that Bfp are not involved in the initial stages of bacterial adhesion (Hicks *et al* 1998). The study used various EPEC strains, including a wild-type strain and mutants in intimin and/or Bfp, to investigate the role of intimin and Bfp in adhesion to paediatric intestinal tissue. Intimin-positive, Bfp negative mutants adhered and caused attaching and effacing lesions whereas intimin negative, Bfp positive mutants did not adhere to the intestinal tissue. This is confirmed by further unpublished observations that Bfp are not essential for, but do promote, full EPEC adherence (Haigh R, personal communication). A second adhesin is therefore probably present and responsible for the initial interactions with the host cell (Haigh R, personal communication). Similar observations were made with the AF/R1 pili in the rabbit EPEC strain RDEC-1 some time ago (Wolf *et al*

1988). Such a hypothesis is supported by recent evidence from a study on serotypes of EHEC (Dytoc *et al* 1994). This study showed that serotype O157:H7 possesses an *eae* homologue but serotype O113:H21 does not thereby suggesting that another adherence mechanism is possessed by the latter serotype. Although Bfp may not be fully responsible for the LA phenotype, recent data have confirmed that they are required for full virulence of EPEC (Bieber *et al* 1998). The diarrhoeal response of human volunteers in a randomised double-blind study to wild-type EPEC, *bfpA*-negative EPEC and *bfpT*-negative EPEC was studied. Only administration of wild-type EPEC resulted in full diarrhoeal disease.

In addition to Bfp, two other surface structures have been characterised from EPEC (Girón *et al* 1993). These have been termed rod-like fimbriae and fibrillae and the authors suggested that these structures were involved in the interaction of EPEC with host cells. It has also been suggested by the same authors that the LA phenotype is a multifactoral process involving both bacterial and host cell factors. Other authors suggest that additional factors may also be required for adherence; the genes encoding such factors may be present on the EAF plasmid or the chromosome (Stone *et al* 1996). LA may also be due in part to bacterial surface hydrophobicity conferred by lipopolysaccharides (LPS) and outer membrane proteins (OMP), although these probably have a small role in overall adherence since other *E. coli*, including non-pathogenic strains, also exhibit surface hydrophobicity due to the possession of LPS and OMPs. Recent evidence suggests that the above hypotheses may contain some truth. A second adhesin has been found in strains of attaching and effacing *E. coli* that are pathogenic to rabbits, particularly those of serotype O103:K$^-$:H2 (Pillien *et al* 1996). This adhesin enables the bacteria to adhere to HeLa cells with a diffuse pattern. A major subunit of the adhesin can be purified from surface extracts of the bacteria and is protein of 32 kDa. As mentioned above, the presence of a second adhesin in strains of EPEC pathogenic to humans would therefore not come as a surprise. A second adhesin would also fully explain the events which occur during the initial interaction between bacteria and enterocytes.

Another genetic locus on the EAF plasmid in EPEC consists of four genes called *perA, B, C, D* (plasmid encoded regulator) (Gómez-Duarte and Kaper 1995). Genes *bfpTVW*, described by Tobe *et al* (1996) are essentially the same as *perABC*. The *per* locus has been shown to regulate *eaeA* expression (Gómez-Duarte and Kaper 1995). However, the *per* genes may be involved in the regulation of other virulence genes because presence of the EAF plasmid has been linked to the development of localised adherence and the development of AE lesions (Jerse and Kaper 1991; Gómez-Duarte and Kaper 1995).

1.10.2 The attaching and effacing lesion of enteropathogenic *Escherichia coli.*

Polotsky *et al* (1977) showed that EPEC produce a characteristic ultrastructural lesion whereby bacteria make intimate contact with the outer enterocyte membrane causing localised destruction of the intestinal brush border microvilli and distortion of the apical enterocyte membrane. This results in cytoskeletal rearrangement and the formation of an actin-rich cup-like indentation (pedestal) at the site of bacterial contact. This was subsequently termed the attaching and effacing (AE) lesion (Moon *et al* 1983). This lesion has been shown to bend and undulate when viewed by video microscopy (Sanger *et al* 1995) and individual attached bacteria can also move on the cell surface. The formation of AE lesions results in a reduction in the absorptive capacity of the intestinal mucosa which inevitably leads to a disruption of the electrolyte balance and subsequent diarrhoea. EPEC are also able to invade tissue culture cells *in vitro*, a phenomenon not thought to occur *in vivo* (Donnenberg *et al* 1989; Francis *et al* 1991).

The ability of EPEC to attach and efface is associated with a large chromosomal locus termed the locus of enterocyte effacement (LEE). The LEE is a pathogenicity island and is present in all bacterial enteropathogens that produce the AE lesion (McDaniel *et al* 1995) including enterohaemorrhagic *E. coli*, *Hafnia alvei* and *Citrobacter rodentium* (Nataro and Kaper 1998). The LEE has now been sequenced and contains 41 predicted open-reading frames arranged in at least five polycistronic operons (**figure 1.3**) (Elliot *et al* 1998b). The insertion site of the LEE varies amongst EPEC and EHEC strains according to their evolutionary lineage (Wieler *et*

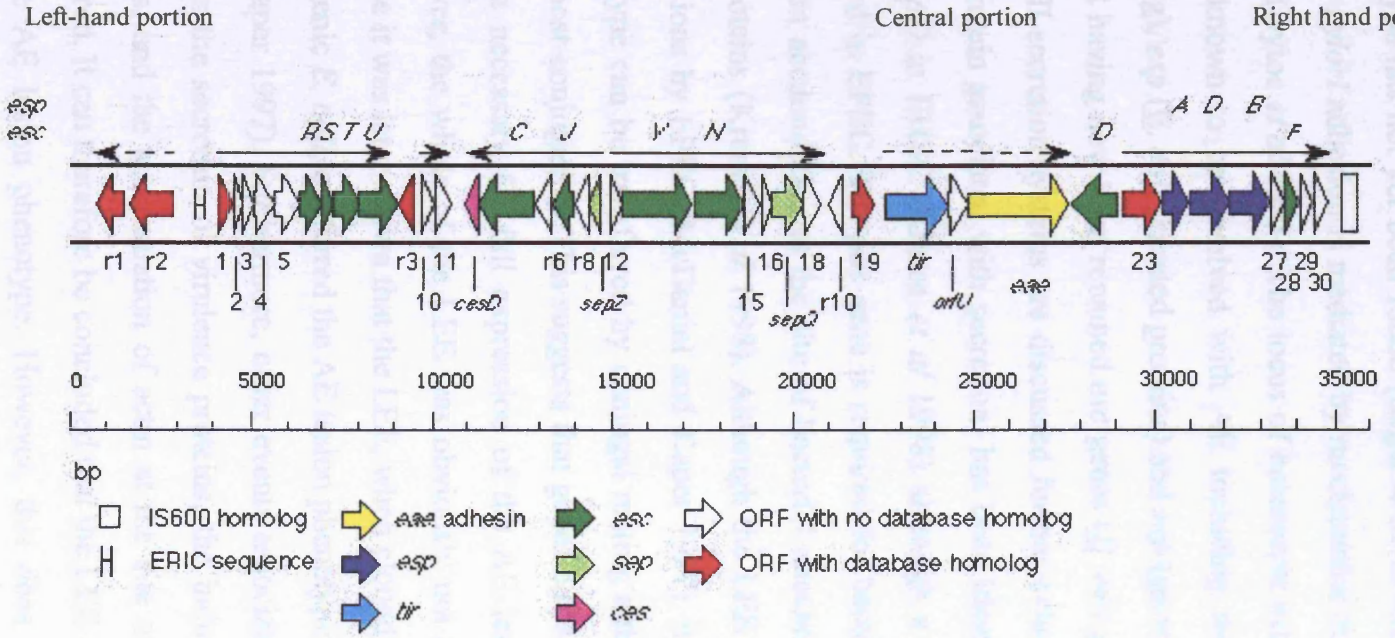


Figure 1.3 Schematic representation of the EPEC E2348/69 locus of enterocyte effacement (LEE)

Reproduced with kind permission Dr S. Elliot and Dr J. Kaper, University of Maryland
First published in Elliot *et al* (1998a).

al 1997). This suggests that the LEE inserted into the EPEC or EHEC chromosome a number of times and at multiple sites during their evolution. Although *Helicobacter pylori* induces cytoskeletal rearrangements and causes tyrosine phosphorylation of host cell proteins, a chromosomal locus associated with this phenotype has not yet been found (Segal *et al* 1996); data from one study suggest that *H. pylori* adhesion is mediated by mechanisms distinct from those produced by EPEC (Dytoc *et al* 1993). The locus of enterocyte effacement encodes many of the genes known to be involved with AE including the *eae (E. coli* attaching and effacing), *esp* (E. *coli* secreted proteins) and *sep* (secretion of EPEC proteins) genes, the last having now been renamed *esc* genes (*E. coli* secretion) (Elliot *et al* 1998b). Type III secretion systems are discussed further below. In addition, a gene named *pas* (protein associated with secretion) has been identified upstream of *espA*, *espB* and *espD* in EHEC (Kresse *et al* 1998) although a homologue has not yet been reported in EPEC. The *pas* gene is required for bacterial attachment and invasion, for actin accumulation at the site of bacterial attachment, and for the secretion of Esp proteins (Kresse *et al* 1998). Although the LEE is required for production of AE lesions by EPEC (McDaniel and Kaper 1997), it has been shown that the AE phenotype can be transferred by conjugal mating although its expression was only weak post-conjugation. This suggests that genes elsewhere on the chromosome may also be necessary for full expression of the AE lesion (O'Gorman *et al* 1996). However, the whole of the LEE was obviously not transferred in this experiment because it was later shown that the LEE, when cloned as a single fragment into non-pathogenic *E. coli*, conferred the AE lesion phenotype upon those *E. coli* (McDaniel and Kaper 1997). Furthermore, other events associated with AE lesion formation, such as the secretion of virulence proteins, the induction of host protein tyrosine kinases, and the accumulation of actin at the site of bacterial contact, were also conferred. It can therefore be concluded that the LEE contains all the genes required for the AE lesion phenotype. However, this does not mean that all the genes contained within the LEE are involved in virulence, although this is probably the case as it is a pathogenicity island (McDaniel and Kaper 1997).

The first gene identified within the LEE was *eaeA* (Jerse *et al* 1990), now known simply as *eae* due to recent nomenclature changes, and is located within the central region of the LEE (Elliot *et al* 1998b). The protein product of *eae*, termed intimin, shows significant homology with the invasin of *Yersinia pseudotuberculosis* and is present in all AEEC. However, homology of the gene varies between animal and human AEEC (Zhu *et al* 1996). Agin and Wolf (1997) suggested that there were three families of related intimins but it has recently been found that there are four different intimin types, designated α, β, δ and γ from different serotypes of AEEC (Adu-Bobie *et al* 1998). There are at least six secreted proteins, five of which are known to be encoded by genes within the LEE (**figure 1.3**). The first is a 25 kDa protein encoded by *esp*A (Kenny *et al* 1996); EspA is necessary for signal transduction, intimate adherence and formation of the attaching and effacing lesion. The second is a 37 kDa protein encoded by *espB*, previously known as *eaeB* (Donnenberg *et al* 1993; Foubister *et al* 1994a), which is located 5 kb downstream of *eae* and is also necessary for intimate attachment and the production of AE lesions (Abe *et al* 1998). It has recently been shown that EspB is translocated into the host cell after cell contact (Wolff *et al* 1998) and is targeted to the host cell cytoplasm (Taylor *et al* 1998). Furthermore, EspB may be a cytoskeletal toxin that causes actin redistribution (Taylor *et al* 1999) and its translocation is mediated by EspA-associated filamentous structures (Knutton *et al* 1998). These structures are only expressed during the early stages of AE lesion formation. EspA and EspB from human strains of EPEC are maximally expressed at the hosts' body temperature of 37°C (Abe *et al* 1997). The third protein has a molecular weight of 39 kDa and is encoded by *esp*D which lies between *espA* and *espB* in the LEE (Lai *et al* 1997). EspD is secreted via the type III secretion system (see below) and *espD* mutant is unable to induce the signal transduction events which activate the pathways leading to the attaching and effacing lesion. The genes *espA*, *espB* and *espD* are located in the right-hand portion of the LEE (Elliot *et al* 1998b). The fourth protein, known as Tir (for translocated intimin receptor), is 78 kDa and acts as the intimin receptor in host cells (Kenny *et al* 1997). It is located in the central portion of the LEE (Elliot *et al* 1998b). This protein has also subsequently been termed EspE (Deibel *et al* 1998) but Tir is the accepted name. Tir is discussed further in **section 1.10.3**. The fifth

protein is 110 kDa, is IgA protease-like and is encoded by the *esp*C gene which is not located within the LEE (Stein 1996). EspC is not necessary for mediating signal transduction in epithelial cells and does not play a role in adherence or invasion of tissue culture cells *in vitro*, and its role has therefore yet to be deduced. The sixth protein, EspF, is encoded by the *espF* gene which is located at the extreme right-hand end of the LEE (McNamara and Donnenberg 1998). EspF is approximately 21 kDa and requires the type III secretion apparatus for its secretion. However, EspF is not required for production of AE lesion formation (McNamara and Donnenberg 1998). In EHEC a gene encoding a predicted protein of 156 amino acids is located immediately upstream of *eae* (Zhao *et al* 1995) and a similar open reading frame is present in EPEC; this gene may be associated with pathogenicity in these organisms. It has recently been found that CesD (chaperone for *E. coli* secreted protein D) is a chaperone required for proper EspB and EspD secretion and shows sequence homology to other chaperone proteins from type III secretion pathways (Wainwright and Kaper 1998).

EPEC also possess a putative type III secretion system that is involved with the secreted proteins already described and is necessary for the secretion of virulence factors (Jarvis *et al* 1995). Four well-characterised secretion pathways have been described from Gram negative bacteria and are named type I to type IV (Finlay and Falkow 1997, Hueck 1998). A potential type V system has been described but is not well characterised and will not be described further here (Hueck 1998). Secretion pathway types I to IV are involved with the export of bacterial virulence factors. Type I secretion systems are *sec*-independent and are involved in the transport of signal sequence-lacking molecules (Finlay and Falkow 1997). Type II and type IV secretion systems are *sec*-dependent and require amino-terminal processing of the secreted protein. These systems are involved in the transport of many different molecules including some virulence factors (Russel 1998). Type III secretion systems are *sec*-independent although assembly of the actual system may require the *sec* pathway. They are involved with protein secretion in a number of bacterial pathogens of both animals and plants (Lee 1997, Hueck 1998). In EPEC there are ten genes located throughout the LEE which are associated with the system (**table**

1.3) (Elliot *et al* 1998b). These were recently renamed from the Sep proteins to Esc proteins (Elliot *et al* 1998b) in line with the nomenclature used for type III secretion systems in other bacterial pathogens such as *Pseudomonas* (Hueck 1998), *Salmonella* (Collazo and Galán 1997), *Shigella* (Bahrani *et al* 1997) and *Yersinia* (Cornelis and Wolf-Watz 1997; Fällman *et al* 1997; Cornelis 1998).

Table 1.3 Revised nomenclature for genes encoding the EPEC type III secretion apparatus

Old name	New name		Old name	New name
sepA	*escV*		*sepG*	*escT*
sepB	*escN*		*sepH*	*escS*
sepC	*escC*		*sepI*	*escR*
sepD	*escJ*		-	*escD*
sepF	*escU*		-	*escF*

The ability of EPEC to induce AE lesions is also dependent on a number of physiological and environmental conditions (Rosenshine *et al* 1996a). It is dependent upon the bacteria being in the early to mid-logarithmic phase of growth and also being at 37°C; AE lesions are not induced at 28°C. A number of studies (Rothbaum *et al* 1982; Clausen *et al* 1982; Taylor *et al* 1986) have described the typical AE lesion and pedestal formation in infants. The AE lesion is different from the lesions caused by enterotoxigenic *E. coli* (ETEC) or enteroinvasive *E. coli* (EIEC). EHEC infection results in the formation of similar lesions at the point of bacterial contact but are limited to the terminal ileum or colon (Hart *et al* 1993). EAggEC and DAEC, although able to adhere to HEp-2 (human laryngeal epidermoid carcinoma cells) and HEL (human embryonic lung) cells, do not produce AE lesions.

The attaching and effacing lesion produced by EPEC forms the basis of the fluorescent actin staining (FAS) test (Knutton *et al* 1989). The test relies on the accretion of actin at the site of intimate adherence and utilises the fungal toxin phalloidin for staining the actin. The FAS test can be used for visualising all AE lesion-producing bacteria such as EPEC, EHEC, *Hafnia alvei* and *Citrobacter rodentium*. However, recent evidence suggests that some bacteria may be FAS test negative although they possess the LEE (Wieler *et al* 1998), but this is more to do with the lack of adherence to certain cell lines than to a lack of genes required for actin accretion. As EAggEC and DAEC are unable to produce AE lesions they are FAS test negative.

1.10.3 Protein phosphorylation and signal transduction events during EPEC infection.

The protein phosphorylation and signal transduction events which occur during EPEC infection are complex and may involve proteins in both the infecting bacteria and the host cell. After attachment, intimate adherence and formation of the AE lesion, a number of changes are brought about in the host cell by EPEC. One of these is host cell protein tyrosine phosphorylation which has been reported as a potential virulence mechanism in a number of pathogenic bacteria (Foubister *et al* 1994a; Dehio *et al* 1995). In both EPEC (Foubister *et al* 1994a) and *Shigella flexneri* (Dehio *et al* 1995) it appears that such activity leads to cytoskeletal rearrangement. In EPEC it has recently been shown that a sixth secreted protein is involved in host cell signalling events (Kenny *et al* 1997). This protein of 78 kDa is inserted into the host cell via the type III secretion system and is dependent upon the presence of EspA and EspB to facilitate this delivery. An open reading frame immediately upstream of the *tir* gene and previously known as *orfU* is now known to be a chaperone for Tir and has been renamed *cesT* (Elliot S and Kaper J, personal communication). Upon insertion, the 78 kDa protein is tyrosine phosphorylated to form a protein with an apparent weight of 90 kDa and then acts as the receptor for intimin. As such the protein has been named Tir (translocated intimin receptor) (Kenny *et al* 1997). A subsequent study confirmed these observations although the authors named the protein EspE (Deibel *et al* 1998). Tir was previously known as

Hp90 (Rosenshine *et al* 1992, Foubister *et al* 1994b; Rosenshine *et al* 1996b) and was originally thought to be a mammalian cell protein.

It has also been shown that non-tyrosine protein phosphorylation occurs upon attachment of EPEC cells (Manjarrez-Hernandez *et al* 1992; Baldwin *et al* 1990). The major phosphoprotein is myosin light chain (20 kDa), a cytoskeletal component which, when phosphorylated, regulates cell actin organisation in non-muscle cells (Manjarrez-Hernandez *et al* 1992). This was identified from infected small intestine mucosal biopsies and Caco-2 cells (Manjarrez-Hernandez *et al* 1991). It has been proposed (Manjarrez-Hernandez *et al* 1992) that extensive phosphorylation of myosin light chain in host cells leads to irreversible destruction of microvillus function and the accretion of actin at the site of bacterial attachment. However, this alone probably does not account for the diarrhoea associated with EPEC infection and therefore other mechanisms must be involved. It has also been shown that any increase in phosphorylation of myosin light chain by EPEC results in perturbation of intestinal epithelial barrier function (Yuhan *et al* 1997). This is probably an additional contributory factor for the diarrhoea seen in EPEC infection. More recently, it has been shown *in vitro* that there is a transient increase in short circuit current after AE lesion formation by EPEC on Caco-2 cells (Collington *et al* 1998a) which is mediated by EspA, EspB and EspD (Collington *et al* 1998b). This is possibly linked to an influx of sodium and amino acids across the apical membrane. In addition to the above-mentioned proteins there are also two other phosphoproteins of 21 kDa and 29 kDa which have been identified from infected HEp-2 cells although their role remains unknown (Baldwin *et al* 1990).

Host protein phosphorylation only occurs upon infection with *E. coli* that are able to cause the AE lesion resulting in actin accretion (Baldwin *et al* 1990) and does not occur with those *E. coli* that simply adhere. Although it has been shown that host cell phosphorylation leads to AE lesion formation, intimate adherence is not required in order to activate the signal transduction pathways (Haigh *et al* 1995). This suggests that *eae* is not required to produce this phenomenon. Instead, *espB* may have a role in the activation of host cell signal transduction pathways (Haigh *et*

al 1995). This is supported by evidence that EPEC *espB* mutants are unable to activate epithelial cell signals including tyrosine phosphorylation and cytoskeletal rearrangements (Foubister *et al* 1994a).

EPEC infection results in an increase in the level of inositol phosphates (IP) within HeLa (human cervical adenocarcinoma cells) cells (Foubister *et al* 1994b). IP's are responsible for releasing calcium from calcium-sequestering compartments. The flux was proposed to occur as a result of host cell protein tyrosine phosphorylation but was not a result of actin rearrangement within the host cells. However, such protein tyrosine phosphorylation is probably of the EPEC Tir protein recently described by Kenny *et al* (1997) and not of host cell proteins. The increase in IP level takes approximately two hours whereas the normal host response to hormones inducing IP formation takes seconds or minutes. Those attached EPEC that have not caused cytoskeletal rearrangement still cause an increase in IP levels within HeLa cells. EPEC also causes a decrease in transepithelial electrical resistance in Caco-2 and MDCK cells (Canil *et al* 1993). The decrease is due to the disruption of an intracellular pathway. This phenomenon may be involved in the diarrhoea seen during EPEC infection.

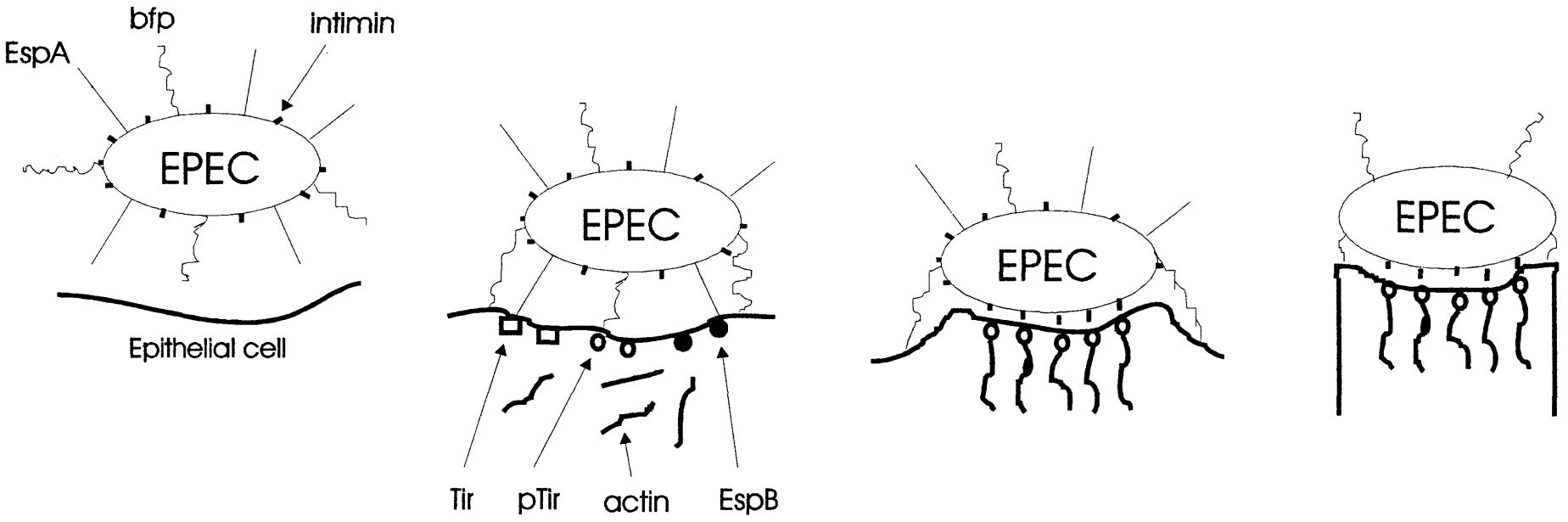
Early studies of the signal transduction events during EPEC pathogenesis also demonstrated increased levels of intracellular calcium which, along with the increases in IP, gave rise to a plausible model for microvilli effacement and subsequent AE lesion formation (Baldwin *et al* 1991). However, recent data have shown that increased levels of intracellular calcium cannot be detected at the site of the AE lesion, or in EPEC or EHEC-infected cells compared to non-infected cells, and are therefore not required for AE lesion formation by EPEC or EHEC (Bain *et al* 1998). Conflicting data are thought to be due to the methods employed in each study, the latter of which employed a ratiometric technique whereby the risk of calcium compartmentalisation is eliminated (Bain *et al* 1998).

1.10.4 Four-stage model of EPEC pathogenesis.

The data acquired from studies on the pathogenesis of EPEC infection led to the description of a three-stage model of EPEC pathogenesis (Donnenberg and Kaper 1992; Nataro and Kaper 1998). However, data gained over the last few years has led to the description of a four-stage model of pathogenesis which updates and provides a more detailed model than the three-stage model (**figure 1.4**) (Knutton *et al* 1998). In the first stage and in the presence of the correct environmental conditions, EPEC express Bfp and intimin and produce EspA filaments. The expression of these products depends on genes located on the EAF plasmid and in the LEE. In stage two, EPEC adhere to the epithelial cell via Bfp and EspA filaments resulting in the stimulation of EspB and the translocation of Tir, and probably other as yet unidentified secreted proteins, into the host cell. This leads to the activation of tyrosine protein kinase (TPK), formation of the intimin receptor, Tir, which is tyrosine phosphorylated and actin rearrangements within the host cell. In stage three, EspA and other surface structures are removed from the region of bacterial adherence, intimin binds to phosphorylated Tir and actin accumulation occurs. During stage four, further actin accumulation leads to the AE lesion.

1.11 Plasmids of enteropathogenic *Escherichia coli*

EPEC possess a number of plasmids depending on the actual strain; for example, strains of EPEC may contain seven or more plasmids (Scaletsky *et al* 1995). The relevance of many of these plasmids is unknown. EPEC strain E2348/69 possesses two plasmids; a large EAF (EPEC adherence factor) plasmid called pMAR2 and a small 5.5 kb cryptic plasmid. The EAF plasmid, as already described, is associated with the presence of the *bfp* gene cluster which codes for bundle-forming pili. The role of the cryptic plasmid is not known. However, it may have a role in adherence because strains of E2348/69 which are cured of the EAF plasmid are still able to adhere to cultured epithelial cells while if both the EAF and cryptic plasmids are removed no adherence occurs (Haigh R, personal communication). Previous work in this laboratory has shown that two regions of the cryptic plasmid have 96% and 88% sequence similarity respectively with regions of the plasmid pWQ799 of



Stage One **Stage Two** **Stage Three** **Stage Four**

Figure 1.4 Four-stage model of EPEC pathogenesis. See text for full description of each stage.

Adapted from Knutton *et al* 1998

Salmonella enterica serovar *borreze*. The homologous regions in the *Salmonella* plasmid are mobilisation and partitioning regions. The role of a small plasmid in EPEC strain 0041 serotype 0111:H$^-$ has also recently been described (Scaletsky *et al* 1995). It was shown that genes on a 6.6 kb plasmid encoded epithelial cell invasion and kanamycin resistance; cell invasion also appeared to be dependent on the synthesis of a 32 kDa protein. Furthermore, a recent paper (Scaletsky *et al* 1996) described a clinical isolate, EPEC serotype 018ab, which may cause diarrhoea by an, as yet, undescribed virulence mechanism. This isolate, confirmed to be enteropathogenic *E. coli* and possessing the *eae* gene, was able to invade HeLa cells in a gentamicin invasion assay, and also invaded rabbit intestinal epithelium. However, the isolate did not contain the EAF plasmid and would therefore not possess bundle-forming pili. Bacterial adhesion assays confirmed the absence of Bfp because the bacteria adhered to HeLa cells in an indefinite pattern. The fact the bacteria are still able to adhere, although they lack Bfp, suggests that other genes may be present on the chromosome or a cryptic plasmid, although the presence of such plasmids was not described. Therefore the various small plasmids to be found in strains of EPEC may have important functions involved in virulence.

1.12 Diagnosis and identification of enteropathogenic *Escherichia coli*

Diagnosis and identification of EPEC may be achieved by routine biochemical and serotyping methods. However, these tests are not 100% sensitive or specific and other methods are therefore required. A test which may be employed to identify EPEC from intestinal biopsies or cultured cell lines is the fluorescent actin staining test (FAS test) (Knutton *et al* 1989) as described in **section 1.10.2**. The test is highly sensitive and specific for EPEC, EHEC, *Hafnia alvei* and *Citrobacter rodentium*.

There are probes available for the main virulence properties of EPEC (Nataro and Kaper 1998). Use of these in one study (Giammanco *et al* 1996) showed that 80% of classical and 25% of non-classical EPEC serotypes possessed the *eae* gene. However, 60% or fewer of classical EPEC serotypes possessed the gene for localised adherence, the EAF plasmid or the ability to accrete actin at sites of

adherence as determined by the FAS test. In non-classical EPEC serotypes, the incidence of such probe-positive serotypes was 10% or lower. FAS test positive, EAF negative strains of EPEC have been isolated (Knutton *et al* 1991) and these have been of classical and non-classical serotypes, but their role in infantile diarrhoea remains to be assessed. Similarly, EAF-cured EPEC are positive in the FAS test (Knutton *et al* 1989). Although Shigella appear to cause an increase in actin concentrations beneath plasma membranes during invasion, EIEC, which are similarly invasive, are not strongly positive in the FAS test (Knutton *et al* 1989).

1.13 Tyrosine phosphorylation in enteropathogenic *Escherichia coli*

One of the covalent modifications proteins undergo after translation is phosphorylation and this is now recognised as the most important method of acute protein regulation in eukaryotic cells. Protein phosphorylation is responsible for controlling cellular activity by being the major mechanism by which cells respond to extracellular signals and by being responsible for the timing of cell cycle events (Hardie 1993). Protein phosphorylation occurs in eukaryotes principally on serine, threonine and tyrosine residues but, in prokaryotes, it was thought for many years that protein phosphorylation occurred only on serine and threonine residues (Cozzone 1988). Evidence of protein phosphorylation on tyrosine residues was reported in *E. coli* by Cortay and colleagues (1986a, 1986b) although this was later shown to be adenylation not phosphorylation (Foster *et al* 1989). Further legitimate reports of tyrosine phosphorylation in prokaryotes have been made over the past decade in *Clostridium thermohydrosulfuricum* (Londesborough 1986), *Bacillus subtilis* (Mitchell *et al* 1992), *Acinetobacter calcoaceticus* (Dadssi *et al* 1990), *Pseudomonas solanacearum* (Atkinson *et al* 1992), *Streptococcus pyogenes* (Chiang *et al* 1989), *Salmonella typhimurium* (Ostrovsky and Maloy 1995), *Pseudomonas aeruginosa* (Kelly-Wintenberg *et al* 1990), *Pseudomonas syringae* (Ray *et al* 1994) and enteropathogenic *Escherichia coli* (Freestone *et al* 1995; Freestone *et al* 1998b). Autophosphorylation at tyrosine has also been reported in *Acinetobacter johnsonii* (Duclos *et al* 1996); an 81 kDa protein which is located in the inner membrane is autophosphorylated and its kinase exists in at least five isoforms (Duclos *et al* 1996,

Grangeasse *et al* 1997). Cloning of the kinase gene revealed an open reading frame of 2199 nucleotides encoding a 82 kDa protein (Grangeasse *et al* 1997). The authors suggest a possible involvement in cell recognition and bacterial pathogenicity.

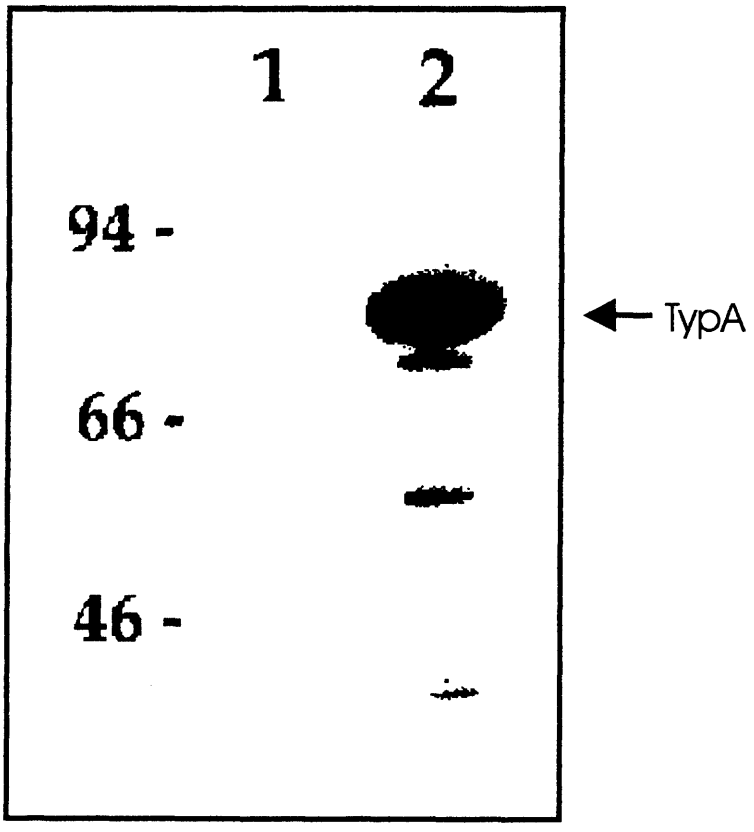
More recently, three phosphotyrosine proteins of 27, 36 and 52 kDa were identified in the cyanobacterium *Anabaena* sp. strain PCC7120 (McCartney *et al* 1997). A number of bacteria in the *Archaea* have also been shown to possess proteins which are phosphorylated on tyrosine (Smith *et al* 1997). The bacteria in this group so far shown to possess the tyrosine phosphorylated proteins are *Sulfolobus sulfataricus* ATCC 35091, *Haloferaz volcanii* and *Methanosarcina thermophila* TM-1. However, no clear function has yet been described for the phosphotyrosine proteins in any of the above organisms.

Specifically in *E. coli*, the presence of a number of phosphoproteins in enteropathogenic *E. coli* derivative MAR001 has been demonstrated (Freestone *et al* 1995). Two tyrosine phosphorylated proteins of 80 and 84 kDa have been detected. The N-terminal sequence of the 80 kDa protein was MIEKLYNIAIIAHVD. The sixth residue is thought to be a modified tyrosine residue and 14 of the 15 residues are identical to the predicted N-terminal amino acid sequence of an open reading frame called *o591* (Plunkett *et al* 1993) at 88 minutes on the *E. coli* chromosome. ORF *o591* has a predicted product of 65.5 kDa; the difference between its predicted size and the size seen on a protein gel may be due to its structure. ORF *o591* is approximately 1.77 kb in length and has been sequenced as part of the *E. coli* genome project. ORF *o591* is present in both *E. coli* K-12 and EPEC strains and its product is found in extracts of both EPEC and *E. coli* K-12 laboratory strains (Freestone *et al* 1998b). However, the product of ORF *o591* is tyrosine phosphorylated in EPEC (Freestone *et al* 1995; Freestone *et al* 1998b) and in an *E. coli* L-form (Freestone *et al* 1998a) but not in *E. coli* K-12 (Freestone *et al* 1995) (**figure 1.5**). Genuine tyrosine phosphorylation of this protein was shown by its elimination when treated with protein-tyrosine phosphatase (Freestone *et al* 1998b). Furthermore, *in vivo* tyrosine phosphorylation was shown by labelling *E. coli* MAR001 with [^{32}P]-orthophosphate followed by analysis of cytoplasmic extracts

Figure 1.5 *In vivo* phosphorylation of TypA in total cell extracts of *E. coli* probed with anti-phosphotyrosine antibodies. Note that TypA is tyrosine phosphorylated only in EPEC MAR001 as indicated by the arrow.

Lane 1: *E. coli* K-12 strain 3301

Lane 2: EPEC MAR001.



Reproduced with kind permission of Dr P. Freestone, University of Leicester. First published in Freestone *et al* 1995.

(Freestone *et al* 1998b). The protein also appears to be phosphorylated in its purified form but it is not known whether this is due to autophosphorylation or whether its associated kinase remains in a bound form with the purified protein. Although kinase activity of the product of ORF *o591* has been investigated, no independent kinase has been found (Freestone *et al* 1998b). These results suggest that the 80 kDa protein is not autophosphorylated. It is known that the protein co-purifies with RNA polymerase and therefore the product of ORF *o591* may have a role in transcription (Freestone *et al* 1995), although it may be more likely that it has a role in translation because the DNA sequence of *o591* has 40% sequence identity with the gene encoding elongation factor G. Due to its known characteristics, ORF *o591* was recently named *typA* for tyrosine phosphoprotein A (Freestone *et al* 1998b) and from this point will be described as such.

The fact that protein TypA is tyrosine phosphorylated in EPEC but not in *E. coli* K-12 strains may mean that *typA* has a role in pathogenicity, although tyrosine phosphorylation of TypA has also been observed in an L-form of *E. coli* K-12 (Freestone *et al* 1998a). A role for TypA in pathogenicity is supported by the fact that BipA, the membrane-associated homologue of the product of TypA in *Salmonella typhimurium*, is induced more than seven-fold by the human neutrophil protein BIP (bactericidal/permeability-increasing protein) (Qi *et al* 1995). The protein has homology with members of the GTPase superfamily such as elongation factor G (Freestone *et al* 1998b; Farris *et al* 1998) suggesting an interaction of *typA* with ribosomes (Farris *et al* 1998) and perhaps, therefore, a role in protein synthesis. In addition, recent work by Farris and colleagues suggests that *typA* may have a role in the attaching and effacing phenotype of EPEC (Farris *et al* 1998). However, a wild-type clinical strain was not used for these experiments and therefore these data should be interpreted with caution. A role for TypA in the virulence of EPEC has therefore yet to be proven beyond hypothesis.

1.14 Aims of project

The aim of this project was to investigate the role of *typA* in the virulence of EPEC. Limited data suggest a role of the *typA* homologue, *bipA*, in the virulence of *Salmonella typhimurium*. However, there exist only the data of Farris and colleagues (1998) to support such a role in EPEC. As these data used the EPEC derivative MAR001, which does not possess the EAF plasmid, further work is required to investigate the role of *typA* in virulence. To fulfil this objective the initial aim was to clone and sequence the *typA* gene from EPEC strain E2348/69. After sequencing, the main objective was to create a site-specific deletion in the *typA* gene resulting in a frameshift mutation so that a mutant strain could be constructed using a suicide vector system. Standard phenotypic studies described for the analysis of EPEC virulence, including the FAS test and cell culture assays, could then be peformed on the mutant strain to determine the effect of the *typA* mutation on EPEC virulence. Analysis of the interaction of *typA* with other EPEC virulence genes could also be performed as a *typA* mutation in *E. coli* K-12 has been shown to result in the alteration of expression of certain proteins (Freestone *et al* 1995). Such interactions may be important in gene regulation.

Chapter 2

Materials and Methods

2.1 Bacterial strains

Designation	Relevant characteristics	Source
EPEC E2348/69	O127:H6	Baldini *et al* 1983
EPEC E2348/69 Nal^r	O127:H6, Nal^r	Dr J. Kaper, Univ. of Maryland
EPEC MAR001	O127:H6, Km^r, EAF negative	Knutton *et al* 1989
EPEC JPN15	O127:H6, Nal^r, EAF negative	Dr J. Kaper, Univ. of Maryland
EPEC 6-8-1	O127:H6, pMAR2 *bfpA*::Tn*phoA*	Donnenberg *et al* 1990
EPEC RDEC-1	O15:H-, Nal^r	Dr P. Everest, Imperial College
DH5α	*sup*E44, Δ*lac*, *hsd*R17, *rec*A1, *end*A1, *gyr*A96, *thi*-1 *rel*A1	Hanahan 1983
CC118λ*pir*	*araD139*, Δ(*ara leu*)7697, Δ*lacX74*, Δ*phoa20*, *galE*, *galK*, *thi*, *rpsE*, *rpoB*, *argE*(Am), *recA1*, Rif^r	Manoil and Beckwith 1985
SM10λ*pir*	*thi*1, *thr*1, *leuB6*, *supE*44, *tonA*21, *lacY1*, *recA*::RP4-2-Tc::Mu, Km^r	Simon *et al* 1983
SCC1001	E2348/69, Str^r	Chapter 3
SCC1002	E2348/69 Str^r, *typA*	Chapter 4
SCC1003	E2348/69 Nal^r, *typA*	Chapter 4
SCC1004	MAR001, *typA*	Chapter 4
SCC1005	JPN15, *typA*	Chapter 4
SCC1006	RDEC-1, *typA*	Chapter 4

2.2 Routine culture of bacterial strains

Bacterial strains were inoculated into the required volume of LB supplemented with the required antibiotics at appropriate concentrations and incubated with shaking at

37°C overnight. Alternatively, bacterial strains were plated onto LBA plates and incubated at 37°C overnight. Any modification in culture for specific experiments is described in the appropriate sections.

2.3 Storage of bacterial strains

The method for storage of bacterial strains was modified from Sambrook *et al* 1989. The required bacterial strain was grown overnight at 37°C in LB broth with the addition of the appropriate antibiotics. After growth, 0.5 ml of the culture was added to 0.5 ml of 50% glycerol in a 1.5 ml Eppendorf tube to give a final concentration of 25% glycerol. The tube was labelled and frozen at -70°C. Bacterial strains were recovered by taking a sterile loop-full of the frozen suspension, plating it onto LBA and incubating at 37°C overnight.

2.4 Plasmid vectors

Plasmid	Characteristics	Source
pBluescript	ColE1, *lacI*, *lacZ*, Amp^r	Stratagene
pUC18	ColE1, *lacI*, *lacZ*, Amp^r	Yannisch-Peron *et al* 1985
pUC19	ColE1, *lacI*, *lacZ*, Amp^r	Gibco BRL
pKK223-3	*tac* expression vector, Amp^r	Pharmacia
pAV3	*lacZ*, Chl^r, pBluescript MCS	Van Vliet A, personal comm.
pRDH10	*sacRB*, *cat*, Tet^r, oriR6K, *mobRP4*	Haigh R, personal comm.
pSCC-T	1.8kb *typA* *Dra*I-*Bam*HI fragment cloned into *Sma*I-*Bam*HI sites of pUC19	Chapter 3
pSCC-T2	1.8kb *typA* *Eco*RI-*Hind*III fragment cloned into pKK223-3	Chapter 5
pSCC01	6.5 kb *Not*I-*Hind*III *typA* region from EPEC E2348/69 cloned into pBluescript	Chapter 3
pSCC04	6.3 kb *Sal*I *typA* region from EPEC E2348/69 cloned into pUC18	Chapter 4
pSCC05	1.4 kb *Pst*I fragment downstream of *typA* cloned into pUC18	Chapter 3
pSCC06	500bp *Pst*I-*Hind*III fragment downstream of *typA* cloned into pUC18	Chapter 3
pSCC12	pSCC04 with four base-pair *typA* deletion	Chapter 4
pSCC16	*Sal*I fragment from pSCC12 cloned into pRDH10	Chapter 4

p6-8-1	pMAR2 *bfpA*::Kan^r	Donnenberg *et al* 1990
pRDH7	pBluescript, 5.4 kb *Bam*HI fragment from EAF plasmid containing *bfp* region	Haigh R, personal comm.
pMAR7	Ampicillin-tagged EAF plasmid	Baldini *et al* 1983
pCVD450	3.5 kb *Eco*RI fragment containing *per* genes from pMAR2 cloned into pACYC184 Tet^r	Gómez-Duarte and Kaper 1995

2.5 Handling of hazardous chemicals

All chemicals were handled with care and the necessary precautions taken according to health and safety guidelines such as COSHH. Appropriate safety wear and extractor hoods were used when necessary.

2.6 Sterilisation of buffers, media, solutions and equipment

All buffers, media, solutions and equipment requiring sterilisation before use were autoclaved at 121°C, 15psi for 15 minutes.

2.7 Media

Luria-Bertani broth. Luria-Bertani (LB) broth was prepared by the addition of 2 g NaCl, 4 g tryptone and 2 g bactoyeast to 400 ml distilled water. The broth was then autoclaved.

Luria-Bertani agar. LBA was prepared by the addition of 2 g NaCl, 4 g tryptone, 2 g bactoyeast and 6 g agar to 400 ml distilled water. The agar was then autoclaved.

Sucrose Luria-Bertani broth. Sucrose LB broth was prepared by the addition of 24 g sucrose, 4 g tryptone and 2 g bactoyeast to 400 ml distilled water. The solution was then autoclaved.

Sucrose Luria-Bertani agar. Sucrose LBA was prepared by the addition of 24 g sucrose, 4 g tryptone, 2 g bactoyeast and 6 g agar to 400 ml distilled water. The agar was then autoclaved.

SOC medium. SOC medium was prepared by adding 20 g of bacto-tryptone, 5 g of bacto-yeast, 0.5 g of NaCl and 10 ml of 250mM KCl to 950 ml of deionised water. The solution was autoclaved at 121°C for 15 minutes. After the solution had cooled, 20 ml of sterile 1M glucose solution and 5 mls of 2M $MgCl_2$ were added.

Dulbecco's Modified Eagle Medium (DMEM). DMEM without sodium pyruvate but with 4500 mg/l glucose with pyridoxine and phenol red was obtained from Gibco BRL.

2.8 Media supplements

Supplement	Solvent	Storage	Stock conc.	Working conc.
Ampicillin	dH_2O	4°C	100 mg/ml	100µg/ml
Chloramphenicol	Ethanol	4°C	50 mg/ml	50µg/ml
Kanamycin	dH_2O	4°C	25 mg/ml	25µg/ml
Nalidixic acid	0.1M NaOH	4°C	50 mg/ml	50µg/ml
Rifampicin	50:50 0.1M K_2CO_3:EtOH	4°C	25 mg/ml	25µg/ml
Streptomycin	dH_2O	4°C	50 mg/ml	50µg/ml
Tetracycline	Ethanol	4°C	20 mg/ml	20µg/ml
X-gal	DMF	-20°C	50 mg/ml	40µg/ml
X-P	dH_2O	-20°C	20 mg/ml	20µg/ml

2.9 Miscellaneous solutions

20x SSC. 175.3 g of NaCl and 88.2 g of sodium citrate were added to 800 ml of distilled water. The pH was adjusted to 7.0 using 10N NaOH and the final volume adjusted to 1000 ml with distilled water. The solution was then autoclaved.

20% and 50% glycerol. 20% glycerol was made by the addition of 20 ml of glycerol to 80 ml of distilled water. The solution was then mixed thoroughly and autoclaved. 50% glycerol was made by the addition of 50 ml glycerol to 50 ml of distilled water.

50x ELFO buffer. 242 g of Tris and 18.61 g EDTA were added to 900 ml of distilled water. The pH was adjusted to 7.7 with approximately 50 ml of glacial acetic acid and then the final volume made to 1000 ml with distilled water.

Phenol chloroform. Equal amounts of phenol and DNA grade chloroform were mixed and equilibrated by extracting with 0.1M Tris HCl pH7.6. The equilibrated mixture was stored under a layer of 0.01M Tris HCl pH 7.6 at 4°C.

Chloroform: isoamyl alcohol. DNA-grade chloroform was mixed in the ratio of 24:1 with isoamyl alcohol and stored at 4°C.

3M sodium acetate. 408.1 g of sodium acetate·$3H_2O$ were added to 800 ml of distilled water and mixed well. The pH was adjusted to 5.2 using glacial acetic acid and the volume made to 1000 ml with distilled water. The solution was then autoclaved.

5M sodium chloride. 116.9 g of sodium chloride were dissolved in 300 ml of distilled water. Once dissolved the final volume was made to 400 ml with distilled water. The solution was then autoclaved.

10% sodium dodecyl sulphate (SDS). 100 g of SDS were added to 900 ml of distilled water, mixed well and heated to 60°C to aid the dissolving. The pH was then adjusted to 7.2 using concentrated HCl and autoclaved.

Ethidium bromide. To make a 10 mg/ml solution of ethidium bromide, 1 g of ethidium bromide powder was added to 100 ml of distilled water and mixed to

ensure that the powder had completely dissolved. The container was wrapped with aluminium foil and stored at 4°C.

Agarose gel loading buffer. To make 20 ml of DNA electrophoresis loading buffer, 2 ml of 50x ELFO buffer and 4 ml of 20% glycerol were added to 14 ml of sterile distilled water. Bromothymol blue was added to give a final concentration of approximately 0.001%.

2.10 DNA manipulations

2.10.1 Restriction enzyme digest.

The required amount of DNA was incubated with the appropriate amount of restriction enzyme, 10x React buffer and sterile distilled water. Typically, reactions were performed in 20 µl volumes as shown below and all restriction enzymes/React buffers were obtained from GIBCO BRL:

DNA	x ng/µg
Restriction enzyme	1 µl
10x React buffer	2 µl
Water	y µl to 20µl

2.10.2 Alkaline lysis mini-preparation of plasmid DNA.

Solution I was made by adding 10 ml 40% glucose, 5 ml 1M Tris HCl and 4 ml 0.5M EDTA pH8.0 to 200 ml of distilled water. Solution II was made by adding 20 ml 10% SDS and 8 ml 5M NaOH to 200 ml of distilled water. Solution III was made by adding 23 ml glacial acetic acid and 120 ml 5M potassium acetate to 200 ml of distilled water. A culture of the required bacterium was set up in 5 ml of LB broth and incubated overnight at 37°C. Following growth, 1.5 ml of the culture was transferred to an Eppendorf tube and centrifuged for 3 minutes at 13000 rpm in a bench-top minifuge. The supernatant was discarded, another 1.5 ml of the culture transferred and the tube again centrifuged. The pellet was resuspended in 100 µl of

solution I (DNA preparation solution). 200 µl of solution II (lysis solution) were then added and the solution mixed gently until the solution went clear. 150 µl of solution III (protein precipitation solution) were then added and the tube vortexed upside down to form a white precipitate composed of protein and chromosomal DNA. The tube was centrifuged for 3 minutes at 13000 rpm. After centrifugation the supernatant was removed and transferred to a fresh tube. Phenol chloroform extraction was then performed. 800 µl volumes of ethanol were added to the resultant upper phase, the solution mixed and then centrifuged for 15 minutes at 13000 rpm to obtain a pellet of RNA and DNA. The supernatant was aspirated and washed in 1 ml of 70% ethanol by centrifugation for 3 minutes at 13000 rpm. The supernatant was again aspirated and the pellet dried in a vacuum drier. The DNA was then resuspended in an appropriate amount of sterile distilled water, 1 µl of RNAse (10 mg/ml) added and the solution incubated at 37°C for 15-30 minutes.

2.10.3 Qiagen® QIAprep Spin mini-preparation of plasmid DNA.

This method was adapted from the Qiagen® QIAprep Spin method and was used for the preparation of high-quality DNA for cloning where a small amount of DNA was required. The required bacterial strain was inoculated into 5 ml of LB broth with the addition of the appropriate antibiotics and incubated overnight at 37°C. After incubation the culture was centrifuged at 3000 x g for 10 minutes to pellet the cells. The pellet was resuspended in 250 µl of buffer P1 and transferred to a 1.5 ml Eppendorf tube. 250 µl of buffer P2 was added to the tube and the solution mixed gently. 350 µl of buffer N3 was then added to the tube and the solution again mixed gently. The suspension was centrifuged in a bench-top centrifuge at 13000 rpm for 5 minutes. After centrifugation the supernatant was carefully transferred to a QIAprep spin column placed within a 2 ml collection tube. The QIAprep spin column was washed with 0.75 ml of buffer PE and centrifuged in a bench-top centrifuge at 13000 rpm for 30 seconds. The flow-through was discarded. The QIAprep spin column was centrifuged again at 13000 rpm for 1 minute to remove residual buffer. The QIAprep spin column was placed in a clean 1.5 ml Eppendorf tube and the DNA eluted by adding 50 µl of sterile distilled water to the centre of the column

and leaving to stand for 1 minute. The column was then centrifuged at 13000 rpm for 1 minute to collect the DNA solution.

2.10.4 Qiagen® Tip100 midi-preparation of plasmid DNA.

This method was adapted from the Qiagen® Tip100 protocol and was used for the preparation of high-quality DNA for sequencing where a medium amount of DNA was required. The required bacterial strain was inoculated into 25 ml of LB broth with the addition of the appropriate antibiotics and incubated overnight at 37°C. After incubation the culture was centrifuged at 3000 x g for 10 minutes to pellet the cells. The supernatant was discarded and the pellet resuspended in 4 ml of buffer P1. 4 ml of buffer P2 were added, mixed gently and then incubated at room temperature for 5 minutes. 4 ml of chilled buffer P3 were then added, mixed gently and incubated on ice for 15 minutes. The suspension was then centrifuged at 3000 x g for 30 minutes at 4°C. During this time a Qiagen® Tip 100 was equilibrated by the addition of 4 ml of buffer QBT and allowing the buffer to move through the column by gravity flow. After centrifugation, the supernatant was carefully removed avoiding the transfer of cellular debris and applied to the Qiagen® Tip 100 and again allowed to move through the column by gravity flow. The flow-through was collected in a 50 ml Falcon tube and discarded. The Qiagen® Tip 100 was washed twice with 10 ml of buffer QC and the flow-through discarded. The DNA was eluted into a fresh tube by adding 5 ml of buffer QF to the Qiagen® Tip 100. 0.7 volumes (3.5 ml) of room-temperature isopropanol was added to the eluted DNA and mixed well. The mixture was then transferred to six 1.5 ml Eppendorf tubes and centrifuged in a bench-top centrifuge at 13000 rpm for 15 minutes. The supernatants were aspirated and discarded. The DNA pellets were then washed with 1 ml of 70% ethanol and the tubes centrifuged at 13000 rpm for 15 minutes. Again the supernatant was aspirated and discarded. The pellet was dried in a vacuum dessicator for 10 minutes. Each DNA pellet was resuspended in 40 µl of sterile distilled water after which the contents of each tube were combined.

2.10.5 Caesium chloride preparation of plasmid DNA.

This method was used for the preparation of high-quality plasmid DNA for sequencing and cloning and where a large amount or high concentration of DNA was required. A culture of the required bacterium was set up in 100 ml LB broth and incubated overnight at 37°C. Following growth, the culture was centrifuged in two Falcon tubes for 10 minutes at 3000 x g. The supernatant was discarded and the pellet resuspended in 5 ml of solution I. 10 ml of solution II were then added and the solution mixed gently until the solution went clear. 7.5 ml of solution III were then added and the solution shaken vigorously to form a white precipitate composed of protein and chromosomal DNA. The solution was centrifuged for 10 minutes at 3000 x g. After centrifugation the supernatant was removed, taking care not to remove the upper protein layer, and transferred to a fresh tube. Two volumes of ethanol were added, the solution mixed and then centrifuged for 10 minutes at 3000 x g to obtain a pellet of RNA and DNA. The supernatant was discarded and the tubes inverted and left to drain. When the pellets were dry they were resuspended to a final volume of 1.6 ml with sterile distilled water. 1.76 g of caesium chloride and 66µl of ethidium bromide were added to each tube, mixed and left for 10-15 minutes at room temperature. The tubes were then centrifuged for 5 minutes at 3000 x g to remove any RNA deposit and the supernatants transferred to separate Beckman ultracentrifuge tubes. The tubes were balanced, sealed and placed in a Beckman ultracentrifuge for at least 4 hours at 100 000 rpm at 20°C. The rotor acceleration speed was set at maximum and the deceleration speed was set at minimum. After centrifugation the samples were removed from the centrifuge and, using appropriate bore needle and syringe, the lower plasmid DNA band removed with care. The DNA solution was then washed at least four times with caesium chloride/water-saturated isopropanol to remove the ethidium bromide. Once clean an equal amount of water and 0.6 volumes of isopropanol were added. The solution was then centrifuged for 15 minutes at 13000 rpm in a bench-top minifuge to obtain a colourless pellet of DNA. The supernatant was then aspirated and the pellet washed twice with 1 ml 70% ethanol by centrifugation for 3 minutes at 13000 rpm to remove the caesium chloride. The supernatant was then aspirated and the pellet

vacuum dried for 10 minutes. The DNA was then resuspended in an appropriate amount of sterile distilled water.

2.10.6 Mini-preparation of chromosomal DNA.

This method is adapted from Chen *et al* (1993). The required bacterial strain was inoculated into 5 ml of LB broth with the addition of appropriate antibiotics and was incubated overnight at 37°C. After incubation, three 1.5 ml aliquots were placed into 1.5 ml Eppendorf tubes and centrifuged for 3 minutes at 13000 rpm in a bench-top centrifuge. The supernatants from each were discarded and the deposits resuspended in 200 µl of lysis buffer (40 mM Tris-acetate pH 7.8, 20 mM sodium-acetate, 1mM EDTA, 1% SDS). To each tube, 66 µl of 5M NaCl were then added, the solutions mixed well and the mixtures then centrifuged for 10 minutes at 13000 rpm in a bench-top centrifuge. The supernatants were recovered and combined in one Eppendorf tube. An equal volume of phenol was added and gently mixed but not vortexed. The mixture was briefly centrifuged at 13000 rpm and the supernatant recovered to a clean Eppendorf tube. An equal volume of chloroform was added and mixed gently but not vortexed. The mixture was briefly centrifuged at 13000 rpm and the supernatant recovered to a clean Eppendorf tube. The chloroform step was repeated as necessary. The upper phase was removed to a clean Eppendorf tube and 800 µl of ethanol added, the solution mixed and then centrifuged for 15 minutes at 13000 rpm to obtain a pellet of RNA and DNA. The supernatant was aspirated and washed in 1 ml of 70% ethanol by centrifugation for 3 minutes at 13000 rpm. The supernatant was again aspirated and the pellet dried in a vacuum drier. The DNA was then resuspended in an appropriate volume of sterile distilled water, 1 µl of RNAse (10 mg/ml) added and the solution incubated at 37°C for 15-30 minutes.

2.10.7 CTAB preparation of chromosomal DNA.

This method, based on that of Ausubel *et al* (1995), was used to prepare large amounts of chromosomal DNA where the quality of the DNA was not necessarily required to be of high quality. The required bacterial strain was inoculated into 50 ml of LB broth with the addition of the appropriate antibiotics and incubated

overnight at 37°C. After incubation the culture was centrifuged at 3000 x g for 10 minutes.

TE buffer: 10 mM Tris-HCl pH 7.5, 1 mM EDTA

CTAB/NaCl soln: 4.1 g NaCl were dissolved in 80 ml distilled water followed by 10% w/v CTAB (hexadecyltrimethyl ammonium bromide) (Sigma) with gentle heating and stirring. The final volume was adjusted to 100 ml. The solution was not autoclaved.

The supernatant was discarded and the pellet resuspended in 9 ml TE buffer. 2 ml of 10% SDS and 100 µl of 20 mg/ml Proteinase K were added to the resuspended pellet, mixed gently and incubated at 37°C for one hour. After incubation, 1.8 ml of 5 M NaCl were added and the solution mixed gently but thoroughly. 1.5 ml of CTAB/NaCl solution were then added, mixed and the solution incubated at 65°C for 20 minutes. After incubation an equal volume of 1:24 chloroform/isoamyl-alcohol was added to the solution, mixed thoroughly and then centrifuged at 3000 x g for 20 minutes at 4°C. After centrifugation the supernatant was transferred to a clean tube and 5 ml of phenol/chloroform added. The solution was mixed gently and then centrifuged at 3000 x g for 10 minutes. The supernatant was again transferred to a clean tube and 5 ml of phenol/chloroform added. The tube was then mixed gently and centrifuged at 3000 x g for 10 minutes. After centrifugation the supernatant was transferred to a clean tube and a 0.6 volume of ice-cold isopropanol added to the supernatant and the solution mixed gently until the DNA precipitated. The DNA was transferred to 1 ml of 70% ethanol using a sterile Pasteur pipette and washed. The tube was then centrifuged at 3000 x g for five minutes to pellet the DNA. The 70% ethanol was aspirated and discarded, and the pellet dried in a vacuum dessicator. Once dry the DNA was resuspended in an appropriate volume of sterile distilled water, 1 µl of RNAse (10 mg/ml) added and the solution incubated at 37°C for 15-30 minutes.

2.10.8 Caesium chloride preparation of chromosomal DNA.

This method was used to prepare large amounts of high quality chromosomal DNA. The required bacterial strain was inoculated into 50 ml of LB broth and incubated overnight at 37°C. Following growth, the culture was centrifuged in two Falcon tubes for 10 minutes at 3000 x g and the supernatants discarded. The pellets were then resuspended in 20 ml of STE (50 mM sucrose, 10 mM Tris, 25 mM EDTA). 2 ml of 10% SDS and 100 µl of 20 mg/ml proteinase K were added to each tube and the solutions mixed. The solutions were incubated at 37°C for approximately one hour or until they were clear. An equal volume of phenol chloroform was then added to each tube, mixed by inversion and then centrifuged for 10 minutes at 3000 x g. The upper phase was recovered with care, an equal volume of chloroform added and the tubes mixed well by inversion. The tubes were then centrifuged for 10 minutes at 3000 x g. The upper phase was recovered and a 1/10 volume of 3M sodium acetate and two volumes of ethanol added. The solutions were mixed by inversion to form a DNA precipitate. The DNA was hooked out by using a bent glass pipette and left to air-dry. Each DNA precipitate was then placed in a suitable amount of distilled water (final volume 3.2 ml) and eluted at 50°C. 3.52 g of caesium chloride and 132 µl of 10 mg/ml ethidium bromide were then added to each tube. The solutions were gently mixed and left for 10-15 minutes before being centrifuged for 5 minutes at 3000 x g to remove any RNA deposit. The supernatants were removed and transferred to separate Beckman tubes which were then balanced, sealed and placed in an ultracentrifuge for at least 4 hours at 100 000 rpm at 20°C. The rotor acceleration speed was set at maximum and the deceleration speed at minimum. After centrifugation the samples were removed and, using appropriate bore needle and syringe, the upper chromosomal DNA band removed with care. The DNA was washed at least four times with caesium/water-saturated isopropanol to remove the ethidium bromide. Once clean an equal amount of water and 0.6 volumes of isopropanol were added. The solution was then centrifuged for 15 minutes at 13000 rpm in a bench-top minifuge to obtain a pellet of DNA. The supernatant was then aspirated and the pellet washed twice with 1 ml 70% ethanol by centrifugation for 3 minutes at 13000 rpm to remove the caesium chloride. The supernatant was then aspirated and the pellet vacuum dried for 10 minutes. The

DNA was then resuspended in an appropriate amount of sterile distilled water (approximately 250 µl) and the DNA eluted at 50°C.

2.11 Preparation of agarose gel

The appropriate amount of agarose powder was added to 100 ml ELFO buffer to obtain the required percentage agarose. The usual concentrations used were 1% and 2%, and therefore 1 g or 2 g of agarose powder were added to 100 ml ELFO buffer respectively. The solution was microwaved with care until the agarose powder was fully dissolved. For each 100 ml of solution, 5 µl of ethidium bromide (10 mg/ml) were added. The agarose solution was then stored at 55°C until use.

2.12 Agarose gel electrophoresis

Appropriate amounts of DNA were diluted in appropriate volumes of sterile distilled water in an clean Eppendorf tube. 2 µl of loading buffer were then added and the solution mixed. The samples were transferred to the appropriate wells of an agarose gel in an electrophoresis tank. λ*Hind*III and/or ϕX174RF DNA/*Hae*III standard DNA molecular weight markers were included as required. A voltage of 100V DC run through the gel. The DNA was run until the loading buffer had travelled approximately two-thirds of the way down the gel. The gel was removed from the tank and observed safely using an ultraviolet transilluminator.

2.13 Phenol chloroform extraction

The DNA solution was washed in an equal volume of phenol chloroform and centrifuged for 3 minutes at 13000 rpm in a bench-top microfuge. The supernatant was recovered and placed in an Eppendorf tube. An equal volume of 24:1 chloroform: isoamyl alcohol was added and the solution mixed. The tube was then centrifuged as before. The supernatant was recovered and transferred to a fresh Eppendorf tube.

2.14 Ethanol precipitation

A 1:10 volume of sodium acetate and two volumes of 95% ethanol (and optional 1µl of 10 mg/ml tRNA) were added to the DNA solution. The tube was centrifuged for 15 minutes at 13000 rpm in a bench-top microfuge. The supernatant was aspirated taking care not to dislodge the DNA pellet. The DNA was then washed in 1 ml of 70% ethanol and centrifuged for 3 minutes at 13000 rpm. The supernatant was again aspirated taking care not to dislodge the pellet or leave any alcohol behind. The tube was placed in a vacuum dessicator for 10 minutes or until the DNA pellet was dry. The DNA was then redissolved in an appropriate volume of distilled water.

2.15 Preparation of calcium chloride-competent *E. coli*

A culture of *E. coli* was set up in 5 ml of LB broth and incubated overnight at 37°C. Following growth, a 1:50 dilution culture of *E. coli* in LB broth was grown at 37°C to an optical density of 0.6 at 600 nm to obtain an exponential culture. After such time, 1 ml of the culture was centrifuged in an Eppendorf tube for 3 minutes at 13000 rpm in a bench-top microfuge. The pellet was resuspended in 1 ml of ice-cold 100mM magnesium chloride and centrifuged for 3 minutes at 13000 rpm. The supernatant was discarded and the pellet resuspended in 1 ml of 100 mM calcium chloride. The suspension was left on ice for 30 minutes after which time it was centrifuged for 3 minutes at 13000 rpm. The supernatant was then discarded and the pellet resuspended in 100 µl of 100 mM calcium chloride. The suspension was left on ice until use. If the cells were not required for immediate use they were frozen at -80°C in a final concentration of 25% glycerol.

2.16 Plasmid DNA transformation using calcium chloride-competent *E. coli*

This method was used where a high transformation efficiency was not required. Calcium chloride-competent *E. coli* were taken and the required amount of DNA added. The preparation was left on ice for 10 to 20 minutes after which it was heat

shocked at 42°C for 90 seconds. 1 ml of SOC medium was added and incubated at 37°C for 30 to 60 minutes. Appropriate dilution of the transformed cells were plated onto LBA with the addition of appropriate antibiotics.

2.17 Determination of DNA concentration

A Pharmacia Biotech Ultrospec 2000 was used for this purpose. The UV lamp was turned on and allowed to warm up. 1 ml of distilled water was transferred to a clean quartz cuvette and the spectrophotometer zeroed at 260 nm and 280 nm. 10 µl of DNA was then added to 990 µl of distilled water in a clean quartz cuvette and the absorbance again read at 260 nm and 280 nm to provide the DNA concentration reading. The quality of the DNA was ascertained by noting the absorbance ratio (maximum quality 1.8).

2.18 Isolation of DNA fragments by gel electrophoresis

The required amount of DNA was run on an agarose gel by electrophoresis and after the appropriate time was observed safely using an ultraviolet transilluminator. The required DNA band or bands were excised from the gel using a clean scalpel. A small hole was made through the bottom of a 0.5 ml Eppendorf tube and a small amount of polyallomer wool placed inside the tube. The lid was broken off a large Eppendorf tube and the small tube placed inside. No more than 1 cm^2 of agarose gel containing the DNA was then placed in the small Eppendorf tube and centrifuged for 5 minutes at 13000 rpm in a desk-top minifuge. The liquid was recovered from the large tube and purified using the Qiagen® QIAquick PCR purification method. Briefly, five volumes of buffer PB were added to one volume of the unpurified DNA solution and mixed well. The solution was then applied to a QIAquick column placed inside a 2 ml collection tube and then centrifuged in a bench-top centrifuge at 13000 rpm for 30 seconds. The flow-through was discarded. The column was washed by 0.75 ml buffer PE and centrifuged at 13000 rpm for 30 seconds. The flow-through was again discarded and the column re-centrifuged for one minute to remove residual buffer. The QIAquick PCR column was then placed in a clean 1.5

ml Eppendorf tube and the DNA was eluted by adding 50 µl of sterile distilled water to the centre of the column. The column was centrifuged in a bench-top centrifuge at 13000 rpm for one minute to collect the DNA solution.

2.19 Dephosphorylation of DNA

A 10 µl volume of appropriately diluted vector DNA was placed in a clean Eppendorf tube. 2 µl of 10x alkaline phosphatase buffer, 1 µl of alkaline phosphatase (0.01 units) and 7 µl of distilled water were then added to give a final total volume of 20 µl. The solution was then incubated at 37°C for one hour. After incubation, the DNA was cleaned by phenol/chloroform and chloroform extraction unless gel electrophoresis was being performed immediately after dephosphorylation.

2.20 DNA ligations

Standard ligations were set up as below. Volumes of individual components were varied as necessary although it was usual to use the molar concentration of DNA insert to DNA vector at a ratio of 2:1.

5x ligation buffer	2 µl
10mM ATP	1 µl
DNA vector	x (dependent on DNA concentration)
DNA insert	y (dependent on DNA concentration)
Ligase	0.1 µl
Total volume	10 µl

The ligations were incubated at 16°C for at least four hours or overnight as convenient. Following incubation the ligation mixture was ethanol precipitated with the addition of tRNA. When performing ligations ready for transformations, control ligations were included as well as test ligation(s). Controls included reactions with

and without ligase for each vector to be used and a transformation control using undigested vector.

2.21 Preparation of electro-competent *E. coli* cells

A culture of *E. coli* (DH5α) was set up in 5 ml of LB broth and incubated overnight at 37°C. Following growth, a 1:50 dilution culture of *E. coli* in 50 ml LB broth was grown at 37°C to an optical density of 0.6 to obtain an exponential culture. The culture was then chilled on ice for 10 minutes and centrifuged for 10 minutes at 3000 x g. The supernatant was removed and discarded, the pellet resuspended in 50 ml of ice cold distilled water and centrifuged as before. This was repeated at least four times to ensure that the cells were salt free. After the last centrifugation step the pellet was resuspended in 100 µl of distilled water in an Eppendorf tube and left on ice until use.

2.22 Electroporation of plasmid DNA

Electroporation is an efficient method for the transformation of bacteria with plasmid DNA (Miller 1994). The method employed here was based upon that of Dower and colleagues (1988). The appropriate number of sterile Bio-Rad Gene Pulser®/ *E. coli* Pulser™ cuvettes with an electrode gap of 0.1 cm were obtained and chilled on ice. 40 µl of competent DH5α cells were transferred to a cuvette and exposed to a voltage of 1.5 kV with a resistance of 1000 Ω and a capacitance of 25 µFD to check the time constant. If the time constant was good, approximately 23.0, then the method was continued for the ligation mixture. 40 µl of competent cells were mixed with each ligation mixture and then each transferred to a cuvette and exposed as before. The cells were immediately transferred to 1 ml of SOC medium and incubated at 37°C for 45 minutes. After such time the cells were plated onto an appropriate medium.

2.23 Southern blot analysis of DNA

The method was performed according to the Amersham Life Science *Fluorescein Gene Images*™ labelling and detection system protocol and based on the method of Southern (1975).

Preparation of solutions. Depurinating solution was 0.25 M HCl. Denaturing solution consisted of 0.5 M NaOH and 1.5 M NaCl. Neutralising solution consisted of 1.0 M Tris HCl pH7.5 and 1.5 M NaCl. Hybridisation buffer was prepared by mixing 25 ml of 20xSSC, 1 ml of 10% SDS, 5 g of dextran sulphate and 5 ml of liquid block (as supplied by manufacturer) and making up to 100 ml with distilled water. Stringency buffer 1 was 1xSSC, 0.1% SDS whilst stringency buffer 2 was 0.5xSSC, 0.1% SDS. Diluent buffer was prepared by mixing 100 mM Tris HCl pH 7.5 with 300 mM NaCl and making up to 1 litre with distilled water. Liquid block was as supplied by the manufacturer.

Preparation of labelled probe. DNA to be used as the probe was diluted to a concentration of between 2 and 25 ng/µl in distilled water and then denatured by boiling for 5 minutes in a minimum volume of 20 µl. After boiling the DNA was rapidly cooled on ice. The labelling reaction was made by mixing 10 µl of nucleotide mix, 5 µl of primers, 50 ng minimum of denatured DNA, 1 µl of enzyme and distilled water to a final volume of 50 µl. The reaction mixture was incubated at 37°C for 1 hour (or room temperature overnight) and then terminated by the addition of EDTA to a final concentration of 20 mM. The probe was stored at -20°C if not for immediate use.

Transfer of DNA to nylon membrane. Excess agarose gel was removed so as to reduce its volume for washing. The gel was then placed in a suitable container and given two 10 minute washes in depurinating solution followed by two 10 minute washes in denaturing solution and two 10 minutes washes in neutralising solution. The DNA was then transferred to a nylon membrane using the standard blot method

either overnight or in a few hours. After the DNA had transferred it was fixed to the membrane by exposing it to ultraviolet light at an intensity of 70 mJ/cm^2 using an Amersham UV crosslinker.

Hybridisation and stringency washes. The membrane was washed in pre-warmed hybridisation buffer for at least 30 minutes at 65°C. The required amount of probe was taken, and if not freshly boiled, was boiled for 5 minutes. The probe was then added to the hybridisation buffer for at least 6 hours during which time the stringency buffer solutions were pre-heated to 65°C. After hybridisation had occurred the blot was transferred to stringency buffer 1 and washed for 15 minutes at 60°C. The first buffer was then removed and replaced with stringency buffer 2 and the blot washed for 15 minutes at 60°C.

Blocking, antibody incubation and washes. The stringency buffer 2 was poured off and the blot rinsed briefly in an excess of diluent buffer. The buffer was then poured off and the blot was incubated in 1:10 liquid block:diluent buffer for one hour at room temperature. Again the blot was rinsed briefly in an excess of diluent buffer and the buffer poured off. Anti-fluorescein-AP conjugate was diluted 1:5000 in 0.5% bovine serum albumin in diluent buffer. This was added to the blot and incubated at room temperature for one hour. After incubation the conjugate was poured off and the blot washed three times for 10 minutes each at room temperature in 0.3% Tween-20 in diluent buffer. This was then poured off and the blot rinsed briefly in an excess of diluent buffer.

Signal generation and detection. Excess fluid was drained from the blot. The blot was then sprayed with dioxetane spray, any excess drained off and the blot then placed in Saran™ wrap film. The blot was exposed to photographic X-ray film in a film cassette for at least one hour or overnight depending on the DNA being analysed. After exposure the film was developed. If necessary, the blot was washed three times in 50 ml of 0.1% SDS for 15 minutes each to strip the probe/antibody complex from the blot. The blot could then be re-used at a future date.

2.24 Colony blot analysis of DNA

Colonies were lifted onto a nylon membrane and incubated for at least four hours at 37°C. Alternatively colonies were inoculated directly onto a gridded nylon membrane and again incubated for four hours at 37°C. Denaturing solution was then poured over the membrane and left for 5 minutes in a suitable container to lyse the cells and denature the DNA. This was repeated with neutralising solution. The membrane was then baked at 80°C for one hour to fix the DNA. Excess bacterial growth was scrubbed off with a towel moistened with stringency buffer 1. Hybridisation and stringency washes were then performed as already described for the Southern blot method followed by blocking, antibody incubation, washes, signal generation and detection.

2.25 Creation of sub-clones by nested deletions

Nested deletions were performed to create sub-clones and allow subsequent sequencing. These were performed according to the method of Henikoff (1984) using the Pharmacia double-stranded Nested Deletion kit.

2.26 Primer preparation

Primers were designed according to their use from known DNA sequences and were prepared by the Protein and Nucleic Acid Chemistry Laboratory of Leicester University. Stock primers were prepared for use using ethanol precipitation by adding 200 µl of primer to 20 µl 3 M sodium acetate and 600 µl ethanol. The solution was centrifuged in a bench-top centrifuge at 13000 rpm for 15 minutes after which the supernatant was aspirated and discarded. The pellet was washed in 1 ml 70% ethanol by centrifugation for 5 minutes. The supernatant was again aspirated and discarded and the pellet dried using a vacuum dessicator. Once dry, the pellet was resuspended in 100 µl sterile distilled water and the concentration of primer determined in pmol/µl.

2.27 Polymerase chain reaction

Stock 10x dNTP mix was prepared by mixing 100 mM dATP, 100 mM dCTP, 100 mM dGTP and 100 mM dTTP (all Advanced Biotechnologies). 10x PCR buffer was prepared by mixing 500 mM KCl, 100 mM Tris-HCl, pH8.3, 1.8 mg/ml purified bovine serum albumin (Gibco BRL) and 50 mM $MgCl_2$.

A typical PCR reaction consisted of 1-10 pg/µl of the DNA template, 1 µl of 10x dNTPs, 1 µl of 10x PCR buffer, 1 µl of primer A, 1 µl of primer B, 5 µl of sterile distilled water and 0.05 µl of Taq DNA polymerase (Advanced Biotechnologies). The mixture was overlaid with 20 µl of PCR grade mineral oil. PCR reactions were performed in a Hybaid OmniGene thermal cycler, the conditions for which varied between experiments. Individual conditions are described in the relevant results chapters. Upon completion of the PCR reaction, 5 µl of loading buffer was added to the PCR mixture and the product analysed by agarose gel electrophoresis.

2.28 Automated DNA sequencing

Automated sequencing of DNA was performed according to the Perkin Elmer ABI PRISM™ Dye Terminator Cycle Sequencing protocol. 8 µl of sequencing mix, containing A-Dye terminator, C-Dye terminator, D-Dye terminator, T-Dye terminator, dITP, dATP, dCTP, dTTP, Tris-HCl (pH9.0), $MgCl_2$, thermostable pyrophosphatase and AmpliTaq DNA polymerase FS, followed by 3.2 pmol of the appropriate primer and 200-300 ng of DNA, was placed in a 0.5 ml thin-walled Eppendorf tube. The final volume was made to 20 µl with sterile distilled water followed by 20 µl of PCR-grade oil. The tube was then vortexed for five seconds to ensure that the contents were mixed well and then centrifuged briefly to deposit the contents. The tube was placed in a Hybaid OmniGene thermal cycler and the cycling set for denaturing at 96°C for ten seconds followed by annealing at 50°C for five seconds and extending at 60°C for four minutes. This cycle was repeated 25 times. After thermal cycling the extension products were purified by removing the

20 µl sequence reaction and adding to 50 µl of ethanol and 2 µl of 3 M sodium acetate. The reaction was then centrifuged in a bench-top centrifuge at 13000 rpm for 15 minutes, washed once with 500 µl of 70% ethanol and the pellet dried in a vacuum dessicator. The reaction was then sent to the Protein and Nucleic Acid Chemistry Laboratory of Leicester University for automated sequencing on a Perkin-Elmer ABI Sequencer (model 377).

2.29 Computer analysis of DNA and predicted protein sequences

DNA sequences were received via the Protein and Nucleic Acid Chemistry Laboratory Macintosh computer server. These sequences were compiled and analysed using the Seqed, Translate, Gap, Bestfit, Map and Frames programs on the Wisconsin Molecular Biology Package version 9.1-UNIX (Genetics Computer Group, Madison, Wisconsin). DNA and predicted protein sequences were compared with others on the GenBank databases using the BLAST (Altshul *et al* 1990) or Gapped-BLAST (Althsul *et al* 1997) programs available on-line at the National Centre for Biotechnology Information website (http://www.ncbi.nlm.nih).

2.30 Isolation of a spontaneous streptomycin-resistant variant of EPEC E2348/69

A fresh strain of the standard clinical isolate EPEC strain E2348/69, as used by Baldini *et al* 1983, was obtained from Dr S. Knutton, Institute of Child Health, Birmingham. Upon receipt the strain possessed no useful antibiotic resistance marker required for genetic manipulations. Spontaneous chromosomal streptomycin resistance was introduced by the following method. 5 ml of LB were inoculated with 5 µl of stock streptomycin to give a sub-lethal final concentration in the medium of 25 µg/ml. The LB was then inoculated with EPEC E2348/69 from a pure LBA culture and inoculated overnight at 37°C with shaking. 1 ml of the culture was removed from the culture and centrifuged in a bench-top centrifuge at 13000 rpm for 5 minutes, the supernatant discarded and the deposit resuspended in 50 µl of

sterile distilled water. The 50 µl suspension was then plated onto LBA containing streptomycin at a concentration of 50 µg/ml and incubated overnight at 37°C. A dozen streptomycin-resistant colonies were selected and plated onto LBA containing streptomycin concentrations of 50 µg/ml, 100 µg/ml, 150 µg/ml and 200 µg/ml. Those EPEC that possessed a chromosomal mutation conferring streptomycin resistance grew on streptomycin concentrations up to 150 µg/ml. One such variant was selected for further genetic manipulation.

2.31 Construction of a mutant bacterial strain

2.31.1 Construction of a gene mutation by frameshift deletion.

Genes can be mutated by introducing a frameshift at the desired base-pair using certain enzymes (Ausubel 1995). To create a frameshift deletion the following was performed:

a) Nucleotide degradation using Klenow exonuclease: to achieve nucleotide degradation ('chew-back'), approximately 2 µg of plasmid DNA were digested at 37°C for 30 minutes in a solution containing 4 µl of ligase buffer, 2 µl of Klenow exonuclease (Gibco BRL) and 2 µl of sterile distilled water.

b) Repairing 3' overhanging ends to generate blunt ends: to repair the 3' overhanging ends so as to generate blunt ends, 2 µl of 10 mM dNTPs were added to the above 18 µl reaction and incubated at room temperature for 30 minutes. After incubation the DNA was purified by phenol/chloroform extraction and ethanol precipitation. A standard ligation reaction was then performed to religate the blunt ends and this was transformed into the appropriate bacterial strain as previously described.

2.31.2 Transfer of a frameshift gene mutation to the suicide vector pRDH10.

The method used here for the homologous recombination of a frameshift gene mutation is based upon that of Miller and Mekalanos (1988). The method involves homologous recombination between the wild-type allele and mutant allele using a

suicide vector containing the *sacRB* gene. A sucrose selection procedure on low salt concentrations and at 30°C can then be used since sucrose is toxic for those strains possessing the *sacRB* gene. Selection therefore results in a homologous recombination event and loss of the suicide vector.

The frameshift gene mutations were transferred to the suicide vector pRDH10 (Haigh R, personal communication) using standard cloning techniques as described in **sections 2.18 to 2.20** and transformed into the *E. coli* strain CC118λ*pir* by electroporation as described in **sections 2.21 and 2.22**. Transformants were checked for chloramphenicol resistance and tetracycline sensitivity. Plasmid was prepared from a chosen transformant using the plasmid mini-preparation method described in **section 2.10.3** and transformed into SM10λ*pir* by electroporation. Transformants were checked for chloramphenicol and kanamycin resistance.

2.31.3 Conjugation of pRDH10 from SM10λ*pir* to wild-type strains.

SM10λ*pir* and the wild-type strains were each inoculated into 5 ml of LB broth and incubated at 37°C to OD 0.6 at 600 nm without the addition of antibiotics. One millilitre of each culture was then centrifuged for 10 minutes at 4000 rpm, the supernatant discarded and each deposit resuspended in 50 µl of sterile distilled water. Pairs of 50 µl cultures for conjugation, being SM10λ*pir* and each wild-type strain, were mixed with each other and the total 100 µl pipetted onto a 25 mm Sartorius cellulose nitrate 0.45 µm pore-size filter and placed on the surface of a LBA plate (three filters to a plate). The agar plates were incubated for at least one hour at 37°C. After incubation the cells were washed off the filter into 1 ml of LB broth. The cells were plated at the appropriate dilution onto media containing the necessary antibiotics and incubated overnight at 37°C. After incubation, colonies were patch plated onto LBA containing the appropriate antibiotics and incubated overnight at 37°C for further analysis.

2.31.4 Sucrose selection of strains that have lost pRDH10.

A representative merodiploid colony resulting from the conjugation of pRDH10 from SM10λ*pir* to the wild-type strain was plated to single colonies on LBA containing the appropriate antibiotics. After growth, the merodiploid was grown in 5 ml 6% sucrose in LB (without salt) to stationary phase at 30°C. The resultant culture was plated onto 6% sucrose in LBA at a dilution of 1/100 000 and incubated overnight at 30°C. Revertant colonies were patch-plated onto LBA containing chloramphenicol and also onto the wild-type resistance marker and incubated overnight at 37°C.

2.31.5 Screening of revertants for the presence of chromosomal gene mutation.

Revertants were screened by PCR, using the method described in **section 2.27**, using primers that annealed either side of the mutation followed by digestion of the PCR product with the appropriate enzyme which was present in wild-type strains but absent in mutant strains. Mutants were confirmed by Southern blot analysis using the method described in **section 2.23**.

2.32 Growth curve assay

This assay was performed in duplicate for each bacterial strain. An overnight culture of the desired strain was diluted, usually 1/50 (1 ml culture in 50 ml medium), in the pre-warmed culture medium, either LB or DMEM, without the addition of antibiotics. The culture was incubated in a New Brunswick Scientific gyrotory water bath shaker (model G76D) at 37°C with shaking at 270 rpm. 1 ml samples were taken at 0 minutes and every 30 minutes, placed in a 2 ml disposable plastic cuvette and the OD at 600 nm determined using a Pharmacia Biotech Ultrospec 2000 spectrophotometer after it had been blanked using 1 ml of the appropriate medium. Readings were taken until the bacterial culture reached stationary phase which was assumed when three similar consecutive readings were recorded.

2.33 Maintenance of tissue culture cells

HEp-2 cells (derived from human laryngeal carcinoma cells; Toolan 1954) were obtained from Dr S. Knutton, Institute of Child Health, Birmingham. Cell lines were routinely cultured as monolayers on 80 cm^2 tissue culture flasks (Nunc) in DMEM containing 10% ($^{v}/_{v}$) heat inactivated, mycoplasma screened foetal calf serum (FCS) and incubated at 37°C, 5% CO_2. After growth, confluent cell monolayers were washed twice with 10 ml sterile PBS. 2 ml of PBS and 1 ml of 0.25% trypsin in 5 mM KCl, 2.6 mM Na_2HCO_3, 115 mM NaCl, 5.5 mM glucose, 0.05% phenol red were then added and the monolayers incubated at 37°C for approximately 3 minutes until the cells showed rounding. Cells were dislodged from the monolayer by striking the side of the flask against the palm of the hand. 10 ml of DMEM/FCS were added to the flask and the cells gently resuspended by pipetting. Clean flasks containing DMEM/FCS were taken and 1 ml of cell suspension was added for every 10 ml of fresh DMEM/FCS. The cells were again grown at 37°C, 5% CO_2 until confluent with medium changes every 2 days.

When coverslip tissue cultures were required, cell suspensions were inoculated onto 13 mm glass coverslips in approximately 25 ml DMEM/FCS, in a sterile plastic Petri dish. These were incubated at 37°C, 5% CO_2. If these were for use in adhesion or FAS test assays then cells were grown until 50% confluent.

2.34 Long term storage of tissue culture cells

Trypsinised cell suspensions were pelleted by centrifugation at 1000 rpm for 2 minutes in 10 ml conical plastic tubes (Sterilin). Cell pellets were resuspended in 1 ml of 10% ($^{v}/_{v}$) FCS/10% ($^{v}/_{v}$) DMSO in DMEM by gentle agitation and transferred to 2 ml plastic cryo-tubes (Sarstedt). The tubes were placed in an insulated freezing box and the temperature slowly reduced to -80°C. Tubes were finally stored under liquid nitrogen. Cells were revived from storage by fast thawing at 37°C followed by inoculation into DMEM/FCS and incubation at 37°C, 5% CO_2.

2.35 Fluorescent actin staining (FAS) test

The adherence part of this method was based upon that of Cravioto *et al* (1979) whilst the remaining method was based upon that of Knutton *et al* (1989) and was used for visualising actin accumulation by attaching and effacing bacteria when adhering to cultured epithelial cells. 5 ml of LB were inoculated with the required organism and incubated overnight at 37°C with the addition of the appropriate antibiotics. After incubation, 2 ml of DMEM were placed into each well of a Corning six well (35 mm diameter) flat bottom plate and incubated at 37°C, 5% CO_2 for one hour to equilibrate the medium. Semi-confluent cultures of HEp-2 cells on sterile 13 mm diameter cover slips were then placed in each well followed by 40µl of overnight bacterial culture. The plate was then incubated at 37°C, 5% CO_2 for three hours. After incubation the DMEM was aspirated and the HEp-2 cells washed three times with PBS to remove non-adherent bacteria. The cells were then fixed in 3% formaldehyde for 20 minutes. The fixed cells were permeabilised using 0.1% Triton X-100 in PBS for 5 minutes. The cells were washed three times with PBS and then stained with 20 µl of 10 µg/ml fluorescein isothiocyanate-phalloidin (Sigma) in PBS for 30 minutes. The cells were again washed three times with PBS and each cover slip mounted in 10 µl anti-fade (50% glycerol in PBS containing 1 mg/ml *p*-phenylenediamine) on a microscope slide and visualised at 400x magnification using a Zeiss Axiophot fluorescence microscope.

2.36 Cultured epithelial cell adhesion assay

The assay was performed with triplicate samples. HEp-2 cells from a maintained stock were seeded to 25-50% confluent growth in 24 well tissue culture plates and allowed to grow in DMEM/10% foetal calf serum until almost confluent at 37°C in 5% CO_2. The culture medium was then replaced with 0.5 ml per well of DMEM containing 0.5% w/v D-mannose and the plates preincubated for one hour at 37°C in 5% CO_2. After preincubation, the cells in each well were infected with $2x10^7$ bacteria (5 µl of an overnight LB culture) and incubated at 37°C in 5% CO_2 for 3

hours. After incubation, the cell monolayers were washed 3 times with PBS and the bacteria released by solubilisation of the HEp-2 cells by the addition of 1 ml of 1% Triton X-100 in PBS for 20 minutes at room temperature. The resultant bacteria were enumerated by appropriate dilution in duplicate and plating onto LBA with overnight incubation at 37°C. Plates containing approximately 10^2 colonies were chosen for counting.

2.37 Gentamicin cultured epithelial cell protection (invasion) assay

This method was based upon that of Ellinghorst (1994). HEp-2 cells from a maintained stock were seeded to 25-50% confluent growth in 24 well tissue culture plates and allowed to grow in DMEM/10% foetal calf serum until almost confluent at 37°C in 5% CO_2. The culture medium was then replaced with 0.5 ml per well of DMEM containing 0.5% w/v D-mannose and the plates preincubated for one hour at 37°C in 5% CO_2. After preincubation, the cells in each well were infected with $2x10^7$ bacteria (5 µl of an overnight LB culture) and incubated at 37°C in 5% CO_2 for 3 hours. After incubation, the cell monolayers were washed 3 times with PBS, 0.5 ml DMEM containing 100 µg/ml gentamicin added, and incubation was continued for 2 hours. The monolayers were then washed 3 times with PBS and the bacteria released by solubilisation of the HEp-2 cells with 1 ml of 1% Triton X-100 in PBS for 20 minutes at room temperature. The resultant bacteria were enumerated by appropriate dilution in duplicate and plating onto LBA with overnight incubation at 37°C. Plates containing approximately 10^2 colonies were chosen for counting.

2.38 Alkaline phosphatase assay

This standard assay for measuring the alkaline phosphatase activity of a growing culture was adapted from the method of Brickman and Beckwith (1975). An overnight culture of the required strain was diluted 1/50 in colourless DMEM and grown at 37°C with shaking at 270 rpm in a New Brunswick Scientific gyrotory water bath shaker (model G76D). At regular intervals the A_{600} of the culture was measured and at each time point two 1 ml samples were taken and the bacteria

pelleted by centrifugation at 13000 rpm for 5 minutes. The supernatants were discarded and the pellets resuspended together in 1 ml of 1M Tris-HCl, pH8.0. The bacteria were preincubated in a static waterbath at 37°C for 5 minutes. 100µl of *p*-nitrophenol phosphate (NPP; 4 mg/ml; Sigma) were added to the tube and incubated at 37°C until a yellow colour became visible. The reaction was stopped by adding 100 µl of 1 M K_2HPO_4 noting the time taken since the addition of NPP. The sample was transferred to a 2 ml cuvette and the A_{420} and A_{550} measured. The blank sample was 1 ml Tris-HCl pH8.0, 100 µl NPP, 100 µl 1 M K_2HPO_4. The alkaline phosphatase activity was calculated using the following equation:

$$\text{Activity (in Miller units)} = \frac{A_{420} - (1.75 \times A_{550}) \times 1000}{\text{time (minutes)} \times \text{sample volume} \times A_{600}}$$

2.39 Western blot analysis of proteins

2.39.1 Preparation of bacterial lysates.

100 ml of LB were inoculated with the appropriate bacterial strain and antibiotic(s) in a 250 ml baffled conical flask and grown in a shaking water bath at 270 rpm, 37°C to an OD 0.6 at 600 nm. When the bacterial growth was complete, cell cultures were quickly cooled on ice and centrifuged for 10 minutes at 4°C and 6000 x g in a Sorvall RC5B centrifuge. The harvested cell pellet was then washed by resuspending in chilled TED (20 mM Tris-HCl, pH 7.5 containing 1 mM EDTA and 2 mM DTT) followed by centrifugation as above. The cell pellet was washed and centrifuged once more before re-suspension in TED and extraction by ultrasonication. The washed cell suspension (approximate cell density 30 to 40 mg dry wt. ml^{-1}) was placed in a beaker surrounded by an ice-water mixture and treated with ultrasound for a total period of 2 to 5 minutes, depending on the volume of the suspension, in 15 second treatments with 45 second intervals for cooling. The sonication was performed on a MSE 60W ultrasonic oscillator set at 9 microns. After cell breakage, cell debris was removed by centrifugation for 30 minutes at 4°C, 20 000 x g in a Sorvall RC5B Centrifuge (Dupont).

2.39.2 SDS polyacrylamide gel electrophoresis.

This method from Sambrook *et al* (1989) was used for the separation of secreted proteins. Routine sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) was performed using a discontinuous pH 8.9 system, with Tris-glycine buffer, and 0.1 %, SDS (Sambrook *et al* 1989). A Biorad Protean II vertical minigel system was used for all routine SDS PAGE. The 10 % resolving gel (8 x 7cm x 1 mm) was cast from the following mixture:

4.0 ml of 30 % acrylamide : bis-acrylamide (30 g : 0.8 g)
3.3 ml of 1.5 M Tris-Cl, pH 8.8
100 µl of 10 % (w/v) SDS
100 µl of 10 % (w/v) ammonium persulphate
4 µl of TEMED
4.0 ml of distilled water

The resolving gel was allowed to polymerize at room temperature for at least 45 minutes, before the addition of a 5 % stacking gel. This had the following composition:

0.83 ml of 30 % acrylamide : bis-acrylamide (30 g : 0.8 g)
0.63 ml of 0.5 M Tris-Cl, pH 6.8
50 µl of 10 % (w/v) SDS
50 µl of 10 % (w/v) ammonium persulphate
5 µl of TEMED
3.4 ml of distilled water

Protein samples (approximately 50-100 µg) were mixed with an equal volume of SDS sample buffer (5.0 ml of 0.5 M Tris-Cl, pH 6.8; 2.0 ml of 10 % (w/v) SDS; 4.0 ml of glycerol; 10 mg of bromophenol blue and 9.0 ml of distilled water) containing 10 mM DTT, and heated in a boiling water bath for at least 5 minutes. Pharmacia low molecular weight SDS PAGE marker proteins were used as standards (5 µg of

the standard mixture per gel track). The electrophoresis was performed at room temperature in Tris-glycine running buffer (0.25 M Tris, 1.92 M glycine and 0.1% SDS, pH 8.9) using a constant voltage of 120 V, for an average of 1.5 to 2 hours. After electrophoresis, the gels were either electroblotted on PVDF membranes (see below) or fixed and stained in Coomassie stain (0.05% (w/v) coomassie blue in isopropanol/acetic acid/water (25%:10%:65%). Destaining was performed in 25% isopropanol, 10% acetic acid (v/v). Gels were stored in 10% acetic acid (v/v).

2.39.3 Western blotting.

After SDS PAGE, gels were electroblotted (250 mA, 90 minutes) onto PVDF membranes (Immobilon P, Millipore) at 0-4°C in 25 mM Tris, 192 mM glycine, 0.037% (w/v) SDS, and 10% (v/v) methanol. Membranes were washed briefly with TBS (10 mM Tris-HCl, pH 7.5, 150 mM NaCl) to remove residual transfer buffer, and incubated (overnight at 4°C or 3 hours at room temperature) in 5% (w/v) BSA in TBS containing 0.2% (v/v) Tween 20.

After blocking, membranes were washed in an excess of TBS-Tween for 3 x 10 minutes. Both primary and secondary antibodies were diluted just before use into BSA/TBS-Tween blocking buffer. Primary antibody incubations were performed at room temperature for 2 hours, followed by 3 x 10 minute washes in TBS-Tween. Secondary antibody incubations were performed at room temperature for 90 minutes, followed by 3 x 10 minute washes in TBS-Tween. Antibody cross-reactivity was determined using enhanced chemiluminescence, according to the manufacturer's instructions (Amersham). Autoradiography of the PVDF blots was carried out at room temperature using DuPont Cronex X-ray film. Exposure times were typically between 3 and 10 minutes.

Anti-phosphotyrosine antibody probing was performed using a 1:1000 dilution of anti-phosphotyrosine monoclonal antibodies (Sigma); cross-reactivity was determined using an HRP-conjugated secondary antibody (1:2000 dilution). Anti-TypA IgY antibodies (provided by Dr P. Freestone, University of Leicester) were

used at 1:2000 dilutions; antibody recognition was determined using an HRP-conjugated anti-IgY secondary antibody (Promega) diluted 1:2000 dilution.

Chapter 3

Cloning and sequencing of the *typA* locus

3.1 Introduction

Tyrosine phosphorylation of prokaryotic proteins is a newly-described phenomenon which has now been reported in a number of bacteria and cyanobacteria (Kennelly and Potts 1996). The reason why this type of phosphorylation has only recently been found is unclear. Protein phosphorylation on serine and threonine residues in bacteria has been known for some time and its importance is clearly understood with respect to post-translational modification of proteins (Cozzone 1988). The importance, if any, of tyrosine phosphorylation is not known at present although it appears to play a smaller part in protein phosphorylation overall, perhaps suggesting that this type of phosphorylation may be of greater importance. This is supported by recent data from work done on a tyrosine phosphorylated protein in enteropathogenic *E. coli* derivative MAR001, now known as TypA (Freestone *et al* 1998b) or BipA (Farris *et al* 1998); I shall refer to it as TypA. TypA is tyrosine phosphorylated in EPEC derivative MAR001 (Freestone *et al* 1998b) and an *E. coli* L-form (Freestone *et al* 1998a) but not in *E. coli* K-12 strains (Freestone *et al* 1998b); the reason for this is not known. Mixing extracts of TypA from EPEC MAR001 with extracts of TypA from *E. coli* K-12 does not result in tyrosine phosphorylation of the *E. coli* K-12 TypA (Freestone *et al* 1998b). Furthermore, protein expression is altered in an *E. coli* K-12 *typA* mutant, suggesting a regulatory role for TypA. Interestingly, the expression of the global regulator H-NS is increased, as is the expression of the carbon starvation protein CspA and the universal stress protein UspA. H-NS is an important global regulator and therefore also plays a role in virulence (Atlung and Ingmer 1997). CspA is important in responding to environmental stimuli (Yamanaka *et al* 1998) whilst UspA is important in responding to various stress stimuli including antibiotics (Nyström and Neidhardt 1996; Freestone *et al* 1997). Such conditions may be found in the

gastrointestinal tract (Duncan and Edberg 1995). Although a number of EPEC virulence genes have been described in recent years (Nataro and Kaper 1998), the regulatory aspects of such virulence genes have not yet been determined. A number of environmentally-regulated genes, such as *fur* and *ompR*, have now been well characterised (Mahan 1996). Work in EPEC has found that there are a number of global regulatory genes, such as *hns* (Finlay and Falkow 1997), and also regulatory genes which regulate specific virulence genes and are therefore non-global, such as *perA* (Gómez-Duarte and Kaper 1995). The potential exists that genes such as *typA* may be involved in the regulation of recently-described virulence genes in EPEC.

As discussed in **section 1.13**, the *typA* gene is 1776 bp in length and its sequence in *E. coli* K-12 is known from the published genome project sequence (Plunkett *et al* 1993; Blattner *et al* 1997). The predicted product of *typA* is 65.5 kDa but the protein runs at 80 kDa on SDS-PAGE gel electrophoresis, probably due to a conformational irregularity. Although the *typA* sequence is known from *E. coli* K-12, the differences seen in tyrosine phosphorylation of the gene product between pathogenic and non-pathogenic strains of *E. coli* may be accounted for by differences in the sequence of *typA*. Therefore, characterisation of *typA* is important to determine if there are differences between pathogenic and non-pathogenic strains at the DNA and protein level such as additional tyrosine residues.

The initial hypothesis speculated that the sequence of *typA* differs between EPEC and *E. coli* K-12 thereby accounting for differences seen in tyrosine phosphorylation of TypA between these two strains. Furthermore, it was hypothesised that *typA* plays a role in the regulation of virulence genes on a global or non-global level. Additionally, since many virulence genes occur in operons, it was also reasonable to suggest that genes upstream or downstream of *typA* in EPEC may be further responsible for the differences seen between EPEC and *E. coli* K-12. The availability of the *E. coli* genome sequence (Blattner *et al* 1997), and indeed genome sequences of other bacteria, could enable gene identification and characterisation. This would be particularly useful if novel genes or genes without a known phenotype are found so that comparisons could be made between the novel

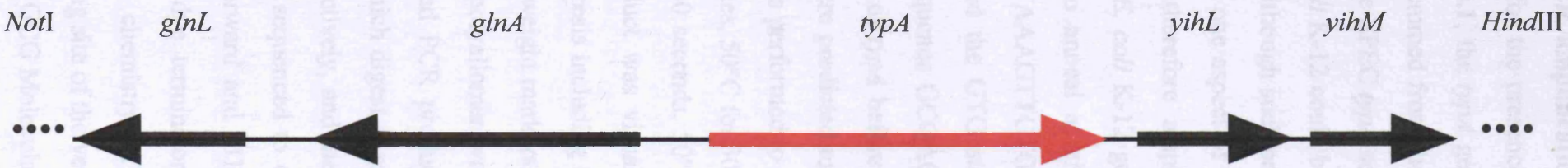
DNA sequence and published sequences. If a potentially novel DNA sequence were found which showed homology to a published sequence then it may provide a clue to its function. To test these hypotheses it would be first necessary to clone and sequence *typA* and the regions upstream and downstream. This chapter describes the rationale for cloning the *typA* gene and the flanking regions from enteropathogenic *E. coli* strain E2348/69 and their subsequent characterisation. Chapters 4 and 5 describe the mutation of *typA*, the construction of an EPEC E2348/69 *typA* mutant and the subsequent phenotypic analysis of this mutant.

3.2 Results

3.2.1 Cloning of the *typA* locus from enteropathogenic *E. coli* strain E2348/69.

The strategy described here for cloning the *typA* gene from enteropathogenic *E. coli* (EPEC) was based upon the published sequence of *E. coli* K-12 (Blattner *et al* 1997) (**figures 3.1 and 3.2**). High quality chromosomal DNA from EPEC strain E2348/69 was prepared using the caesium-chloride gradient method. Two aliquots of 2 µg of chromosomal DNA were digested with the restriction enzymes *Not*I, *Hind*III and *Not*I, *Kpn*I respectively in the appropriate reaction buffers, 0.5 µl of enzyme being used in each case, to create two genomic libraries. According to the *E. coli* K-12 genome sequence (Blattner *et al* 1997), these restriction enzymes digest either side of *typA* resulting in fragments of approximately 8 kb and 9 kb in size (**figures 3.1 and 3.2**). The resultant fragments from the genomic library were cloned into the cloning vector pBluescript SK- by ligating 1 µg of digested chromosomal DNA with 0.5 µg of *Not*I, *Hind*III and *Not*I, *Kpn*I digested pBluescript SK- respectively. The ligation product was transformed into *E. coli* DH5α by electroporation and transformants selected on LBA containing 100 µg/ml ampicillin and 40 µg/ml X-gal. Approximately 60% of the transformants that grew on this medium were white indicating inactivation of the *lacZ* gene due to the presence of an insert; the remaining colonies were blue indicating the absence of an insert.

Figure 3.1 Diagrammatic representation of the *typA* locus in *E. coli* K-12 according to the *E. coli* K-12 genome sequence (Blattner *et al* 1997). The diagram indicates the *Not*I, *Hind*III and *Kpn*I sites used for attempts at cloning the *typA* gene from EPEC E2348/69.



Function:

glnL	negative regulator of *glnA*	*glnA*	glutamine synthesis	*typA*	formerly o591, function unknown
yihL	formerly o236, temporary name	*yihM*	formerly o326, temporary name		

Figure 3.2 Diagrammatic representation of the *Not*I-*Hind*III *typA* chromosomal locus of *E. coli* K-12. Open-reading frames within the *Not*I-*Hind*III locus which contains *typA* are indicated. Information was obtained from the *E. coli* K-12 genomic sequence (Blattner *et al* 1997) and the *E. coli* K-12 genomic traditional genetic and physical maps (Berlyn 1998; Rudd 1998). The *typA* gene is highlighted in red. Note that the *Not*I-*Hind*III locus does not include the complete open-reading frames of *glnL* and *yihM*.

3.2.1.1 PCR amplification and cloning of *typA* from *E. coli* K-12. To enable screening for the presence of the EPEC *typA* in the transformants obtained from **section 3.2.1**, the *typA* gene from *E. coli* K-12 was cloned. The necessity for this strategy stemmed from the fact that the *E. coli* K-12 *typA* sequence was published whereas the EPEC *typA* sequence was not. Therefore amplification of the *typA* gene from *E. coli* K-12 could be achieved after the design of appropriate oligonucleotide primers. Although such primers could be designed blindly for amplifying the EPEC *typA* gene, one aspect of this project was to also clone the region downstream of *typA* and therefore amplification of *typA* alone was not justified. Using the published *E. coli* K-12 genome sequence (Blattner *et al* 1997), two primers were designed to anneal at either end of *typA*. The first, MIR1A (primer sequence CTGATTTAAAGTTGTGATCGAA), was designed to anneal at the 5' end and incorporated the GTG start codon of *typA* and a *Dra*I site. The second, MIR2 (primer sequence GCGAGGATCCTTTTCCGTCAGT), was designed to anneal at the 3' end of *typA* before the terminator loop and incorporated a *Bam*HI site. The primers were predicted to amplify a product of 1805 bp. Optimisation of the PCR cycling was performed to achieve a single clean band, which was one cycle at 96°C for 2 minutes, 50°C for 30 seconds and 72°C for 2 minutes followed by 36 cycles at 96°C for 30 seconds, 50°C for 30 seconds and 72°C for 2 minutes. The resultant DNA product was visualised at approximately 1800 bp using 2% agarose gel electrophoresis including both λ*Hind*III and ϕX174 RF DNA/*Hae*III standard DNA molecular weight markers. The 1800 bp band was excised from the gel and purified using the polyallomer wool/ QIAGEN QIAquick™ PCR purification kit method. The isolated PCR product was digested with the restriction enzymes *Dra*I and *BamH*I, which digest at restriction sites that occurred upstream and downstream of *typA* respectively, and then cloned into the *Sma*I and *BamH*I sites of pUC19. The insert was sequenced to confirm the cloning of the *typA* fragment with the M13 (16mer) forward and M13 (24mer) reverse primers using the Perkin-Elmer ABI PRISM™ dye terminator cycle sequencing system with the dichlororhodamine terminator chemistry. These M13 primers anneal at sites adjacent to the multicloning site of the vector allowing for cycle sequencing of the insert. Using the Wisconsin GCG Molecular Biology Package version 9.1-UNIX (Genetics Computer

Group, Madison, Wisconsin) and NCBI BLAST (Altschul *et al* 1997) programs the insert was confirmed as corresponding to *E. coli* K-12 *typA*. The recombinant plasmid was designated pSCC-T.

3.2.1.2 Colony blot preparation and hybridisation of EPEC genomic libraries using a *typA* specific probe. A total of 200 white colonies, corresponding to 100 from the *NotI/Hind*III genomic library and 100 from the *NotI/Kpn*I genomic library, were selected from the transformation in **section 3.2.1**. They were patch-plated onto LBA containing 100 μg/ml ampicillin and onto an Amersham Hybond™ N+ gridded nylon membrane placed onto LBA containing 100 μg/ml of ampicillin. The former plate was stored at 4°C for future reference. The transformants were then screened by colony blot analysis using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system as described in the Materials and Methods. The hybridisation probe was prepared by digesting approximately 2 μg of pSCC-T with *EcoR*I and *BamH*I, two sites which existed within the multi-cloning site of pUC19, and the resultant fragments separated by gel electrophoresis using 1% agarose. The 1.8 kb band, corresponding to the *typA* gene, was removed and purified from the agarose gel using the polyallomer wool/QIAGEN QIAquick™ PCR purification kit method. The probe was then labelled, using the Amersham Life Science Gene Images random prime labelling module, as described in the Materials and Methods. Approximately 10 ng of labelled *typA* probe was used for hybridisation overnight at 65°C. Unhybridised probe was removed by washing with high stringency buffers. One transformant hybridised with the *typA* specific probe corresponding to a recombinant clone from the *NotI/Hind*III genomic library. The corresponding duplicate was isolated from the stored LBA plate. This recombinant plasmid was designated pSCC01.

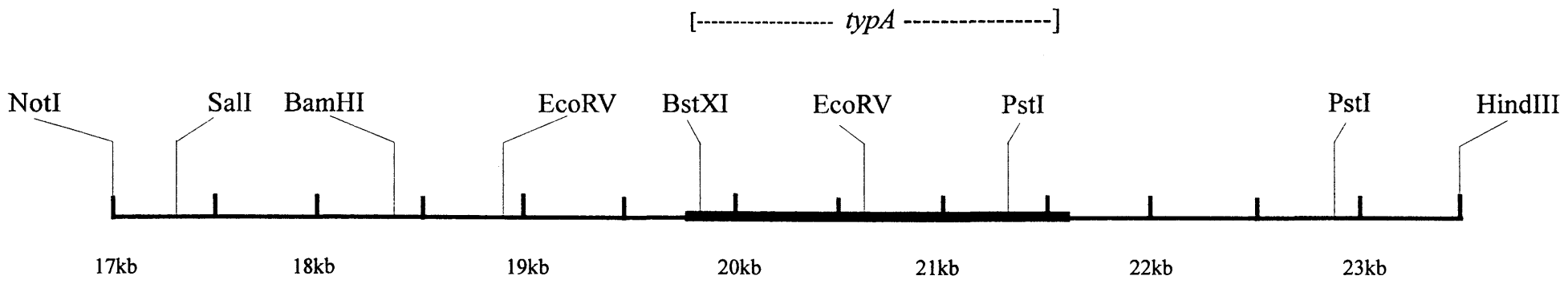
3.2.2 Restriction enzyme analysis of pSCC01.

Restriction analysis is a useful tool for determining the presence of restriction sites and for indicating the presence and size of a DNA fragment (Daniels 1996) although it does not confirm the presence of the correctly cloned fragment. The reasons for

performing the restriction analysis of pSCC01 were two-fold. Firstly, the colony blot analysis of putative *typA* recombinant clones was a screening method requiring further analysis to putatively establish the identity of this clone. Secondly, further sub-cloning and mutational analysis of *typA* will be required therefore requiring identification of suitable restriction sites. Restriction analysis was considered to be a satisfactory interim method for these requirements. Using the published *E. coli* K-12 genome sequence (Blattner *et al* 1997), commonly-used restriction enzymes were chosen for the analysis of pSCC01. Their choice was determined by the maximum amount of information that could be gained from their use as single enzymes or in combination. The restriction enzymes and combinations chosen were as follows:

*Bst*XI	*Dra*I	*Eco*RV	*Hinc*II
*Sal*I	*Xho*I	*Not*I/*Asp*700I	*Not*I/*Bst*XI
*Hind*III/*Eco*RV	*Hind*III/*Hae*II	*Not*I/*Hind*III	*Hind*III/*Xho*I
*Not*I/*Hind*III/*Dra*I			

DNA purified by the alkaline lysis mini-preparation method was digested with the above enzymes, as described in the Materials and Methods, and separated on a 1% gel by electrophoresis. The sizes of the resultant fragments from each restriction digest were calculated and a restriction map of the pSCC01 insert constructed (**figure 3.3**). These restriction digests indicated that the correct fragment was obtained and indicated that it was smaller than predicted at approximately 6.5 kb in size. It also indicated that many of the restriction sites present in this region in *E. coli* K-12 were also present at the same position in the corresponding region in EPEC E2348/69. However, a number of differences were noted. The *Asp*700I site was at 19.1 kb in EPEC E2348/69 compared to 17.4 kb in *E. coli* K-12 and the *Hind*III site in EPEC E2348/69 was at 23.8 kb compared to 25.1 kb in *E. coli* K-12. Also, the *Xho*I site which was present in *E. coli* K-12 at 22.2 kb was absent in EPEC E2348/69. It was also suspected that the *Hae*II and *Dra*I sites were different in EPEC E2348/69 compared to *E. coli* K-12 due to the resultant number and size of



Fragment size approximately 6.5kb

Figure 3.3 Restriction map of the *typA* locus from EPEC E2348/69. The map indicates restriction sites which would be useful for sub-cloning and mutation analysis. The locus is at 88 minutes in the EPEC chromosome; the kb markings relate to the position of the locus with relation to this region of the chromosome. The map was mostly based on data from the *E. coli* K-12 genome sequence (Blattner *et al* 1997).

the fragments produced, although they were too numerous for accurate analysis by this method.

3.2.3 Creation of sub-clones of pSCC01 by nested deletions

A library of sub-clones can be created from a recombinant plasmid containing a large insert (Henikoff 1984). Using the Pharmacia Nested Deletion kit, sub-cloning of the entire insert of pSCC01 was attempted so that each sub-clone would possess a deletion of approximately 500 bp. This was performed by digesting one end of the fragment with *Not*I/*SSt*I in reaction buffer 2 and the other with *Hind*III/*Kpn*I in reaction buffer 4; the rest of the method was followed according to the manufacturers instructions. The potential sub-clone library DNA of 14 samples were run on an 1% agarose gel by electrophoresis to check that the reactions had performed. However, after attempts with two different kits, the reactions resulted in complete degradation of the plasmid or no exonuclease III digestion at all. Time points and enzyme quantity were varied to reduce degradation time without success.

3.2.4 Automated sequencing of pSCC01 by primer walking.

Various methods exist for the sequencing of DNA, such as the chemical cleavage method devised by Maxam and Gilbert (1977) and the dideoxy nucleoside triphosphate chain-terminating method of Sanger *et al* (1977). There are also variations of the Sanger method such as the radioisotope labelling method (Evans and Read 1992) and the chemiluminescent detection system (Beck *et al* 1989, Tizard *et al* 1990). More recently, automated sequencing has become available which is a variation of the Sanger method but where the dideoxynucletides are labelled with a fluorescent dye (Ansorge *et al* 1987, Connell *et al* 1987, Middendorf *et al* 1992). Automation of this method has led to its wide use for single-stranded and double-stranded DNA sequencing (Anon 1998).

Although the nested deletion method of creating sub-clones was not successful, other strategies exist to enable rapid sequencing of DNA including sub-cloning,

transposon-based sequencing and primer-walking so that the above sequencing methods may be used (Anon 1998). Sub-cloning of the insert of a recombinant clone into a common vector plasmid may be carried out and then sequencing performed using M13 forward and M13 reverse primers and this strategy is ideal for large inserts. Transposon-based sequencing is also ideal for large cloned inserts and has the advantage of possessing an antibiotic resistance marker for clone selection but also the disadvantage that transposons may insert into the chromosome randomly and may not insert into the sequence of interest; however, this strategy is not useful for small inserts. Alternatively, a potentially rapid but also more expensive method is primer walking. This method is established for sequencing recombinant clone inserts of moderate length where sub-cloning is not warranted (Giesecke *et al* 1992).

3.2.4.1 Automated sequencing of regions flanking *typA*. Initially, two autosequencing reactions were performed with the M13 (16mer) forward and M13 (24mer) reverse primers using the Perkin-Elmer ABI PRISM™ system. Using this strategy, sequence data were obtained from either end of the cloned *typA* fragment. The sequence from the 5' end of the cloned insert indicated that the region shared 92% amino acid sequence identity with the same region in the *E. coli* K-12 chromosome. This region contains the *glnL* gene (**figure 3.4**). The sequence of the 3' end of the cloned insert, corresponding to data obtained using the M13 forward primer, showed no amino acid sequence identity with the same region in *E. coli* K-12. Instead, it possessed 57% sequence identity over 271 amino acids with ORF *o300* at 35.4-36.3 kb in the 88 minute region of the *E. coli* K-12 chromosome. The product of ORF *o300* is similar to members of the phosphofructokinase B (*pfkB*) family of carbohydrate kinases.

As the 5' end of the cloned EPEC insert possessed 93% amino acid sequence identity to the corresponding *E. coli* K-12 sequence and previous work showed that the sequence directly upstream of *typA* was identical to that in *E. coli* K-12 (Trinei M, personal communication), sequencing of the full length of the region upstream of *typA* was not considered to be necessary. Instead, the 3' region downstream of *typA*, was considered to be of greater interest and this region was therefore subcloned as

Figure 3.4 Nucleotide and predicted amino acid sequences of the upstream end of the *Not*I-*Hind*III pSCC01 fragment. The fragment was sequenced with the M13 reverse primer, edited as necessary and translated using the Wisconsin GCG molecular biology program. The predicted amino acid sequence was entered into the BLAST web-site and homologous proteins searched for within the GenBank databases. The sequence possessed 92% identity with the *glnL* gene of *E. coli* K-12 indicating that this chromosomal position was very similar in both EPEC E2348/69 and *E. coli* K-12 (| indicates identical amino acids whilst + represents similar amino acids).

Nucleotide sequence

```
  1   ACCGCGGTGG CGGGCGCTGA CCATCGGGTA AAACAGCGTA TCCTGCAAAT
 51   GAGGCGGAAT GCCCGGCCCG TTATCTTCCA CATCAATCCG CGCCGCCAGC
101   CGATAGCGCT CGCCGTGTAA GGTCAGTTGA AACGCTGTGC GGGTACGCAG
151   AATGATTTCA CCACCCTCCG GCCCCAGCGC CTGTAGCGCA TTGCGCACAA
201   TATTCAGCAA GACCTGTTCA ATTTGATCCG GGTCGTGCGC CAGTTCCGGC
251   AGGCTGGGGT CGTAATCCCG AATCAACCGC ACGTTGTTCG GCAGTTCCAT
301   CGACACCATC GTCACCACGC GTTCAGCCAC TTTGTGAATA CTTTCGGTAA
351   TGCGCGTATC GGGCAGCTGC GGCCCCAACA GACGGTCGAC CATATTTCGC
401   AACCGGTCCG CCTGTTCGAT AATCACTTTT GGTATATCCA A
```

Predicted amino acid sequence

```
  1   GRHRXSVMPY FLTDQLHXPI GPGNXEVDIR AALRYREGHL TLQFATRTRL
 51   IIEGGEPGLA QLANRVINLL VQEIQDPDHA XEPXSPDYDR ILRVNNPLEM
101   SVMXVVREAV KHISETIRTD PLQPGLLRDV MNRLRDAQEI IVKPIDL
```

Amino acid sequence homology

```
pSCC01:       LDIPKVI IEQADRLRNM VDRLLGPQLP DTRITESIHK VAERVVTMVS 101
              |+  ||| ||||||||||+ ||||||||||  ||+|||||| |||||||+||
K-12 glnL:    LEYTKVI IEQADRLRNL VDRLLGPQLP GTRVTESIHK VAERVVTLVS 204

          MELPNNVRLI RDYDPSLPEL AHDPDQIEQV LLNIVRNALQ ALGPEGGEII  51
          ||||+|||||  ||||||||||  ||||||||||  ||||||||||  ||||||||||
          MELPDNVRLI RDYDPSLPEL AHDPDQIEQV LLNIVRNALQ ALGPEGGEII 154

          LRTRTAFQLT LHGERYRLAA RIDVEDNGPG IPPHLQDTLF YPMVSARHRG   1
          ||||||||||  ||||||||||  ||||||||||  ||||||||||  |||||  |  |
          LRTRTAFQLT LHGERYRLAA RIDVEDNGPG IPPHLQDTLF YPMVSGREGG 104
```

two inserts into pUC18. This was achieved by digestion of 1 μg of pSCC01 with *Pst*I and *Hind*III to give predicted fragment sizes of 4800, 1300 and 400 bp. The 1300 bp fragment possessed *Pst*I-digested ends, whereas the 400 bp fragment possessed a 5' *Pst*I end and a 3' *Hind*III end. These fragments were isolated by 2% agarose gel electrophoresis and purified by the polyallomer wool/ QIAGEN QIAquick™ PCR purification kit method. The purified fragments were lighted into *Pst*I and *Pst*I-*Hind*III-digested pUC18 respectively and the ligation products transformed into DH5α by electroporation. Transformants were selected on LBA containing 100 μg/ml ampicillin and 40 μg/ml X-gal. Approximately 70% of the transformants that grew on this medium were white. The presence of the correct insert in the respective transformants was confirmed by the analysis of a dozen transformants in each case by alkaline lysis mini-preparation of plasmid DNA followed by restriction analysis with *Pst*I for the putative 1300 bp clone and *Pst*I-*Hind*III for the putative 400 bp clone. Correct clones were found after restriction analysis and were designated pSCC05 and pSCC06 respectively. Automated sequencing of the inserts of pSCC05 and pSCC06 was performed using the M13 forward and reverse primers with the Perkin-Elmer ABI PRISM™ system. Sequencing data indicated that pSCC05 possessed an insert of 1313 bp (**figure 3.5**) whilst pSCC06 possessed an insert of 422 bp (**figure 3.6**). The 1.3 kb downstream of the *typA* RNA terminator loop, contained within pSCC05, possessed no homology with the same region in *E. coli* K-12 (**figure 3.7**). In fact, it was found that this region was entirely novel after searching the BLAST sequence databases. The sequence of pSCC06 corresponded to homology with ORF *o300* at 35.4-36.3 kb in the 88 minute region of the *E. coli* K-12 chromosome, as stated above. In summary the complete region after *typA* is probably novel because the complete open reading frames have not yet been cloned. The homology shared with the last 422 bp of this region is probably not genuine because it only shows homology with another putative open reading frame. The overall %GC content of the novel region is 50.3% thereby suggesting that the region may have originated from *E. coli*; however, there is also the possibility that the DNA was acquired from another organism with a similar %GC content.

Figure 3.5 Nucleotide and predicted amino acid sequences of the 1.3 kb *Pst*I fragment downstream of *typA*. The *Pst*I fragment was sub-cloned as described in **section 3.2.4.1**. The recombinant clone pSCC05 was then sequenced with M13 forward and reverse primers using the Perkin Elmer ABI PRISM™ model 377 system. The sequence was edited as necessary using the Wisconsin GCG molecular biology program, translated and then homology with sequences in the GenBank and TIGR determined using the BLAST program (Altschul 1997). The nucleotide sequence shows the stop codon of *typA* in bold and the end of homology between EPEC and *E. coli* K-12 after *typA* in bold underlined. The first 61 amino acids of the pSCC05 insert corresponded to the last 61 amino acids of EPEC *typA* as expected. Predicted amino acid identity after the end of *typA* reduced dramatically (see **figure 3.7**). The predicted amino acid sequence for the remainder of the pSCC05 insert possessed no homology with any sequence in the GenBank or TIGR databases thereby suggesting that this region is novel.

Nucleotide sequence

```
   1  GACCTGACTG TAAACTGCCT GACCGGTAAG AAACTGACCA ACATGCGTGC
  51  TTCCGGTACT GACGAAGCCG TTGTTCTGGT TCCGCCTATC CGCATGACTC
 101  TGGAACAAGC TCTGGAATTC ATCCATGATG ACAAACTGGT AGAAGTGACT
 151  CCGATCTCTA TCCGTATTCG TTAACGTCAC CTGACGGAAA ACGATCGTCG
 201  CCGCGCCAAC CGCGCACCGA AAGACGATTA ATTTCGTTTT TTAGTCGTAA
 251  AAAACCTGCC AGCGATGGCA GGTTTTTTTT GGCTTATGGA ATAGAATATT
 301  TGCGCATTTA CCTTTCGGAT TTCCGAAATT TCAATTCCGG ATGCTTCTTG
 351  GCTCGGGAAA CACTATTGCA AATTCTCTTC CTTAAAACCG GCAAAAAAAT
 401  TCGGAGGCTG GCGCCTTTAT AAACCTAACT TACGTTCATT ATCATTTGAA
 451  TTTGCCCATA ACCGGTTAAA TCTTATGATT GGCCTAACCA TTTTTTCCTC
 501  ACCTTTTTAT CCTCCTGTCC CCATATATTC CTCGAAGGGG GCCCGCGGGA
 551  GGGGGGAATA TATAAAGGAG GCGCGGGATA AAAAAGGGGG GTTATAAGAG
 601  GGGAAGGGTT AAATAAGAGA GGGCGGATGG CCTATAAGAG GTGGATTTGG
 651  CCACGGCTTA TTTATATGGA TAAGAGGATA TAGGATAATA TTATACCCGC
 701  GTTGGCTTTA CTAAAAAGGG TGTTTATTTT TAGAGGGAAG ACATGGGAGC
 751  AGAAGAGCCC GAGACCCCAT GGCCGAATGA AGGACGGGAA ATATTCAAAC
 801  CCATTGTTAA AAAGGGGAAG GTTAGGATAT CCCAGGGGGA ACGGTTTAGA
 851  GTAAGAAACA TGCGAGCCGT GATGCCAAAT CTTTCAAATA TTCAAAAAAC
 901  TGGGTTGTGT GATTAGAATA AGGAGATGTC ATGGCGTGGA ATACCGGTCA
 951  GTGGGATGAT ATTCGTGAAT AGTCCGTGGC CCATCGTGGG TAGATGTTTA
1001  CCATGCTACG CATAAAGTGA GAAAAATGAT TCGGTTTGCT AATGATTAGG
1051  CATGCCGGCC CGACCGACAG GTAGCGTACA TTTAATTGCG GCAACCGCAC
1101  TGGCAAAAAG GGTAGCCTTG CATTGCTGGC ATTTTTTATG CCAACGCTAC
1151  CGGGAGTGCG CCGTGAAAAA CGTCGCCTGC TCCAGTGGTA TCAACCACAT
1201  TGACAGTCAA TGCCTCCTGT TGACACAAAT GACCATCTTC AATCCACAAC
1251  GATCCTTCAG GACCCGGCGT AACATACACT TTACCTGCGG AAAGTGTAGT
1301  TGTTTTGAAC AAAC
```

Predicted amino acid sequence

```
   1  DLTVNCLTGK KLTNMRASGT DEAVVLVPLI RMTLEQALEF IHDDKLVEVT
  51  PISIRIR*RH *TEKDLRRAN RPPKDINFVF LIVKNLPPMA GFFWLMENKI
 101  P*YSFRISEI SIPDASWLGK HYCKFSSLKP AKKFGGWRLY KPNLRSLSFE
 151  FAHNRLNLMI GLTIFSSPFY PPVPIYSSKG ARGRGEYIKE ARDKKGGL*E
 201  GKG-IREGGW PIRGGFGHGL FIWIRGYRII LYPRWLY*KG CLFLEGRHGS
 251  RRARDPMAE* RTGNIQTHC* KGEG*DIPGG TV*SKKHASR DAKSFKYSKN
 301  WVV*LE*GDV MAWNTGQWDD IRE*SVAHRG *MFTMLRIK* EK*FGLLMIR
 351  HAGPTDR*RT FNCGNRTGKK GSLALLAFFM PTLPGVRREK RRLLQWYQPH
 401  *QSMPPVDTN DHLQSTTILQ DPA*HTLYLR KV*LF*TN
```

* indicates a stop codon

Figure 3.6 Predicted amino acid sequence of the 422bp *Pst*I-*Hind*III fragment downstream of *typA*. The *Pst*I-*Hind*III fragment was sub-cloned as described in **section 3.2.4.1**. The recombinant clone pSCC06 was then sequenced with M13 forward and reverse primers using the Perkin Elmer ABI PRISM™ model 377 system. The sequence was edited as necessary using the Wisconsin GCG molecular biology program, translated and then homology with sequences in GenBank determined using the BLAST program (Altschul 1997). The region possessed 72% identity with with ORF o300 at 35.4-36.3 kb in the 88 minute region of the *E. coli* K-12 chromosome (| indicates identical amino acids whilst + represents similar amino acids).

Predicted amino acid sequence

```
1   GRVGDDSCGN TLLAELEGWG VNTAFCRRYP NARSSQSAIL

41  VDQHGERIIV NYPSPDLGTD AEWLETIDFS RYDLILADVS

81  LA*RY*KSIL ASAPCGCDNA FRRRYDTAGY FSTCRACRSR

121 RVLNTKAETH DWPAGRL*KI PGY
```

Amino acid sequence homology

```
pSCC06      1 GRVGDDSCGN TLLAELEGWG VNTAFCRRYP NARSSQSAIL
              |||||| || +|||||| || ||| + +||   |+|||||+
ORF o300   61 GRVGDDDTGN SLLAELESWG VNTRYTKRYN QAKSSQSAIM

           41 VDQHGERIIV NYPSPDLGTD AEWLETIDFS RYDLILADV
              ||  |||||+ |||||||  | ||||| |||| ++|++||||
          101 VDTKGERIII NYPSPDLLPD AEWLEEIDFS QWDVVLADV
```

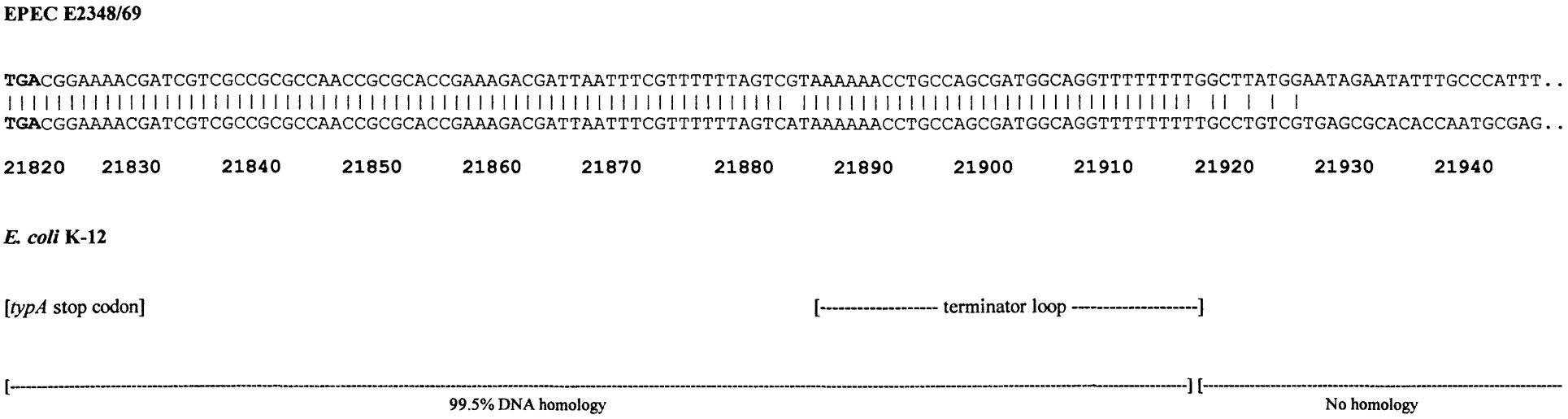


Figure 3.7 DNA homology between EPEC E2348/69 and *E. coli* K-12 at the end of *typA*. The DNA sequences of EPEC E2348/69 *typA* and *E. coli* K-12 *typA* were compared using the Wisconsin GCG molecular biology program. This line-up indicates that nucleotide homology between EPEC E2348/69 and *E. coli* K-12 ends after the *typA* terminator loop.

3.2.4.2 Automated sequencing of *typA*. Seven primers were designed, as indicated in **table 3.1** and **figure 3.8**, to enable complete sequencing of *typA*. These primers were designed based on the published *E. coli* K-12 sequence (Blattner *et al* 1997) but synthesised during primer walking as sequence data was gained. Again the ABI PRISM™ sequencing system was used. Sequencing of *typA* was completed using each primer at least twice so that any sequencing errors could be rectified. The determined DNA sequence and predicted amino acid sequences are shown in **figures 3.9** and **3.10** respectively. Sequencing indicated that the EPEC E2348/69 *typA* gene possessed 99.9% similarity and 98.8% identity with the *E. coli* K-12 *typA* gene at the DNA level (**figure 3.11**). The sequence differed in 18 nucleotides. At the predicted amino acid level there was 99.2% similarity and 98.8% identity (**figure 3.12**) and the differences are shown in **table 3.2**. The nucleotide differences result in six amino acid differences although there was no additional tyrosine residue to account for the tyrosine phosphorylation of *typA* which is observed in EPEC. In addition, there were no alteration in amino acid sequence around the nine existing tyrosine residues. However, there was a premature stop codon that shortened TypA by three amino acids (**figure 3.10**). A premature stop codon could affect the secondary structure of the protein and also expose amino acid residues leading to phosphorylation. To determine whether the amino acid differences between EPEC E2348/69 TypA and *E. coli* K-12 TypA had any significance in protein structure, hydropathy plots were performed using the method of Kyte and Doolittle (1982) (**figure 3.13**), and secondary structure analysis performed using the NNPredict web-site program (**figure 3.14**). Hydropathy profile alignment has been shown to be a valuable tool in protein analysis because the technique compares predicted structural rather than sequence similarities (Lolkema and Slotboom 1998). The hydropathy plots indicated that there are a few differences in the *C*-terminus of EPEC TypA resulting in it being slightly more hydrophobic than that of the *E. coli* K-12 TypA. However, the hydropathy plot predicted no major differences between the proteins. The EPEC E2348/69 *typA* sequence was deposited within the Genbank database, accession number AF058333.

Table 3.1 Nucleotide sequences of the primers used for sequencing of EPEC E2348/69 *typA*. Primers were designed to anneal at various positions within *typA* to enable full and repetitive sequencing of *typA* as indicated in **figure 3.9**.

Primer name	**Primer sequence (5'-3')**
MIR1A	CTG ATT TAA AGT TGT GAT CGA A
MIR12	CAG GTA TTC GAT CTG TTC
SCC5	CGG CCT GCA ACT GGA SCT GCC
SCC6	GCG ATA TCG TTG CGA TCA CGG G
MIR13	CAG AAC GGC GAG GTG TT
SCC1	CAC CTC GCC GTT CTG CG
MIR2	GCG AGG ATC CTT TTC CGT CAG T

Figure 3.8 DNA sequence of *E. coli* K-12 *typA* indicating the position of sequencing primers. The primers indicated on the *E. coli typA* sequence were used for sequencing of the EPEC E2348/69 *typA* gene. The oligonucleotide sequence of each primer is shown in **table 3.1**.

MIR1A→

```
   1  GTGATCGAAA AATTGCGTAA TATCGCCATC ATCGCGCACG TAGACCATGG
  51  TAAAACCACC CTGGTAGACA AGCTGCTCCA ACAATCCGGT ACGTTCGACT
 101  CTCGTGCCGA AACCCAAGAG CGCGTGATGG ACTCCAACGA TTTGGAGAAA
 151  GAGCGTGGGA TTACCATCCT CGCGAAAAAC ACCGCTATCA AATGGAATGA
 201  TTACCGTATC AACATCGTTG ATACCCCGGG GCACGCCGAC TTCGGTGGTG
 251  AAGTTGAACG TGTAATGTCC ATGGTAGACT CAGTGCTGCT GGTGGTTGAC
 301  GCATTTGACG GCCCGATGCC GCAAACGCGC TTCGTAACCA AAAAAGCGTT
 351  TGCTTACGGC CTGAAGCCGA TTGTTGTTAT CAACAAAGTT GACCGCCCTG
 401  GCGCGCGTCC TGATTGGGTT GTGGATCAGG TATTCGATCT GTTCGTTAAC
                                        MIR12→
 451  CTCGACGCGA CCGACGAGCA GCTGGACTTC CCGATCGTTT ACGCTTCTGC
 501  GCTGAACGGT ATCGCGGGTC TGGACCACGA AGATATGGCG GAAGACATGA
                                    ←SCC5
 551  CCCCGCTGTA CCAGGCGATT GTTGACCACG TTCCTGCGCC GGACGTTGAC
 601  CTTGACGGTC CGTTCCAGAT GCAGATTTCT CAGCTCGATT ACAACAGCTA
 651  TGTTGGCGTT ATCGGCATTG GSCGCATCAA GCGCGGTAAA GTGAAGCCGA
 701  ACCAGCAGGT CACTATCATC GATAGCGAAG GCAAAACCCG CAACGCGAAA
 751  GTCGGTAAAG TGCTGGGCCA CCTCGGTCTG GAACGTATCG AAACCGATCT
          SCC6→
 801  GGCGGAAGCT GGCGATATCG TTGCGATCAC GGGCCTTGGC GAACTGAACA
 851  TTTCTGACAC CGTTTGCGAC ACGCAAAACG TTGAAGCGCT GCCGGCACTC
                ←MIR13                  SCC1→
 901  TCCGTTGATG AGCCGACCGT TTCTATGTTC TTCTGCGTTA ACACCTCGCC
 951  GTTCTGCGGT AAAGAAGGTA AGTTCGTAAC GTCTCGTCAG ATCCTGGATC
1001  GTCTGAACAA AGAACTGGTA CACAACGTTG CGCTGCGCGT AGAAGAAACC
1051  GAAGACGCCG ATGCGTTCCG CGTTTCTGGT CGTGGCGAAC TGCACCTGTC
1101  TGTTCTGATC GAAAACATGC GTCGTGAAGG TTTCGAACTG GCGGTATCCC
1151  GTCCGAAAGT TATCTTCCGT GAAATCGACG GTCGTAAACA AGAGCCGTAT
1201  GAAAACGTGA CTCTGGACGT TGAAGAACAG CATCAGGGTT CTGTAATGCA
1251  GGCGCTGGGC GAACGTAAAG GCGACCTGAA AAACATGAAT CCAGACGGTA
1301  AAGGCCGCGT ACGTCTCGAC TACGTGATCC CAAGCCGTGG TCTGATTGGC
1351  TTCCGTTCTG AGTTCATGAC CATGACTTCC GGTACTGGTC TGCTGTACTC
1401  CACCTTCAGC CACTACGACG ACGTACGTCC GGGTGAAGTG GGTCAGCGTC
1451  AGAACGGCGT ACTGATCTCT AACGGTCAGG GTAAAGCGGT CGCGTTCGCG
1501  CTGTTCGGTC TGCAGGATCG CGGTAAGCTG TTCCTCGGTC ACGGTGCAGA
1551  AGTTTACGAA GGTCAGATTA TCGGTATTCA TAGCCGCTCT AACGACCTGA
1601  CTGTAAACTG CCTGACCGGT AAGAAACTGA CCAACATGCG TGCTTCCGGT
1651  ACTGACGAAG CCGTTGTTCT GGTTCCGCCT ATCCGCATGA CTCTGGAACA
1701  AGCTCTGGAG TTCATCGATG ATGACGAACT GGTAGAAGTG ACTCCGACCT
1751  CTATCCGTAT TCGTAAACGT CACTGA
                              ←MIR2
```

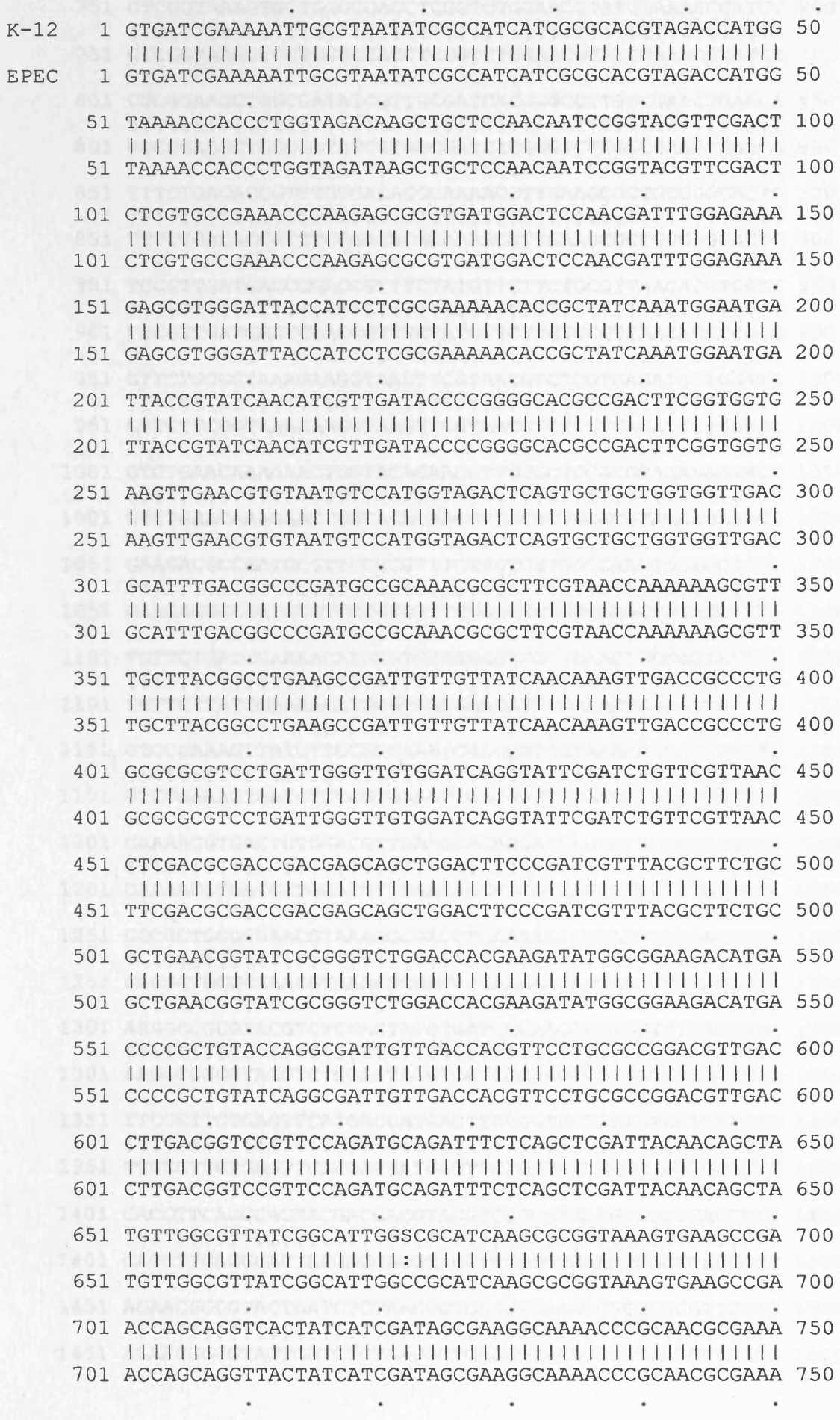
Figure 3.9 The entire DNA sequence of EPEC E2348/69 *typA*. The *typA* gene was cloned into pBluescript on a *Not*I-*Hind*III fragment from EPEC E2348/69 chromosomal DNA. TypA was then sequenced with primers designed to anneal at various locations within *typA* (see **figure 3.8** and **table 3.1**) using the Perkin Elmer ABI PRISM™ model 377 system. The sequence was edited as necessary using the Wisconsin GCG molecular biology program and was deposited within the Genbank database, accession number AF058333. Start and stop codons are in bold and underlined.

```
   1  GTGATCGAAA AATTGCGTAA TATCGCCATC ATCGCGCACG TAGACCATGG
  51  TAAAACCACC CTGGTAGATA AGCTGCTCCA ACAATCCGGT ACGTTCGACT
 101  CTCGTGCCGA AACCCAAGAG CGCGTGATGG ACTCCAACGA TTTGGAGAAA
 151  GAGCGTGGGA TTACCATCCT CGCGAAAAAC ACCGCTATCA AATGGAATGA
 201  TTACCGTATC AACATCGTTG ATACCCCGGG GCACGCCGAC TTCGGTGGTG
 251  AAGTTGAACG TGTAATGTCC ATGGTAGACT CAGTGCTGCT GGTGGTTGAC
 301  GCATTTGACG GCCCGATGCC GCAAACGCGC TTCGTAACCA AAAAAGCGTT
 351  TGCTTACGGC CTGAAGCCGA TTGTTGTTAT CAACAAAGTT GACCGCCCTG
 401  GCGCGCGTCC TGATTGGGTT GTGGATCAGG TATTCGATCT GTTCGTTAAC
 451  TTCGACGCGA CCGACGAGCA GCTGGACTTC CCGATCGTTT ACGCTTCTGC
 501  GCTGAACGGT ATCGCGGGTC TGGACCACGA AGATATGGCG GAAGACATGA
 551  CCCCGCTGTA TCAGGCGATT GTTGACCACG TTCCTGCGCC GGACGTTGAC
 601  CTTGACGGTC CGTTCCAGAT GCAGATTTCT CAGCTCGATT ACAACAGCTA
 651  TGTTGGCGTT ATCGGCATTG GCCGCATCAA GCGCGGTAAA GTGAAGCCGA
 701  ACCAGCAGGT TACTATCATC GATAGCGAAG GCAAAACCCG CAACGCGAAA
 751  GTCGGTAAAG TGCTGGGCCA CCTCGGTCTG GAACGTATCG AAACCGATCT
 801  GGCGGAAGCT GGCGATATCG TGGCGATCAC GGGTCTTGGC GAACTGAACA
 851  TTTCTGACAC CGTTTGCGAC ACGCAAAACG TTGAAGCGCT GCCGGCACTC
 901  TCCGTTGATG AGCCGACCGT TTCTATGTTC TTCTGCGTTA ACACCTCGCC
 951  GTTCTGCGGT AAAGAAGGTA AGTTCGTAAC GTCTCGTCAG ATCCTGGATC
1001  GTCTGAACAA AGAACTGGTA CACAACGTTG CGCTGCGCGT AGAAGAAACC
1051  GAAGACGCCG ATGCGTTCCG CGTTTCTGGT CGTGGCGAAC TGCACCTGTC
1101  TGTTCTTATC GAAAACATGC GTCGTGAAGG TTTCGAACTG GCGGTATCCC
1151  GTCCGAAAGT AATCTTCCGT GAAATCGACG GTCGTAAACA AGAGCCGTAT
1201  GAAAACGTGA CGCTGGACGT TGAACAACAG CATCAGGGTT CTGTGATGCA
1251  GGCGCTGGGC GAACGTAAAG GCGACCTGAA AAACATGAAT CCAGACGGTA
1301  AAGGCCGCGT ACGTCTCGAC TACGTGATCC CAAGCCGTGG TCTGATTGGC
1351  TTCCGTTCTG AGTTCATGAC CATGACTTCT GGTACTGGTC TGCTGTACTC
1401  CACCTTCAGC CACTACGACG ACGTACGTCC GGGTGAAGTG GGTCAGCGTC
1451  AGAACGGCGT ACTGATCTCT AACGGTCAGG GTAAAGCGGT AGCGTTCGCG
1501  CTGTTCGGTC TGCAGGATCG CGGTAAGCTG TTCCTCGGTC ACGGTGCTGA
1551  AGTTTACGAA GGTCAGATTA TCGGTATTCA TAGCCGCTCT AACGACCTGA
1601  CTGTAAACTG CCTGACCGGT AAGAAACTGA CCAACATGCG TGCTTCCGGT
1651  ACTGACGAAG CCGTTGTTCT GGTTCCGCCT ATCCGCATGA CTCTGGAACA
1701  AGCTCTGGAA TTCATCCATG ATGACAAACT GGTAGAAGTG ACTCCGATCT
1751  CTATCCGTAT TCGTTAACGT CACTGA
```

Figure 3.10 The predicted amino acid sequence of EPEC E2348/69 TypA. The *typA* DNA sequence (**figure 3.9**) was translated using the Wisconsin GCG molecular biology program to provide the predicted amino acid sequence.

```
  1  MIEKLRNIAI IAHVDHGKTT LVDKLLQQSG TFDSRAETQE RVMDSNDLEK
 51  ERGITILAKN TAIKWNDYRI NIVDTPGHAD FGGEVERVMS MVDSVLLVVD
101  AFDGPMPQTR FVTKKAFAYG LKPIVVINKV DRPGARPDWV VDQVFDLFVN
151  FDATDEQLDF PIVYASALNG IAGLDHEDMA EDMTPLYQAI VDHVPAPDVD
201  LDGPFQMQIS QLDYNSYVGV IGIGRIKRGK VKPNQQVTII DSEGKTRNAK
251  VGKVLGHLGL ERIETDLAEA GDIVAITGLG ELNISDTVCD TQNVEALPAL
301  SVDEPTVSMF FCVNTSPFCG KEGKFVTSRQ ILDRLNKELV HNVALRVEET
351  EDADAFRVSG RGELHLSVLI ENMRREGFEL AVSRPKVIFR EIDGRKQEPY
401  ENVTLDVEQQ HQGSVMQALG ERKGDLKNMN PDGKGRVRLD YVIPSRGLIG
451  FRSEFMTMTS GTGLLYSTFS HYDDVRPGEV GQRQNGVLIS NGQGKAVAFA
501  LFGLQDRGKL FLGHGAEVYE GQIIGIHSRS NDLTVNCLTG KKLTNMRASG
551  TDEAVVLVPP IRMTLEQALE FIHDDKLVEV TPISIRIR*R H*
```

Figure 3.11 Nucleotide sequence comparison of the *typA* genes from *E. coli* K-12 and EPEC E2348/69. The nucleotide sequences of the *typA* genes from *E. coli* K-12 and EPEC E2348/69 shown below are compared against each other for homology using the Wisconsin GCG molecular biology program. The genes are highly homologous at the nucleotide level indicating a 98.8% homology.


K-12 1 GTGATCGAAAAATTGCGTAATATCGCCATCATCGCGCACGTAGACCATGG 50
EPEC 1 GTGATCGAAAAATTGCGTAATATCGCCATCATCGCGCACGTAGACCATGG 50
51 TAAAACCACCCTGGTAGACAAGCTGCTCCAACAATCCGGTACGTTCGACT 100
51 TAAAACCACCCTGGTAGATAAGCTGCTCCAACAATCCGGTACGTTCGACT 100
101 CTCGTGCCGAAACCCAAGAGCGCGTGATGGACTCCAACGATTTGGAGAAA 150
101 CTCGTGCCGAAACCCAAGAGCGCGTGATGGACTCCAACGATTTGGAGAAA 150
151 GAGCGTGGGATTACCATCCTCGCGAAAAACACCGCTATCAAATGGAATGA 200
151 GAGCGTGGGATTACCATCCTCGCGAAAAACACCGCTATCAAATGGAATGA 200
201 TTACCGTATCAACATCGTTGATACCCCGGGGCACGCCGACTTCGGTGGTG 250
201 TTACCGTATCAACATCGTTGATACCCCGGGGCACGCCGACTTCGGTGGTG 250
251 AAGTTGAACGTGTAATGTCCATGGTAGACTCAGTGCTGCTGGTGGTTGAC 300
251 AAGTTGAACGTGTAATGTCCATGGTAGACTCAGTGCTGCTGGTGGTTGAC 300
301 GCATTTGACGGCCCGATGCCGCAAACGCGCTTCGTAACCAAAAAAGCGTT 350
301 GCATTTGACGGCCCGATGCCGCAAACGCGCTTCGTAACCAAAAAAGCGTT 350
351 TGCTTACGGCCTGAAGCCGATTGTTGTTATCAACAAAGTTGACCGCCCTG 400
351 TGCTTACGGCCTGAAGCCGATTGTTGTTATCAACAAAGTTGACCGCCCTG 400
401 GCGCGCGTCCTGATTGGGTTGTGGATCAGGTATTCGATCTGTTCGTTAAC 450
401 GCGCGCGTCCTGATTGGGTTGTGGATCAGGTATTCGATCTGTTCGTTAAC 450
451 CTCGACGCGACCGACGAGCAGCTGGACTTCCCGATCGTTTACGCTTCTGC 500
451 TTCGACGCGACCGACGAGCAGCTGGACTTCCCGATCGTTTACGCTTCTGC 500
501 GCTGAACGGTATCGCGGGTCTGGACCACGAAGATATGGCGGAAGACATGA 550
501 GCTGAACGGTATCGCGGGTCTGGACCACGAAGATATGGCGGAAGACATGA 550
551 CCCCGCTGTACCAGGCGATTGTTGACCACGTTCCTGCGCCGGACGTTGAC 600
551 CCCCGCTGTATCAGGCGATTGTTGACCACGTTCCTGCGCCGGACGTTGAC 600
601 CTTGACGGTCCGTTCCAGATGCAGATTTCTCAGCTCGATTACAACAGCTA 650
601 CTTGACGGTCCGTTCCAGATGCAGATTTCTCAGCTCGATTACAACAGCTA 650
651 TGTTGGCGTTATCGGCATTGGSCGCATCAAGCGCGGTAAAGTGAAGCCGA 700
651 TGTTGGCGTTATCGGCATTGGCCGCATCAAGCGCGGTAAAGTGAAGCCGA 700
701 ACCAGCAGGTCACTATCATCGATAGCGAAGGCAAAACCCGCAACGCGAAA 750
701 ACCAGCAGGTTACTATCATCGATAGCGAAGGCAAAACCCGCAACGCGAAA 750

```
 751 GTCGGTAAAGTGCTGGGCCACCTCGGTCTGGAACGTATCGAAACCGATCT 800
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 751 GTCGGTAAAGTGCTGGGCCACCTCGGTCTGGAACGTATCGAAACCGATCT 800
           .         .         .         .         .
 801 GGCGGAAGCTGGCGATATCGTTGCGATCACGGGCCTTGGCGAACTGAACA 850
     |||||||||||||||||||| |||||||||| ||||||||||||||||
 801 GGCGGAAGCTGGCGATATCGTGGCGATCACGGGTCTTGGCGAACTGAACA 850
           .         .         .         .         .
 851 TTTCTGACACCGTTTGCGACACGCAAAACGTTGAAGCGCTGCCGGCACTC 900
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 851 TTTCTGACACCGTTTGCGACACGCAAAACGTTGAAGCGCTGCCGGCACTC 900
           .         .         .         .         .
 901 TCCGTTGATGAGCCGACCGTTTCTATGTTCTTCTGCGTTAACACCTCGCC 950
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 901 TCCGTTGATGAGCCGACCGTTTCTATGTTCTTCTGCGTTAACACCTCGCC 950
           .         .         .         .         .
 951 GTTCTGCGGTAAAGAAGGTAAGTTCGTAACGTCTCGTCAGATCCTGGATC 1000
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 951 GTTCTGCGGTAAAGAAGGTAAGTTCGTAACGTCTCGTCAGATCCTGGATC 1000
           .         .         .         .         .
1001 GTCTGAACAAAGAACTGGTACACAACGTTGCGCTGCGCGTAGAAGAAACC 1050
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1001 GTCTGAACAAAGAACTGGTACACAACGTTGCGCTGCGCGTAGAAGAAACC 1050
           .         .         .         .         .
1051 GAAGACGCCGATGCGTTCCGCGTTTCTGGTCGTGGCGAACTGCACCTGTC 1100
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1051 GAAGACGCCGATGCGTTCCGCGTTTCTGGTCGTGGCGAACTGCACCTGTC 1100
           .         .         .         .         .
1101 TGTTCTGATCGAAAACATGCGTCGTGAAGGTTTCGAACTGGCGGTATCCC 1150
     |||||| ||||||||||||||||||||||||||||||||||||||||||
1101 TGTTCTTATCGAAAACATGCGTCGTGAAGGTTTCGAACTGGCGGTATCCC 1150
           .         .         .         .         .
1151 GTCCGAAAGTTATCTTCCGTGAAATCGACGGTCGTAAACAAGAGCCGTAT 1200
     |||||||||| ||||||||||||||||||||||||||||||||||||||
1151 GTCCGAAAGTAATCTTCCGTGAAATCGACGGTCGTAAACAAGAGCCGTAT 1200
           .         .         .         .         .
1201 GAAAACGTGACTCTGGACGTTGAAGAACAGCATCAGGGTTCTGTAATGCA 1250
     |||||||||||| |||||||||||| |||||||||||||||||| |||||
1201 GAAAACGTGACGCTGGACGTTGAACAACAGCATCAGGGTTCTGTGATGCA 1250
           .         .         .         .         .
1251 GGCGCTGGGCGAACGTAAAGGCGACCTGAAAAACATGAATCCAGACGGTA 1300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1251 GGCGCTGGGCGAACGTAAAGGCGACCTGAAAAACATGAATCCAGACGGTA 1300
           .         .         .         .         .
1301 AAGGCCGCGTACGTCTCGACTACGTGATCCCAAGCCGTGGTCTGATTGGC 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1301 AAGGCCGCGTACGTCTCGACTACGTGATCCCAAGCCGTGGTCTGATTGGC 1350
           .         .         .         .         .
1351 TTCCGTTCTGAGTTCATGACCATGACTTCCGGTACTGGTCTGCTGTACTC 1400
     |||||||||||||||||||||||||| ||||||||||||||||||||||
1351 TTCCGTTCTGAGTTCATGACCATGACTTCTGGTACTGGTCTGCTGTACTC 1400
           .         .         .         .         .
1401 CACCTTCAGCCACTACGACGACGTACGTCCGGGTGAAGTGGGTCAGCGTC 1450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1401 CACCTTCAGCCACTACGACGACGTACGTCCGGGTGAAGTGGGTCAGCGTC 1450
           .         .         .         .         .
1451 AGAACGGCGTACTGATCTCTAACGGTCAGGGTAAAGCGGTCGCGTTCGCG 1500
     |||||||||||||||||||||||||||||||||||| |||||||||
1451 AGAACGGCGTACTGATCTCTAACGGTCAGGGTAAAGCGGTAGCGTTCGCG 1500
```

```
                .         .         .         .         .
1501 CTGTTCGGTCTGCAGGATCGCGGTAAGCTGTTCCTCGGTCACGGTGCAGA 1550
     |||||||||||||||||||||||||||||||||||||||||||||||| ||
1501 CTGTTCGGTCTGCAGGATCGCGGTAAGCTGTTCCTCGGTCACGGTGCTGA 1550
                .         .         .         .         .
1551 AGTTTACGAAGGTCAGATTATCGGTATTCATAGCCGCTCTAACGACCTGA 1600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1551 AGTTTACGAAGGTCAGATTATCGGTATTCATAGCCGCTCTAACGACCTGA 1600
                .         .         .         .         .
1601 CTGTAAACTGCCTGACCGGTAAGAAACTGACCAACATGCGTGCTTCCGGT 1650
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1601 CTGTAAACTGCCTGACCGGTAAGAAACTGACCAACATGCGTGCTTCCGGT 1650
                .         .         .         .         .
1651 ACTGACGAAGCCGTTGTTCTGGTTCCGCCTATCCGCATGACTCTGGAACA 1700
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1651 ACTGACGAAGCCGTTGTTCTGGTTCCGCCTATCCGCATGACTCTGGAACA 1700
                .         .         .         .         .
1701 AGCTCTGGAGTTCATCGATGATGACGAACTGGTAGAAGTGACTCCGACCT 1750
     ||||||||| ||||||| |||||||| |||||||||||||||||||| ||
1701 AGCTCTGGAATTCATCCATGATGACAAACTGGTAGAAGTGACTCCGATCT 1750
                .         .
1751 CTATCCGTATTCGTAAACGTCACTGA 1776
     ||||||||||||| ||||||||||
1751 CTATCCGTATTCGTTAACGTCACTGA 1776
```

Figure 3.12 Predicted amino acid sequence comparison of TypA from *E. coli* K-12 and EPEC E2348/69. The predicted amino acid sequences of the *typA* genes from *E. coli* K-12 and EPEC E2348/69 were translated from the respective nucleotide sequences using the Wisconsin GCG molecular biology program. They are compared against each other below for homology. The genes are highly homologous at the amino acid level indicating a 98.8% homology.

```
                .         .         .         .         .
K-12    1 MIEKLRNIAIIAHVDHGKTTLVDKLLQQSGTFDSRAETQERVMDSNDLEK 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||
EPEC    1 MIEKLRNIAIIAHVDHGKTTLVDKLLQQSGTFDSRAETQERVMDSNDLEK 50
                .         .         .         .         .
       51 ERGITILAKNTAIKWNDYRINIVDTPGHADFGGEVERVMSMVDSVLLVVD 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||
       51 ERGITILAKNTAIKWNDYRINIVDTPGHADFGGEVERVMSMVDSVLLVVD 100
                .         .         .         .         .
      101 AFDGPMPQTRFVTKKAFAYGLKPIVVINKVDRPGARPDWVVDQVFDLFVN 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      101 AFDGPMPQTRFVTKKAFAYGLKPIVVINKVDRPGARPDWVVDQVFDLFVN 150
                .         .         .         .         .
      151 LDATDEQLDFPIVYASALNGIAGLDHEDMAEDMTPLYQAIVDHVPAPDVD 200
          :||||||||||||||||||||||||||||||||||||||||||||||||||
      151 FDATDEQLDFPIVYASALNGIAGLDHEDMAEDMTPLYQAIVDHVPAPDVD 200
                .         .         .         .         .
      201 LDGPFQMQISQLDYNSYVGVIGIGRIKRGKVKPNQQVTIIDSEGKTRNAK 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      201 LDGPFQMQISQLDYNSYVGVIGIGRIKRGKVKPNQQVTIIDSEGKTRNAK 250
                .         .         .         .         .
      251 VGKVLGHLGLERIETDLAEAGDIVAITGLGELNISDTVCDTQNVEALPAL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      251 VGKVLGHLGLERIETDLAEAGDIVAITGLGELNISDTVCDTQNVEALPAL 300
                .         .         .         .         .
      301 SVDEPTVSMFFCVNTSPFCGKEGKFVTSRQILDRLNKELVHNVALRVEET 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      301 SVDEPTVSMFFCVNTSPFCGKEGKFVTSRQILDRLNKELVHNVALRVEET 350
                .         .         .         .         .
      351 EDADAFRVSGRGELHLSVLIENMRREGFELAVSRPKVIFREIDGRKQEPY 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      351 EDADAFRVSGRGELHLSVLIENMRREGFELAVSRPKVIFREIDGRKQEPY 400
                .         .         .         .         .
      401 ENVTLDVEEQHQGSVMQALGERKGDLKNMNPDGKGRVRLDYVIPSRGLIG 450
          ||||||||:||||||||||||||||||||||||||||||||||||||||||
      401 ENVTLDVEQQHQGSVMQALGERKGDLKNMNPDGKGRVRLDYVIPSRGLIG 450
                .         .         .         .         .
      451 FRSEFMTMTSGTGLLYSTFSHYDDVRPGEVGQRQNGVLISNGQGKAVAFA 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      451 FRSEFMTMTSGTGLLYSTFSHYDDVRPGEVGQRQNGVLISNGQGKAVAFA 500
                .         .         .         .         .
      501 LFGLQDRGKLFLGHGAEVYEGQIIGIHSRSNDLTVNCLTGKKLTNMRASG 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      501 LFGLQDRGKLFLGHGAEVYEGQIIGIHSRSNDLTVNCLTGKKLTNMRASG 550
                .         .         .         .
      551 TDEAVVLVPPIRMTLEQALEFIDDDELVEVTPTSIRIRKRH* 592
          ||||||||||||||||||||.||.|||||.|||||  |||
      551 TDEAVVLVPPIRMTLEQALEFIHDDKLVEVTPISIRIR*RH* 592
```

Table 3.2 Differences in amino acid composition of TypA in *E. coli* K-12 and EPEC E2348/69 strains. The *typA* gene from EPEC E2348/69 was sequenced as described in **section 3.2.4.2** (see **figure 3.9**) and translated using the Wisconsin GCG molecular biology program to obtain the predicted amino acid sequence (see **figure 3.10**). The sequence was then aligned with that of *E. coli* K-12 *typA* to determine the differences.

Amino acid position	*E. coli* K-12	EPEC E2348/69
151	leucine	phenylalanine
409	glutamic acid	glutamine
573	aspartic acid	histidine
576	glutamic acid	lysine
583	threonine	isoleucine
589	lysine	STOP

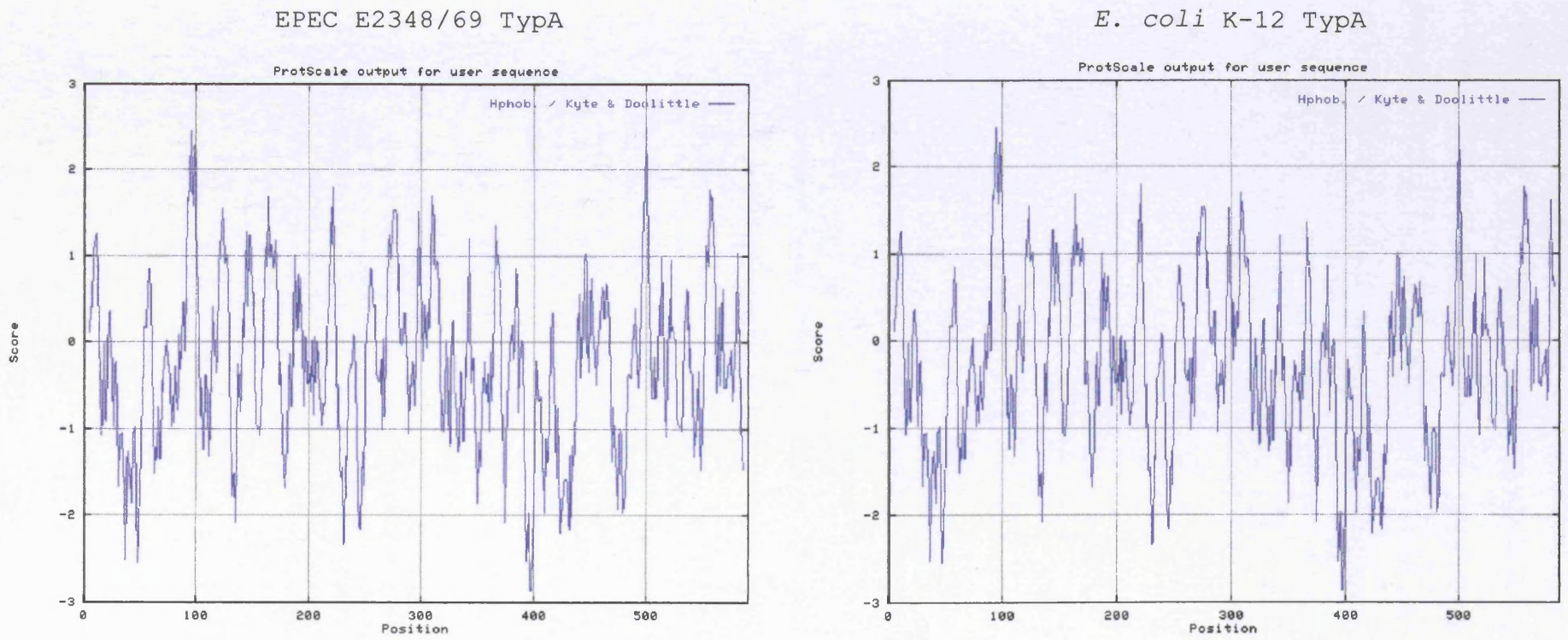


Figure 3.13 Hydropathy plots of EPEC E2348/69 TypA and *E. coli* K-12 TypA. Each sequence was entered into the Protscale program (http://expasy.hcuge.ch/cgi-bin/protscale.p1) and a hydropathy plot performed according to the method of Kyte and Doolittle (1982). The plots indicate that the *C*-terminus of EPEC E2348/69 TypA is slightly more hydrophobic than *E. coli* K-12 TypA.

Figure 3.14 NNPredict secondary structure queries of EPEC E2348/69 TypA and *E. coli* K-12 TypA. Each sequence was entered into the NNPredict program (http://www.cmpharm.ucsf.edu/cgi-bin/nnpredict.pl) and a secondary structure analysis of each performed. The secondary structure predictions were then compared. The results indicate that there is little difference between the two proteins; again there were only slight differences at the *C*-terminus of TypA.

Sequence EPEC E2348/69 TypA:

MIEKLRNIAIIAHVDHGKTTLVDKLLQQSGTFDSRAETQERVMDSNDLEKERGITILAKN
TAIKWNDYRINIVDTPGHADFGGEVERVMSMVDSVLLVVDAFDGPMPQTRFVTKKAFAYG
LKPIVVINKVDRPGARPDWVVDQVFDLFVNFDATDEQLDFPIVYASALNGIAGLDHEDMA
EDMTPLYQAIVDHVPAPDVDLDGPFQMQISQLDYNSYVGVIGIGRIKRGKVKPNQQVTII
DSEGKTRNAKVGKVLGHLGLERIETDLAEAGDIVAITGLGELNISDTVCDTQNVEALPAL
SVDEPTVSMFFCVNTSPFCGKEGKFVTSRQILDRLNKELVHNVALRVEETEDADAFRVSG
RGELHLSVLIENMRREGFELAVSRPKVIFREIDGRKQEPYENVTLDVEQQHQGSVMQALG
ERKGDLKNMNPDGKGRVRLDYVIPSRGLIGFRSEFMTMTSGTGLLYSTFSHYDDVRPGEV
GQRQNGVLISNGQGKAVAFALFGLQDRGKLFLGHGAEVYEGQIIGIHSRSNDLTVNCLTG
KKLTNMRASGTDEAVVLVPPIRMTLEQALEFIHDDKLVEVTPISIRIR

Secondary structure prediction (H = helix, E = strand, - = no prediction):

--HHHHHHHEEEE-------HHHHHHH-----------HH-------------HEEHH--
---------EEE-----------HHHHHHHHHHHHHEEE------------HHHHHHHH-
---EEEE------------EHH-HHHHEE------------EEEH---H-----------
H---HHHHEE----------------EEE--------EEEEEE--H----------EEEE
--------H-H-HHH------HHHHHHHHH--HEEEE----E------------------
--------EEEE------------EE----HHHHHHHHHHHHHH-HH----H--HHEE--
---EEEHHHHHH-HHH---EH-----EEEE---------------HHHH-----HHHHH-
---------------EEEEEE-----EE------EEE-----EEEEE------------
------EEEE-----HHHHHHH------HEEE-----HE--EEEEEE------EEE----
----EE-------EEE------HHHHHHHHHH-----------EEE--

Sequence *E. coli* K-12 TypA:

MIEKLRNIAIIAHVDHGKTTLVDKLLQQSGTFDSRAETQERVMDSNDLEKERGITILAKN
TAIKWNDYRINIVDTPGHADFGGEVERVMSMVDSVLLVVDAFDGPMPQTRFVTKKAFAYG
LKPIVVINKVDRPGARPDWVVDQVFDLFVNLDATDEQLDFPIVYASALNGIAGLDHEDMA
EDMTPLYQAIVDHVPAPDVDLDGPFQMQISQLDYNSYVGVIGIGRIKRGKVKPNQQVTII
DSEGKTRNAKVGKVLGHLGLERIETDLAEAGDIVAITGLGELNISDTVCDTQNVEALPAL
SVDEPTVSMFFCVNTSPFCGKEGKFVTSRQILDRLNKELVHNVALRVEETEDADAFRVSG
RGELHLSVLIENMRREGFELAVSRPKVIFREIDGRKQEPYENVTLDVEEQHQGSVMQALG
ERKGDLKNMNPDGKGRVRLDYVIPSRGLIGFRSEFMTMTSGTGLLYSTFSHYDDVRPGEV
GQRQNGVLISNGQGKAVAFALFGLQDRGKLFLGHGAEVYEGQIIGIHSRSNDLTVNCLTG
KKLTNMRASGTDEAVVLVPPIRMTLEQALEFIDDDELVEVTPTSIRIRKRH

Secondary structure prediction (H = helix, E = strand, - = no prediction):

--HHHHHHHEEEE-------HHHHHHH-----------HH-------------HEEHH--
---------EEE-----------HHHHHHHHHHHHHEEE------------HHHHHHHH-
---EEEE------------EHH-HHHHEE------------EEEH---H-----------
H---HHHHEE----------------EEE--------EEEEEE--H----------EEEE
--------H-H-HHH------HHHHHHHHH--HEEEE----E------------------
--------EEEE------------EE----HHHHHHHHHHHHHH-HH----H--HHEE--
---EEEHHHHHH-HHH---EH-----EEEE---------------HHHH----HHHHH-
---------------EEEEEE-----EE------EEE-----EEEEE------------
------EEEE-----HHHHHHH------HEEE-----HE--EEEEEE------EEE----
----EE-------EEE------HHHHHHHHH------EE-----EEEE---

3.2.5 The distribution of *typA* among other bacteria.

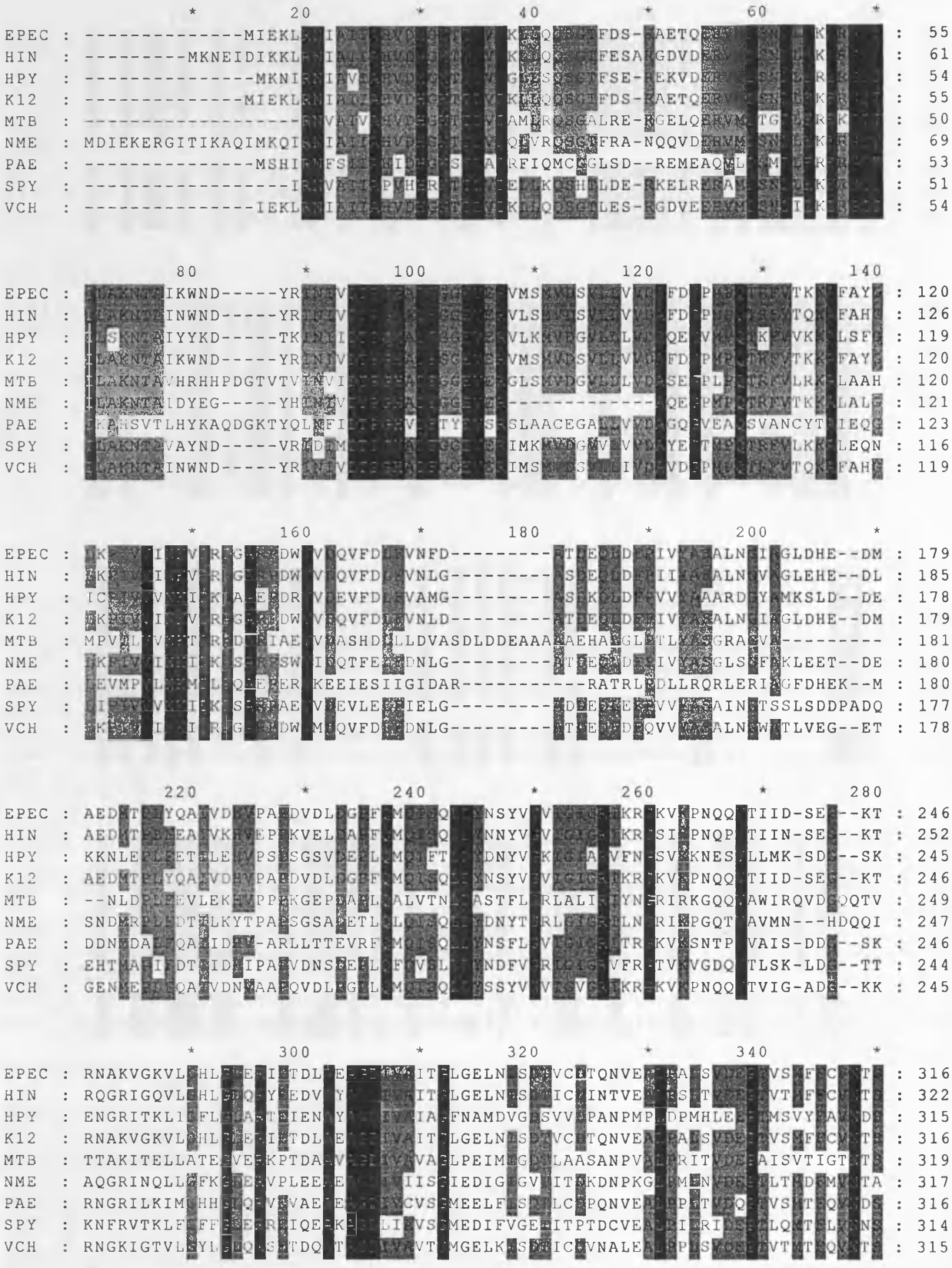
To date it is known that *typA* homologues exist in *E. coli* K-12 (Blattner *et al* 1997), EPEC E2348/69 (Farris *et al* 1998; Freestone *et al* 1998b) and *Salmonella typhimurium* (Qi *et al* 1995). However, sequencing of bacterial genomes has accelerated at unprecedented speed in the last few years. Complete genome sequences are now available for 20 bacteria whilst another 45 are currently in progress. The importance of the *typA* gene in general could be ascertained by determining its presence in other pathogenic and non-pathogenic bacteria. To achieve this, the EPEC E2348/69 *typA* sequence was entered into the BLAST web-page. A gapped BLAST search was then performed on complete and incomplete genome sequences in the Genbank databases (Altschul *et al* 1997). Data from incomplete genome sequences was manipulated using the Wisconsin GCG molecular biology program. It was found that a total of 24 different bacterial genomes possessed TypA homologues (**table 3.3**); amino acid comparisons of eight of these are shown in **figure 3.15** to indicate the high level of homology. Furthermore, *Plasmodium falciparum*, one of the Plasmodium species responsible for malaria, also possessed a TypA homologue. Of the bacterial TypA homologues, 16 are considered to be human pathogens. However, there was no correlation with the presence of TypA with virulence as TypA homologues were present in both humans pathogens such as *Vibrio cholerae* and non-pathogens such as *Deinococcus radiodurans*. However, homology was often seen in pathogens as these are the majority of organisms currently being sequenced. TypA homology ranged from 48% for *Mycobacterium tuberculosis* to 89% for *Yersinia pestis* indicating that homology was very high in some organisms. Interestingly, homology of individual amino acids was often shared between seven or more of the nine organisms compared in figure 3.15; *Mycobacterium tuberculosis* was the organism with the most differences against the other eight organisms although homology was still good. Tyrosine residues were conserved across at least seven of the nine organisms when compared against *E. coli* K-12 apart from three at residues 220, 440 and 480. However, some organisms possessed additional tyrosine residues at various points in the TypA homologues although these were not conserved.

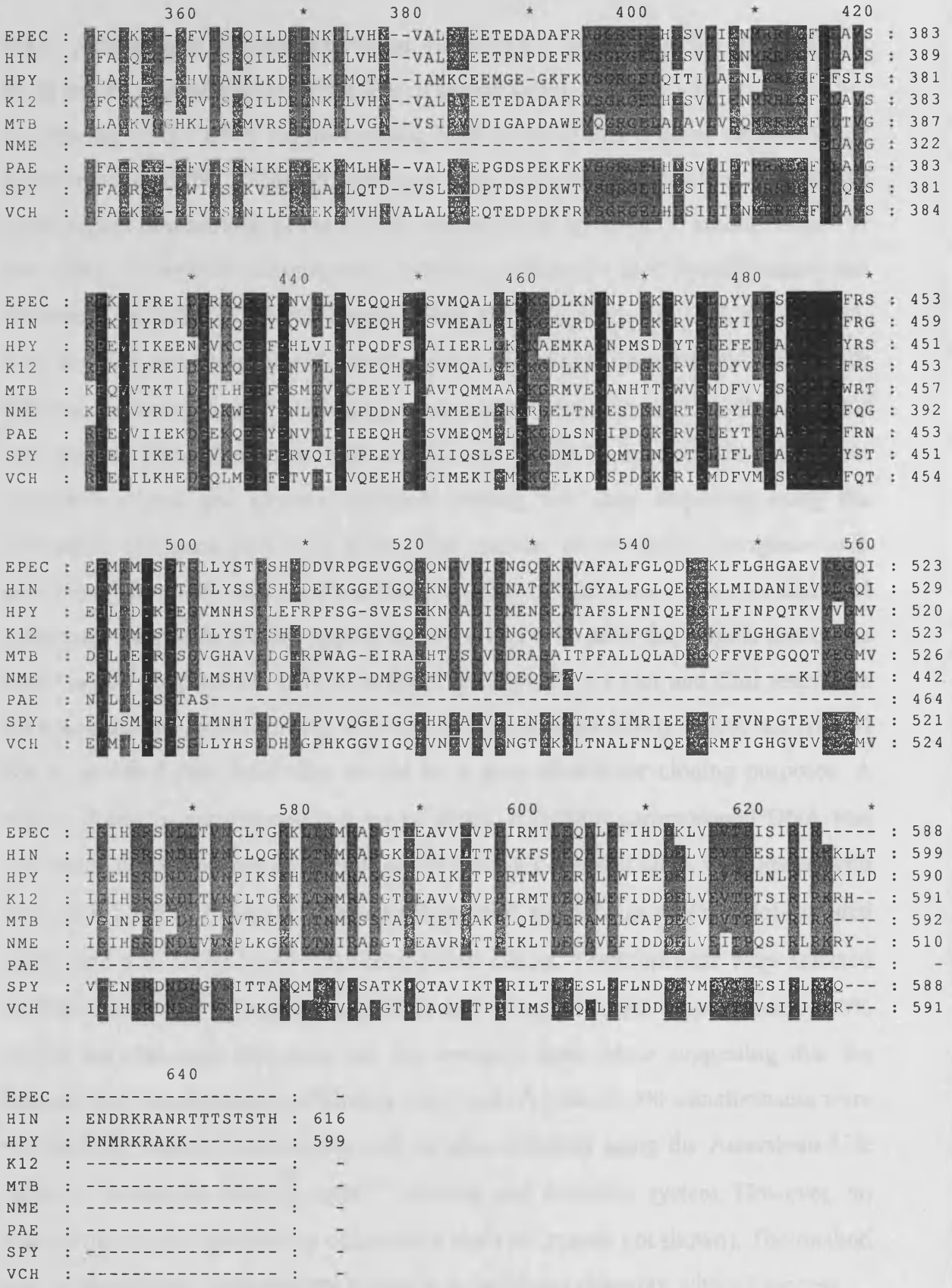
Table 3.3 Genomes in which TypA homologues exist. Homologues of TypA were found in 25 different genomes from the Genbank databases using the BLAST search facility. These included 24 bacteria and *Plasmodium falciparum*. Of the bacteria, 16 are considered human pathogens. This list does not include *E. coli* as many different pathovars, as well as non-pathogenic *E. coli*, possess TypA.

Actinobacillus actinomycetemcomitans	*Bacillus subtilis*
Campylobacter jejuni NCTC 11168	*Chlamydia trachomatis*
Clostridium acetobutylicum	*Deinococcus radiodurans*
Enterococcus faecalis	*Haemophilus influenzae*
Helicobacter pylori	*Mycobacterium tuberculosis* CSU#93
Neisseria gonorrhoeae	*Neisseria meningitidis* MC58
Neisseria meningitidis serogroup A	*Plasmodium falciparum* (chromosome 2)
Porphyromonas gingivalis W83	*Pseudomonas aeruginosa*
Pyrococcus furiosus	*Staphylococcus aureus*
Streptococcus pneumoniae	*Streptococcus pyogenes*
Synechocytis sp.	*Thermotoga maritima*
Treponema pallidum	*Vibrio cholerae*
Yersinia pestis	

Figure 3.15 Predicted amino acid sequence comparison of TypA homologues. Complete and partially complete genome sequence data were obtained from the Genbank database by BLAST search and from the TIGR website search facility. Eight sequences were chosen for comparison with *E. coli* K-12, namely EPEC E2348/69, *Haemophilus influenzae*, *Helicobacter pylori*, *Mycobacterium tuberculosis*, *Neisseria meningitidis*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes* and *Vibrio cholerae* on the basis of their ability to cause human disease. The sequences were edited as necessary using the Wisconsin GCG molecular biology program and a sequence line-up performed using the CLUSTAL program.

LEGEND:
Amino acid sequence comparison of TypA homologues from EPEC E2348/69 (EPEC), *Haemophilus influenzae* (HIN), *Helicobacter pylori* (HPY), *E. coli* K12 (K12), *Mycobacterium tuberculosis* (MTB), *Pseudomonas aeruginosa* (PAE), *Neisseria meningitidis* (NME), *Streptococcus pyogenes* (SPY) and *Vibrio cholerae* (VCH). Black highlight indicates identity in all 9 homologues; grey highlight indicates at least 7/9 amino acids are conserved.





3.2.6 Attempts at cloning the region downstream of the *Hind*III site at 23550 bp in the 87 minute region of the *E. coli* chromosome.

Sequencing of the DNA region flanking *typA* revealed that a novel region existed downstream in EPEC E2348/69. Attempts were therefore made to clone the putative novel region downstream of the *Hind*III site at 23550 bp in the 87 minute region of the EPEC E2348/69 chromosome. Initially, Southern blot hybridisation was performed on EPEC E2348/69 chromosomal DNA digested with *Cla*I, *Eco*RI, *Pst*I, *Pvu*I and *Sty*I and detected with labelled pSCC06 using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system (results not shown). This was done to determine what PCR product size would be required for subsequent restriction digest and cloning. Routine cloning was then employed using the restriction enzymes *Pst*I and *Cla*I. The choice of restriction enzymes was ascertained from the results of the Southern blot hybridisation. The *Pst*I and *Cla*I restriction sites were present approximately 4 and 6 kb after the *Hind*III site of the novel region downstream of *typA* in EPEC E23248/69. As *Pst*I and *Cla*I restriction sites also existed approximately 422 bp and 200 bp respectively before the *Hind*III site it appeared that these sites would be a good choice for cloning purposes. A double digest of approximately 2 μg of EPEC E2348/69 chromosomal DNA was performed using *Pst*I and *Cla*I. Half of the resultant digested DNA was ligated with each of 0.5 μg of *Pst*I/*Cla*I-digested pBluescript and 0.5 μg of *Pst*I/*Cla*I-digested pAV3 and transformed into electrocompetent DH5α. Transformants were selected on LBA containing 100 μg/ml ampicillin and 40 μg/ml X-gal. Approximately 70% of the transformants that grew on this medium were white suggesting that the ligation and transformation efficiency was good. A total of 400 transformants were screened by colony blot analysis with labelled pSCC06 using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system. However, no transformants were positive by colony blot analysis (results not shown). The method was repeated twice in an attempt to gain a recombinant clone but without success.

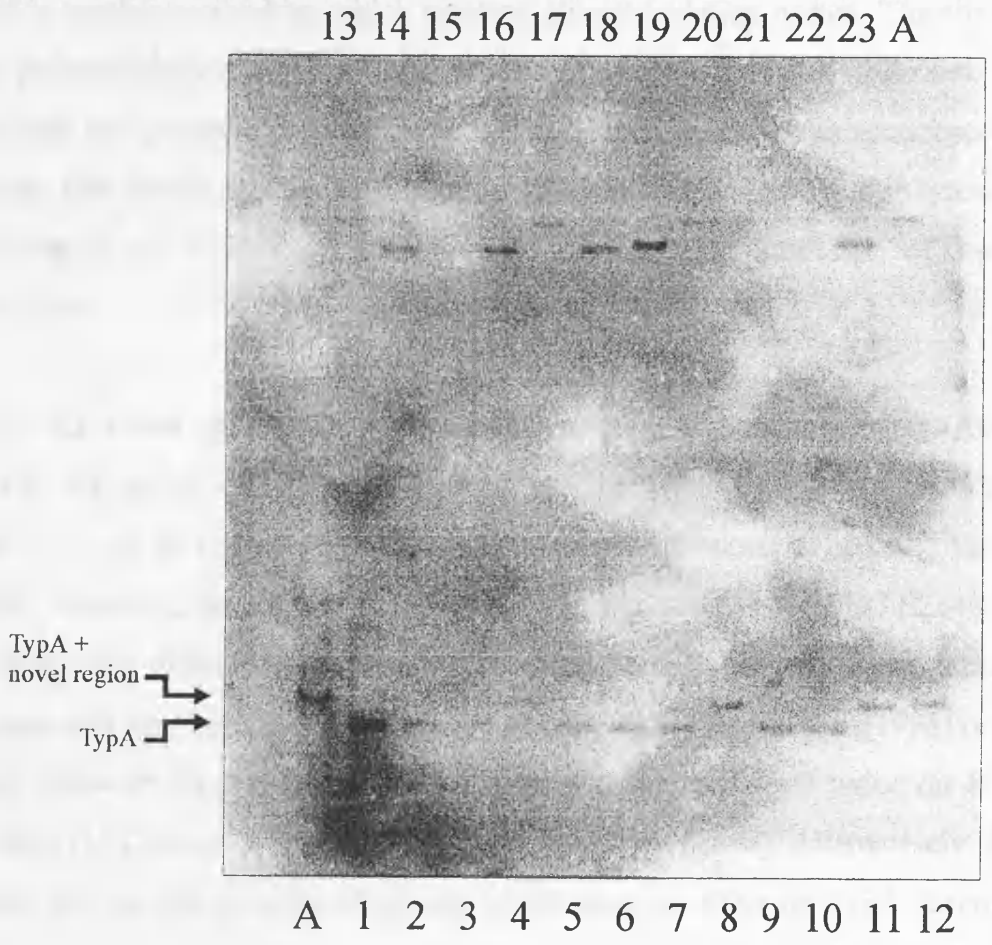
3.2.7 The distribution of the novel region amongst wild-type and laboratory strains of *Escherichia coli.*

It was shown from the sequencing of clone pSCC01 that a region of novel DNA existed after the *typA* gene in EPEC E2348/69. Although cloning attempts were unsuccessful to date, its absence in *E. coli* K-12 suggested that this region may have a role in virulence. To further test this hypothesis in the absence of sequence data, the distribution of the novel region could be ascertained. To achieve this, 23 strains of various pathogenic and non-pathogenic *E. coli* were screened for the presence of the novel region (**table 3.4**). Firstly, chromosomal DNA was prepared from each strain using the mini chromosomal preparation method and 1 μg of DNA from each was then digested with *Bam*HI overnight to ensure full digestion. The digested DNA was subjected to agarose gel electrophoresis followed by Southern blot hybridisation with labelled pSCC-T using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system. It was found that, including EPEC E2348/69, 11 of the 24 strains tested possessed the novel region (**figure 3.16**). The distribution of *typA* was not strictly defined by the virulence potential of each strain. Although those strains which possessed the novel region were virulent strains of *E. coli*, other virulent strains did not possess the region. However, all non-virulent strains did not possess the novel region. Additionally, the results suggest that EHEC strain 86-24, which is serogroup O157:H7, possesses two copies of *typA* but does not possess the novel DNA region. However, there is also the possibility that partial digest of the EHEC chromosomal DNA occurred and this result should therefore be confirmed by repeating with further human and non-human isolates of EHEC.

Table 3.4 The distribution of *typA* and the novel region in various clinical and laboratory strains of *E. coli*. The pathogenic and non-pathogenic strains listed opposite were used for determining the distribution of *typA* and the novel region in various strains of *E. coli*. Chromosomal DNA was prepared from each strain using the mini chromsomal preparation method and 1 μg of DNA from each was then digested with *Bam*HI overnight to ensure full digestion. The digested DNA was subjected to agarose gel electrophoresis followed by Southern blot hybridisation with labelled pSCC-T using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system. The results indicated in the column labelled '*typA*' and 'novel region' should be read in conjunction with the Southern blot results shown in **figure 3.16**.

Lane	Strain	Serotype	Biotype	*typA*	Novel region
A	E2348/69	O127:H6	EPEC	+	+
1	E11881B	O25:H42	ETEC	+	-
2	17-2	O3:H2	EAggEC	+	-
3	1457-75	O124:H3	EIEC	+	-
4	C1845	O75:NM	EIEC	+	+
5	86-24	O157:H7	EHEC	++	+
6	3605-73	O26:H11	EHEC	+	-
7	CH7/6	O26:H1	EPEC	+	-
8	261/88	O86:H34	EPEC	+	+
9	19	O111:H2	EPEC	+	-
10	KH1/8	O114:H2	EPEC	+	-
11	57	O119:H6	EPEC	+	+
12	292	O125:	EPEC	+	+
13	E611	O126:	EPEC	+	+
14	IR3/8	O127:H-	EPEC	+	-
15	44/86	O128:H2	EPEC	+	-
16	DH5α	K-12	N/A	+	-
17	53/85	O142:34	EPEC	+	+
18	1038	Not known	Septicaemia	+	-
19	1187	O75:K5:H-	Cystitis	+	-
20	3071	O6:K2:H1	Septicaemia	+	+
21	1192	O1:K1:H7	Pyelonephritis	+	+
22	1152	O6:K13:H1	Cystitis	+	+
23	HB101	Not known	N/A	+	-

Figure 3.16 Southern blot hybridisation analysis of chromosomal DNA from various *E. coli* strains probed with *typA*. Chromosomal DNA from various pathogenic and non-pathogenic *E. coli* strains (see **table 3.4**) was prepared and digested with *Bam*HI. The resultant DNA fragments were separated using 1% gel electrophoresis, transferred to nylon membrane by Southern blot and hybridised with *typA* purified from pSCC-T. The blot demonstrates that all *E. coli* strains tested possess *typA*. Furthermore, use of the *Bam*HI restriction site shows that some strains also possess the novel region downstream of *typA* whereas others do not. The list is shown and should be read in conjunction with **table 3.4**. Importantly, the blot suggests that EHEC strain 86-24, which is serogroup O157:H7, possesses two copies of *typA* but does not possess the novel region.



3.3 Discussion

3.3.1 General discussion.

The cloning and sequencing of the *typA* gene from enteropathogenic *E. coli* strain E2348/69 has indicated that there are no significant differences between *typA* from EPEC and *E. coli* K-12. Although there are six differences at the amino acid level, none of these appear to be significant in terms of secondary structure and therefore protein function using the prediction programs in this chapter. However, it is possible for single amino acid changes to affect secondary protein structure and thereby affecting protein phosphorylation due to exposure of an appropriate amino acid residue. Importantly, there was no additional tyrosine residue present in the *typA* gene of EPEC E2348/69 and there were no amino acid alterations around the existing tyrosine residues. This is important sequence data as the possibility exists that there is another method by which tyrosine phosphorylation occurs. The site of tyrosine phosphorylation requires analysis by cleavage of TypA followed by labelling with anti-phosphotyrosine antibodies or 'time of flight' mass spectroscopy sequencing. This would provide an insight into the mechanisms involved in tyrosine phosphorylation of EPEC TypA and may confirm the presence of auto-phosporylation.

The EPEC E2348/69 *typA* DNA sequence differs slightly from that of MAR001 *typA* and by six amino acids (Farris *et al* 1998). The MAR001 TypA also differs from that in *E. coli* K-12 by three amino acid sequences (Blattner *et al* 1997; Farris *et al* 1998). However, the predicted amino acid TypA sequences of EPEC E2348/69 and MAR001 also differ. None of these differences between the two EPEC strains are the same and may be due to sequencing errors by either Farris *et al* (1998) or in this thesis although *typA* from each strain was sequenced at least twice on both DNA strands (O'Connor D, personal communication; this thesis). Alternatively, the differences may be due to strain mutations which have no effect on TypA function although these two strains are very closely related so it would be unusual for such mutations to have occurred. Clarification of these differences will be addressed

when further work is performed on determining the role of individual amino acids on the function of *typA*.

Using the BLAST search facility an interesting discovery was made with regards to the existence of *typA* homologues in other bacterial genera. TypA homologues were found in 24 different bacterial genomes as well as in *Plasmodium falciparum*. This is important as it is unlikely that *typA* is a redundant gene present in so many organisms of such diversity. The genera that possess a *typA* homologue are Gram-positive, Gram-negative, chlamydiae, treponeme or *Plasmodia* and therefore exist in many classes of organism known. They probably exist also in other organisms that have not yet been sequenced. Sixteen of the bacteria that possess a *typA* homologue are human pathogens; this therefore raises the question of whether *typA* is a virulence gene. It is not known whether the *typA* homologues are tyrosine phosphorylated in these pathogens or, indeed, the other bacteria. This is therefore an area for further study.

The 1.3 kb region immediately downstream of the *typA* terminator loop in EPEC strain E2348/69 was not present in *E. coli* K-12 strains and was therefore novel DNA. Although the next 422 bp shared homology with ORF *o300* at 35.5-36.2 kb in the 88 minute region of the *E. coli* chromosome, it was hypothesised that this region was also novel. The region may have resulted from a deletion or an insertion event in comparison to *E. coli* K-12; my hypothesis was that the region was deleted from *E. coli* K-12 as its presence no longer conferred selective advantage in this strain. This hypothesis was strengthened by the distribution pattern of the novel region in various *E. coli* strains. The distribution was not random and appeared to be loosely linked to virulence potential. All those strains which possessed the novel region were virulent. Those virulent strains which no longer possess the region may have deleted the DNA in a similar manner to *E. coli* K-12. Laboratory strains of *E. coli* will have lost the region because its presence is not required for survival in laboratory media. Its absence in some pathovars of *E. coli* may be due to the fact that the region is not required for survival in the ecological niche of the particular pathovar. Importantly, EHEC strain 86-24 (serotype O157:H7) appeared to possess

two copies of *typA* as shown by the Southern blot results of *Bam*HI-digested chromosomal DNA probed with *typA*. One other EHEC strain, 3605-73 serotype O26:H11, was used which possessed a single copy of *typA*. Unlike EHEC serotype O157:H7, this serotype does not commonly cause severe human disease and therefore the presence of two copies of *typA* in EHEC strain 86-24 could be linked to virulence. Further strains of various serotypes will need to be studied to confirm this hypothetical link. Also, it would be useful to test strains from various animal carriers, such as cattle, to determine the number of copies of *typA* in EHEC strains carried by them.

To test whether the novel region is an insertion, and therefore a potential pathogenicity island, the DNA sequence should be analysed for the presence of certain characteristics. According to the definition of Hacker *et al* (1997) the potential occurrence of a pathogenicity island can be determined by several criteria. These criteria include the following:

- the presence of many virulence genes within the pathogenicity island
- the occupation of large chromosomal region (often in excess of 30 kb)
- they exist as compact, distinct genetic units which are often flanked by direct repeats
- they are often associated with tRNA genes or insertion sequence (IS) elements
- the presence of mobility genes ie. IS elements, transposases, origins of plasmid replication
- genetic instability

The normal average GC content of *E. coli* is 51%. However, the GC content in pathogenicity islands often decreases to between 39 and 41% (Hacker *et al* 1997) although this is not always the case and depends on the origin of the DNA. The GC content of the novel region downstream of *typA* in EPEC E2348/69 is 50.3%, not much lower than the average GC content of *E. coli*. Furthermore, there are no direct repeats flanking the region. These data support my hypothesis that the novel region is not present due to a horizontal gene transfer event via plasmid, phage or transposon insertion into the chromosome although there is the possibility that the novel region originated from a bacterium with a similar %GC content.

The sequence of the 3' end of the cloned *typA* insert was found to possess 72% sequence identity over 271 amino acids with ORF *o300* at 35.4-36.3 kb in the 88 minute region of the *E. coli* K-12 chromosome. The product of ORF *o300* is similar to members of the phosphofructokinase B family of carbohydrate kinases although the *pfk*B gene itself is located at 35 minutes in the *E. coli* K-12 chromosome. Phosphofructokinases in *E. coli* are responsible for the addition of a phosphate group to fructose-6-P as part of the Embden-Meyerhof pathway (Fraenkel 1996). There are two phosphofructokinase genes, namely *pfkA* and *pfkB*, the former encoding the main phosphofructokinase of *E. coli* whilst the latter accounts for approximately 10% of phosphofructokinase activity. The actual need for *pfkB* in *pfkA*-positive *E. coli* K-12 strains is not known but it is essential in *pfkA*-negative strains (Fraenkel 1996). This may indicate that in virulent *E. coli* strains, such as EPEC, *pfkB* may be essential in the presence of *pfkA* for full virulence. Furthermore, other organisms, including pathogens such as *Streptococcus pneumoniae*, possess phophofructokinases which may be cell surface-associated.

The *E. coli* genome sequence has now been published (Blattner *et al* 1997) and is derived from MG1655, a laboratory strain with limited genetic manipulation (Bachmann 1996). This strain was ideal for genomic sequencing as it has not had genes deleted or added artificially in the laboratory. However, it is non-pathogenic and has been maintained in a laboratory environment. It may have been better to sequence a prototype pathogenic *E. coli* strain; in fact, sequencing of the EHEC genome is now being performed. In E. coli strain MG1655 it is possible that during its evolution any number of chromosomally-encoded virulence genes which may have been present in this strain have now been deleted because their presence is not required for survival. This may lead to future difficulties in the identification of certain genes which are specific for virulent strains of *E. coli*. Even so, microbial genomics is now an important and growing area of microbiology (Strauss and Falkow 1997). The availability of partial or completed genome sequences from a range of bacteria has helped to assign the phenotype to other previously novel genes. The availability of such genome sequences has been particularly useful in

finding genes homologous to *typA* from other pathogens and non-pathogens and can therefore help indicate the importance of this gene.

It was not possible to clone the putative novel region downstream of *typA* and this may have been due to a number of reasons. There did not appear to be a problem with the actual cloning technique because there was good ligation and transformation efficiency. Cloning of certain genes into high-copy vectors can be troublesome due to resultant over-expression of the gene leading to a lethal effect. However, a low-copy vector, namely pAV3, was also used for this cloning attempt. Alternatively, some genes can have lethal effects if cloned and expressed at more than a single copy per cell, often due to a central role in metabolism. This could be a possibility with a gene contained within the novel region and therefore further cloning attempts could be made using a single-copy vector, perhaps after insertion of a suitable antibiotic resistance marker into the *typA* gene to further aid selectivity. This has now been achieved (Haigh RD, personal communication).

3.3.2 Further work.

Although cloning of the novel region downstream of the *Hind*III site after *typA* in EPEC E2348/69 was unsuccessful, further studies remain to be performed. The size of the novel DNA region at 88 minutes in the EPEC E2348/69 chromosome could be ascertained by pulsed-field gel electrophoresis (PFGE). This method has gained much credibility in recent years for use in mapping bacterial genomes according to restriction analysis and is ideal for separating large DNA fragments by gel electrophoresis (Chu 1990). Recently a rapid PFGE technique has also been described for Gram negative bacteria (Gautom 1997). After choosing the appropriate combination of enzymes, PFGE would allow the size of the novel DNA region to be estimated by DNA hybridisation of the resultant PFGE gel.

Although several attempts at cloning the region downstream of the *Hind*III site at 23550 bp in the 87 minute region of the *E. coli* chromosome were made using two different methods, they were unsuccessful. Instead, standard cloning could be performed but with the addition of antibiotic marker selection. Firstly, the entire

typA gene could be amplified using two primers, each of which would be engineered to include a *Sph*HI site. PCR products often possess an overhanging adenine and this fact can be exploited using the PCR cloning vector pDK101 (Kovalenko *et al* 1994). This vector possesses two *Xcm*I sites which, when subjected to restriction digest with *Xcm*I, leave an overhanging thymine which is complementary to the overhanging adenine of the PCR product. The PCR product and pDK101 may therefore be ligated and transformed into electrocompetent DH5α. After confirmation of a *typA* clone, the fragment could be removed from pDK101 by restriction digest, isolated by agarose gel electrophoresis and purified. Subsequent transfer to the suicide vector pRDH10 may then be achieved by ligation with *Sph*I-digested pRDH10 and transformation into electrocompetent DH5α. The 1 kb kanamycin antibiotic resistance gene from the vector pUC-4K could also be isolated by digestion with *Hinc*II. A *Hinc*II restriction site exists approximately 300 bp into the *typA* gene and therefore, following restriction digest of pRDH10 with *Hinc*II and isolation/purification, the kanamycin gene could be ligated with the *typA* fragment within pRDH10. The ligation product would be transformed into electrocompetent DH5α and, after identification of a *typA*-kanamycin clone, transferred into the chromosome of EPEC E2348/69 by homologous recombination. Standard chromosomal cloning could then be performed but with the added advantage of antibiotic selection.

Chapter 4

Construction and genotypic characterisation of non-polar *typA* mutations in different enteropathogenic *Escherichia coli* strains

4.1 Introduction

4.1.1 Modern mutagenesis systems.

Genetic analysis can reveal substantial data required for the understanding of a bacterial phenotype (deLorenzo and Timmis 1994). To ascribe a phenotype to a novel or uncharacterised gene requires the mutation or inactivation of the gene in question, otherwise known as reverse genetics (Reyrat *et al* 1998). In recent years, a number of new technologies have been described for the identification of virulence genes (Heithoff *et al* 1997). These include signature-tagged mutagenesis (STM) (Hensel *et al* 1995) and *in vivo* expression technology (IVET) (Mahan *et al* 1993; Mahan *et al* 1995). These strategies provide relatively simple and accurate ways of identifying virulence genes *in vitro* and/or *in vivo*. Such strategies have revolutionised virulence gene identification (Hensel and Holden 1996). However, *typA* has not yet been identified as a virulence gene using any of these new technologies. Instead, it is hypothesised that *typA* has a role in virulence due to a number of characteristics of *typA* and its *Salmonella* homologue *bipA* which were determined by two research groups (Qi *et al* 1995; Farris *et al* 1998; Freestone *et al* 1998b). This hypothesis has been partially confirmed by recent data from one of these groups (Farris *et al* 1998). However, these data are deemed inconclusive for two reasons. Firstly, the method of *typA* mutagenesis used, namely transposon mutagenesis, is not ideal for reasons that will be discussed below. Secondly, the strain used, EPEC derivative MAR001, is not a standard wild-type strain and its parentage is not fully known. It is thought to be derived from EPEC E2348/69 with the pMAR2 plasmid cured by SDS culture; however, opinion varies according to a number of researchers (Baldwin T, personal communication; Haigh RD, personal communication). It is therefore important that the role of *typA* in EPEC virulence be

determined using a satisfactory mutagenesis strategy and a standard wild-type EPEC strain such as E2348/69 which has not been genetically modified. This strategy will be discussed in detail below.

4.1.2 Introducing mutations into genes.

There are three principal types of gene mutation that may be created. The first is a deletion mutation in which all or part of the gene is removed. The second is a deletion-substitution mutation in which the deleted part of the gene is replaced by other genetic material. Finally, an insertion mutation occurs by inserting genetic material, for example a transposon (Kleckner *et al* 1977) such as the commonly-used Tn*phoA* (Manoil and Beckwith 1985), into the gene without any of the gene being removed. These types of mutation may result in polar effects on downstream genes, an undesirable effect which, in addition to altering the phenotype of the mutated gene, may also affect the genes downstream (Link *et al* 1997; Reyrat *et al* 1998). Types of mutations may have different effects on different genes due to the complexity of gene expression. The most commonly used method for rapidly mutating a gene is transposon mutagenesis. This method may be used to create a deletion-substitution mutation or insertion mutation and is used in any of the new technologies described above. However, both methods may result in polar effects on downstream genes. Deletion mutation is also rapid and effective but again this may result in polar effects due to changes in DNA topology.

4.1.3 Introducing non-polar mutations into genes.

The most reliable method of creating a gene mutation is by defined point mutations (Ausubel 1995). This type of mutation is least likely to result in polar effects. There should be no or little change in DNA topology and the length of the open reading frame in not significantly changed. This is therefore the most desirable method for creating mutations although it should be noted that other mutations may have no polar effect on the gene in question whilst they may in others. Point mutations are made by the substitution of a single base, or the addition or deletion of a few bases (Foster 1992) thereby resulting in a frameshift which alters the reading frame. Alteration of the reading frame results in the synthesis of a different protein and also

in premature protein termination since the two alternative reading frames possess stop codons which are present to protect against such aberrant protein synthesis. The aim is to cause early premature protein termination.

Two main methods exist for introducing frameshift mutations into a specified gene. The first, inverse PCR mutagenesis (IPCRM), was developed by Wren and colleagues (Wren *et al* 1994) as a rapid method for bacterial gene mutagenesis using subgenic fragments amplified from the chromosome by PCR. IPCRM utilises primers inclusive of restriction sites orientated such that amplification of the template DNA occurs in opposite directions thereby creating a linear DNA molecule. The PCR product is digested with the enzyme chosen within the primer design followed by ligation such that the desired mutation is introduced. The second method is restriction site mutagenesis and has been successfully used in our laboratory in EPEC (Haigh RD, personal communication) and *Salmonella enterica* serovar Typhimurium (Kingsley R, personal communication). It does not necessarily require knowledge of the complete gene sequence and results in the deletion of a few bases, usually resulting in a frameshift mutation. A suitably located restriction site within a cloned gene is digested close to the 5' end of the gene. The ends of the linearised DNA molecule are subjected to nucleotide degradation by Klenow exonuclease followed by blunt-ending of 3' overhanging ends. The blunt ends are subsequently ligated to reform the gene which now contains the desired mutation. As the amount of nucleotide degradation is unknown the mutation is confirmed by sequencing for two reasons; firstly, to confirm degradation and secondly, to confirm that such degradation results in a frameshift. Furthermore, sequencing reveals the predicted length of the resultant protein after mutation. Restriction site mutagenesis was the method chosen for the mutation of *typA*.

4.1.4 Introducing gene mutations into the chromosome.

Cloning of a given gene into a plasmid allows the effect of a mutation in that gene to be studied. However, interpretation of data gained from such studies can be difficult due to factors including plasmid copy number and DNA topology

(Blomfield *et al* 1991). For example, multi-copy expression of a mutated gene may be lethal to the bacterium because expression of its product may have an adverse effect on the expression of other gene products even if the mutation does not result in inactivation of that gene. Therefore, if this approach is used then single-copy plasmids should be used. These factors are in addition to the problems associated with the type of mutation made in an open reading frame, as already discussed. The best method of overcoming plasmid-associated problems is to transfer the mutation into its native location in the chromosome. A few methods exist for the introduction of gene mutations into the bacterial chromosome. These include transposon vectors (Simon *et al* 1983; Taylor *et al* 1989), P1 phage transduction (Merryweather *et al* 1987; Shepherd and Smoller 1994) and allelic exchange vectors (Blomfield *et al* 1991; Selbitschka *et al* 1993; Stibitz 1994). However, if the mutated gene does not possess a counterselectable marker then selection of mutants is difficult. To overcome this problem, Blomfield and colleagues (1991) developed a suicide vector system whereby non-selectable genes can be transferred into the chromosome using a temperature-sensitive plasmid and an intermediate strain containing the *Bacillus subtilis sac*B gene as a counterselectable marker. This is described further below. This strategy has now been used for the introduction of mutations into the chromosome of a number of pathogens including EPEC (Donnenberg and Kaper 1991; Donnenberg *et al* 1993; Kenny *et al* 1996; Ramer *et al* 1996; Lai *et al* 1997; Haigh R, personal communication), *Helicobacter pylori* (Copass M *et al* 1997), *Klebsiella pneumoniae* (Geissler and Drummond 1993), *Pseudomonas aeruginosa* (Schweizer 1992), *Salmonella* (Kingsley R, personal communication) and *Yersinia enterocolitica* (Kaniga *et al* 1991).

4.1.5 Introduction of mutations into chromosomes using a suicide vector.

The introduction of mutations into the bacterial chromosome is now easier since the introduction of suicide vectors. A mutated gene or subgenic fragment cloned in a plasmid vector can be sub-cloned into a suicide vector. The suicide vector should contain not only the *sacB* counterselectable marker but also an antibiotic resistance marker. The mutated subgenic fragment can be introduced into the wild-type copy of the target gene by recombination and selected according to the antibiotic

resistance marker present on the suicide vector. The merodiploid possesses two copies of the gene fragment at this stage and therefore has polar characteristics. In order to resolve the merodiploid a second recombination event is required resulting in one of two scenarios. If the recombination event occurs on the same side as the first event then the resolved merodiploid reverts to wild-type. If the recombination event occurs on the opposite side to the first then the mutation remains in the chromosome. The generation of a resolved merodiploid is, however, a rare event which is selected for by loss of the antibiotic resistance marker on the suicide vector. Selection is very labour intensive although suicide vectors have recently been constructed to overcome this. The vector used in our laboratory, pRDH10 (Haigh R, personal communication), uses a modified derivative of the *B. subtilis sacB* gene, namely *sacRB* (Selbitschka *et al* 1993). Modified and unmodified derivatives of *sacB* have been used previously (Blomfield *et al* 1991, Selbitschka *et al* 1993). The *sacB* or *sacRB* gene, when expressed in *E. coli* in the presence of sucrose and a low sodium chloride concentration at 37°C results in lethality. The system selects for the second recombination event and as such can be used as a counterselectable marker for allelic exchange (Reyrat *et al* 1998). The pRDH10 construct has been used successfully in our laboratory for allelic exchange in EPEC (Haigh RD, personal communication) and *Salmonella enterica* serovar typhimurium (Kingsley R, personal communication).

4.1.6 Construction of a *typA* mutation in EPEC strain E2348/69.

This chapter describes the construction of an EPEC E2348/69 *typA* mutant using the restriction site mutagenesis method followed by allelic exchange using a *sacRB* suicide vector system. The genotypic confirmation of the mutation is also described.

4.2 Results

4.2.1 Transfer of the *typA* fragment into pUC18.

Before mutagenesis, the *typA* fragment was transferred to pUC18 and therefor pSCC01 was digested with *Sal*I in order to release the *typA* fragment. *Sal*I was chosen for a number of reasons; firstly, this would be a useful site for cloning *typA*

into the suicide vector pRDH10 and, secondly, the maximum possible *typA* fragment would be retained. Approximately 1 µg of pSCC01 was digested with *Sal*I at 37°C for three hours to obtain predicted fragment sizes of 6.3 kb, corresponding to the *typA* fragment, and 2.9 kb, corresponding to pBluescript. The fragments were isolated by 1% agarose gel electrophoresis and the 6.3 kb fragment purified by the polyallomer wool/ QIAGEN QIAquick™ PCR purification kit method. The purified 6.3 kb fragment was ligated into *Sal*I-digested pUC18 and the ligation products transformed into DH5α by electroporation. Transformants were selected on LB agar containing 100 µg/ml ampicillin and 40 µg/ml X-gal. Approximately 90% of the transformants that grew on this medium were white, the remaining 10% indicating the presence of undigested pUC18. The presence of the correct insert in the respective transformants was confirmed by plasmid analysis of a dozen transformants using the alkaline lysis extraction method followed by restriction analysis with *Sal*I and analysis by 1% gel electrophoresis. After identification of a correct clone, automated sequencing of the insert was performed for confirmation. This was achieved using the M13 forward and reverse primers with the Perkin-Elmer ABI PRISM™ system. After confirmation the clone was designated pSCC04.

4.2.2 Deletion of four base-pairs in cloned *typA* fragment of pSCC01.

The method chosen for the mutation of *typA* was restriction site mutagenesis, as described above, and is summarised in **figure 4.1**. Restriction digest mapping of *typA* (**section 3.2.2**) indicated that a single *Bst*XI site existed 145 bp downstream of the *typA* start codon. This site was therefore selected for mutagenesis. Initially, approximately 1 µg of pSCC04 was digested with *Bst*XI at 55°C for three hours to create a linear fragment. One tenth of this digestion product was visualised using 1% gel electrophoresis to confirm digestion. To create a deletion, nucleotide deletion at the *typA Bst*XI site in pSCC04 was performed with Klenow exonuclease as described (**section 2.31.1**). Nucleotide degradation was chosen due to the 3' overhang left after the digestion of pSCC04 with *Bst*XI. After degradation, the 3' overhanging ends were filled-in to generate blunt ends and the plasmid religated as previously described. The religated plasmid was transformed into *E. coli* DH5α by

Figure 4.1 Summary of the strategy used for constructing pSCC16.

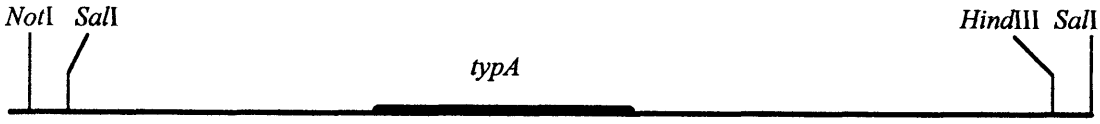
A. The *Not*I-*Hind*III fragment including *typA* from the chromosome of EPEC E2348/69 was cloned into the *Not*I-*Hind*III sites of pBluescript SK-. The construct was designated pSCC01.

B. Two *Sal*I sites were used for transferring the fragment to pUC18 for further analysis. The first site was internal to the *Not*I-*Hind*III fragment just slightly upstream from the *Not*I site. The second was external to the *Not*I-*Hind*III fragment and was within the multi-cloning site of pBluescript SK-. The construct was designated pSCC04.

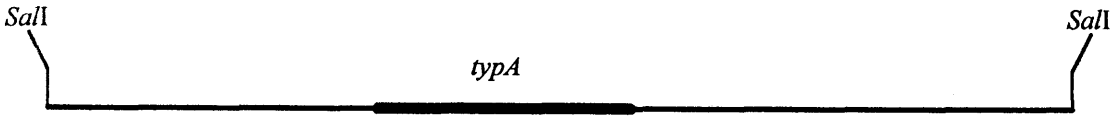
C. A four base-pair deletion was then introduced at the *Bst*XI site of *typA* resulting in a frameshift mutation. The construct was designated pSCC12.

D. The *Sal*I fragment was transferred to the suicide vector pRDH10 and was designated pSCC16.

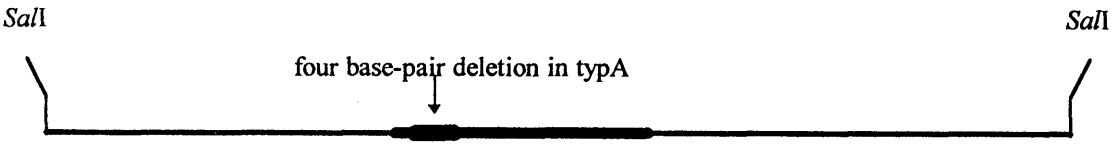
A. pSCC01



B. pSCC04



C. pSCC12



D. Transferred as *Sal*I fragment into suicide vector pRDH10

(pSCC16)

electroporation. Transformants were selected on LB agar containing 100 μg/ml ampicillin and 40 μg/ml X-gal. Although greater than 70% of the transformants were white, inefficient nucleotide degradation was predicted. To overcome the potential problem of screening a great number of transformants, all white colonies on an entire Petri dish (in excess of 10 000 colonies) were taken and the plasmid DNA extracted using the midi-preparation method (**section 2.10.4**). Approximately 1 μg of the plasmid DNA was then digested with *Bst*XI at 55°C for three hours to destroy all those clones which continued to possess a *Bst*XI site. After digestion the DNA was transformed into *E. coli* DH5α by electroporation. Transformants were again selected on LB agar containing 100 μg/ml ampicillin and 40 μg/ml X-gal. A dozen transformants were analysed by alkaline lysis mini-preparation of plasmid DNA followed by restriction analysis with *Bst*XI. The resultant digestion product was analysed using 1% gel electrophoresis. After identification of a clone which did not digest with *Bst*XI, automated sequencing of the insert was performed to confirm the presence of nucleotide deletion. This was achieved using the primer MIR1A with the Perkin-Elmer ABI PRISM™ system. Sequencing revealed that the correct insert was present and that the nucleotide deletion had resulted in a four base-pair deletion (**figure 4.2**); this clone was designated pSCC12. The four base-pair deletion in *typA* results in a frameshift resulting in premature termination of translation. This premature termination was predicted to result in a protein product of 87 amino acids (approximately 10 kDa) as opposed to the typical 592 amino acids (approximately 65 kDa) (**figure 4.3**).

4.2.3 Transfer of the mutated *typA* fragment into the suicide vector pRDH10

Transfer of the *typA* mutation into the suicide vector pRDH10 is necessary in order to facilitate allelic exchange of the mutation into the EPEC chromosome. This was achieved by sub-cloning the *Sal*I fragment containing *typA* from pSCC12 into pRDH10 (**figure 4.1**). Using the *Sal*I fragment, the four base-pair mutation is flanked by 2.7 kb DNA upstream and 3.5 kb DNA downstream. Although it is ideal to have the flanking DNA equally proportioned either side of the mutation to

Figure 4.4 Diagrammatic representation of homologous recombination. The *typA* frameshift mutation was introduced into the chromosome of EPEC strains by homologous recombination using the suicide vector pRDH10.

A The typA gene carrying the frameshift mutation (pSCC12) was conjugated into the recipient strain using the suicide vector (pRDH10).

B Screening for chloramphenicol resistance selected for transconjugants where the plasmid has inserted into the chromosome by homologous recombination. The resultant merodiploids contained one wildtype and one mutant version of the gene as well as the vector sequence between them.

C Growth on sucrose agar at 30°C selected for those bacteria that lost their *sacB* gene due to a second homolgous recombination event.

D The released plasmid was unable to replicate and was rapidly lost. Sucrose resistant clones were screened for the loss of the plasmid resistance marker and then analysed by Southern blot hybridisation for the wild-type or mutant allele.

Figure 4.2 Nucleotide sequence comparison of pSCC12 *typA* and wild-type EPEC E2348/69 *typA* indicating the four base deletion at the *Bst*XI site. Plasmid DNA was prepared from a putative clone which did not digest with *Bst*XI and automated sequencing of *typA* was performed using the primer MIR1A with the Perkin Elmer ABI PRISM™ system. The sequence was manipulated using the Wisconsin GCG molecular biology program which revealed that nucleotide degradation had resulted in a four base-pair deletion approximately 140 bp into *typA*.

```
                .         .         .         .         .
EPEC   80 AACAATCCGGTACGTTCGACTCTCGTGCCGAAACCCAAGAGCGCGTGATG 129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
K-12   80 AACAATCCGGTACGTTCGACTCTCGTGCCGAAACCCAAGAGCGCGTGATG 129
                .         .         .         .         .
      130 GACTCCAA....TTGGAGAAAGAGCGTGGGATTACCATCCTCGCGAAAAA 179
          ||||||||    ||||||||||||||||||||||||||||||||||||||
      130 GACTCCAACGATTTGGAGAAAGAGCGTGGGATTACCATCCTCGCGAAAAA 179
                .         .         .         .         .
      180 CACCGCTATCAAATGGAATGATTACCGTATCAACATCGTTGATACCCCGG 229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
      180 CACCGCTATCAAATGGAATGATTACCGTATCAACATCGTTGATACCCCGG 229
                .         .         .         .         .
      230 GGCACGCCGACTTCGGTGGTGAAGTTGAACGTGTAATGTCCATGGTAGAC 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
      230 GGCACGCCGACTTCGGTGGTGAAGTTGAACGTGTAATGTCCATGGTAGAC 279
                .         .
      280 TCAGTGCTGCTGGTGGTTGAC 300
          |||||||||||||||||||||
      280 TCAGTGCTGCTGGTGGTTGAC 300
```

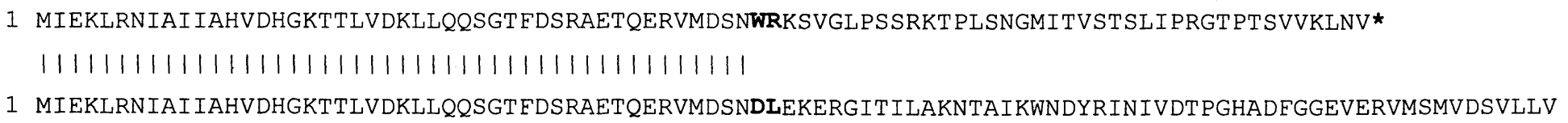


Figure 4.3 Predicted amino acid partial sequence comparison of pSCC12 *typA* and wild-type EPEC E2348/69 *typA* at the region of the four base deletion. The line-up indicates that the four base-pair deletion result in premature termination of *typA* translation as shown by the premature stop codon. The *typA* gene product is predicted to be reduced from 592 amino acids (approximately 65 kDa) to approximately 87 amino acids (approximately 10 kDa).

achieve maximum efficiency during homologous recombination, it was felt that the ratio of 2.7 kb to 3.5 kb would be sufficiently efficient. Approximately 1 μg of pSCC12 was digested with *Sal*I at 37°C for two hours and the resultant DNA isolated by 1% agarose gel electrophoresis. The 6.3 kb fragment was purified using the polyallomer wool/ QIAGEN QIAquick™ PCR purification kit method. The purified 6.3 kb fragment was ligated with 0.5 μg of *Sal*I-digested pRDH10 and the ligation products transformed into *E. coli* strain CC118λ*pir* by electroporation. This strain was used because it encodes the π protein which is essential for the replication of pRDH10. Furthermore, CC118λ*pir* is a better strain than SY327λ*pir* for these purposes as it grows faster and is more easily transformed by electroporation. Transformants were selected on LB agar containing 50 μg/ml chloramphenicol. Approximately 10^8 colonies were isolated on this medium. A representative 12 transformants were patch-plated onto two LB agar plates, one containing 50 μg/ml chloramphenicol, the other containing 100 μg/ml ampicillin and 20 μg/ml tetracycline. The chloramphenicol agar was used to select for those transformants possessing pRDH10. In addition, the tetracycline was used to select against those transformants that possessed pRDH10 but did not possess an insert. The ampicillin was used to select against the presence of pSCC12 or other contaminating ampicillin-resistant vector used during mutant construction. One transformant was found to be chloramphenicol resistant and ampicillin/tetracycline sensitive indicating the presence of pRDH10 with an insert. The transformant was plated onto three separate LB agar plates containing chloramphenicol, ampicillin or tetracycline for confirmatory purposes. Growth on these plates was as expected; therefore the transformant grew on chloramphenicol only. To confirm the presence of the correct insert with a four base-pair deletion in *typA*, plasmid DNA was extracted using the Qiagen® QIAprep Spin method. Approximately 0.5 μg of plasmid DNA was digested with *Sal*I at 37°C for one hour and analysed by 1% gel electrophoresis. Two bands were visualised; one of 6.3 kb corresponding to the *typA* *Sal*I fragment and one of 8.2 kb corresponding to pRDH10. Automated sequencing of the insert was performed using primer MIR1A with the Perkin-Elmer ABI PRISM™ system. This confirmed that the *Sal*I fragment containing *typA* from

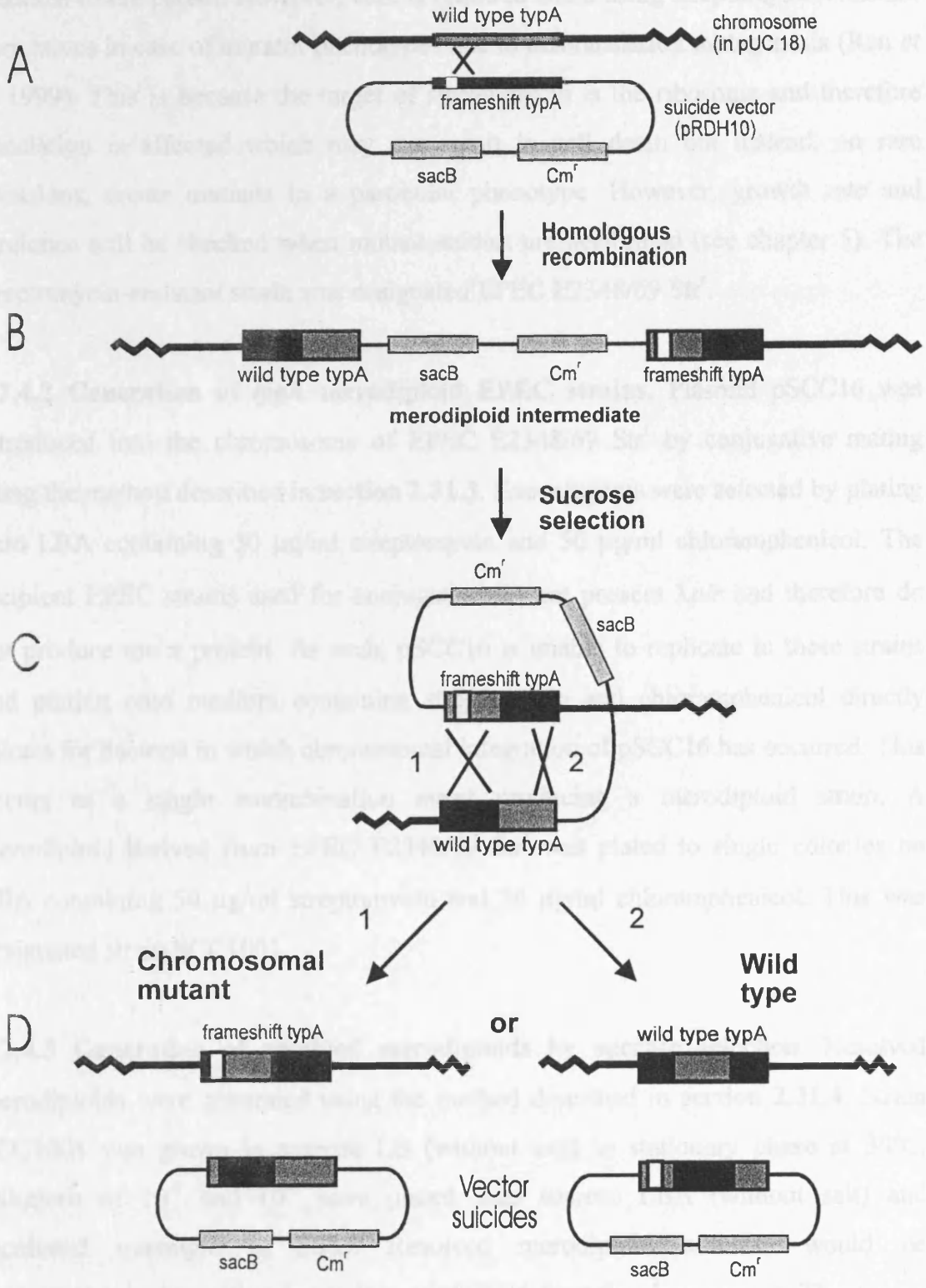
pSCC12 had been successfully transferred into pRDH10. The clone was designated pSCC16.

The next stage of the mutant construction was to transfer the *typA* mutation into the chromosome of one or more wild-type EPEC strains. However, the *E. coli* strain CC118λ*pir* is unable to mobilise pRDH10 by conjugation. Therefore it was necessary to transfer pSCC16 from CC118λ*pir* to the *E. coli* strain SM10λ*pir*; this strain encodes the π protein and also possesses a chromosomally-integrated RP4 plasmid enabling conjugative transfer of pSCC16 into wild-type EPEC. Using the Qiagen® QIAprep Spin method, pSCC16 was extracted from CC118λ*pir* and transformed into *E. coli* strain SM10λ*pir* by electroporation. As SM10λ*pir* is kanamycin resistant, transformants were selected on LB agar containing 25 μg/ml kanamycin and 50 μg/ml chloramphenicol. A representative transformant was taken and used in the next stage detailed below.

4.2.4 Introduction of the *typA* mutation into the chromosome of EPEC.

The four base-pair mutation in *typA* was introduced into the chromosome of EPEC E2348/69 Str^r, and later into EPEC strains MAR001 and JPN15, using the strategy described in **sections 2.31.2** to **2.31.5** of the Materials and Methods, and also summarised in **section 4.1.5** and **figure 4.4**.

4.2.4.1 Isolation of a spontaneous streptomycin-resistant variant of EPEC strain E2348/69. Conjugative mating of SM10λ*pir* containing pSCC16 with wild-type EPEC requires the presence of an antibiotic resistance marker in the latter strain for the selection of conjugants possessing a chromosomally-integrated plasmid. The wild-type EPEC strain E2348/69 did not possess a useful antibiotic resistance marker. Therefore a spontaneous streptomycin-resistant variant of this strain was selected using the method described in **section 2.30**. Spontaneous high-level streptomycin resistance occurs via a single-step mutation in one of the ribosomal proteins (Garrod *et al* 1981). Although some streptomycin-resistant strains have reduced growth rate, and sometimes reduced virulence, most are phenotypically


A
wild type typA
chromosome
(in pUC 18)
frameshift typA
suicide vector
(pRDH10)
sacB
Cm^r
Homologous
recombination
B
wild type typA
sacB
Cm^r
frameshift typA
merodiploid intermediate
Sucrose
selection
C
Cm^r
sacB
frameshift typA
1
2
wild type typA
1
2
Chromosomal
mutant
Wild
type
D
frameshift typA
or
wild type typA
Vector
suicides
sacB
Cm^r
sacB
Cm^r

identical to the parent. However, care is required when using streptomycin-resistant derivatives in case of mutator phenotypes due to mistranslation mutagenesis (Ren *et al* 1999). This is because the target of streptomycin is the ribosome and therefore translation is affected which may not result in cell death but instead, on rare occasions, create mutants in a particular phenotype. However, growth rate and virulence will be checked when mutant studies are performed (see chapter 5). The streptomycin-resistant strain was designated EPEC E2348/69 Str^r.

4.2.4.2 Generation of *typA* merodiploid EPEC strains. Plasmid pSCC16 was introduced into the chromosome of EPEC E2348/69 Str^r by conjugative mating using the method described in **section 2.31.3**. Exconjugants were selected by plating onto LBA containing 50 μg/ml streptomycin and 50 μg/ml chloramphenicol. The recipient EPEC strains used for conjugation do not possess λ*pir* and therefore do not produce the π protein. As such, pSCC16 is unable to replicate in these strains and plating onto medium containing streptomycin and chloramphenicol directly selects for bacteria in which chromosomal integration of pSCC16 has occurred. This occurs as a single recombination event producing a merodiploid strain. A merodiploid derived from EPEC E2348/69 Str^r was plated to single colonies on LBA containing 50 μg/ml streptomycin and 50 μg/ml chloramphenicol. This was designated strain SCC1001.

4.2.4.3 Generation of resolved merodiploids by sucrose selection. Resolved merodiploids were generated using the method described in **section 2.31.4**. Strain SCC1001 was grown in sucrose LB (without salt) to stationary phase at 30°C. Dilutions of 10^{-5} and 10^{-6} were plated onto sucrose LBA (without salt) and incubated overnight at 30°C. Resolved merodiploid colonies would be chloramphenicol sensitive due to loss of pSCC16 from the chromosome. These were therefore identified by patch-plating onto LBA containing 50 μg/ml streptomycin and LBA containing 50 μg/ml streptomycin and 50 μg/ml chloramphenicol respectively. Resolved merodiploids were plated to single colonies on LBA containing 50 μg/ml streptomycin.

4.2.4.4 Screening of resolved merodiploids for retention of the *typA* mutation. Resolved merodiploids were screened using the method described in **section 2.31.5.** As described previously, the second recombination event resulting in loss of pSCC16 from the chromosome can result in one of two scenarios. This depends on the side of cross-over resulting in revertion to wild-type or generation of the mutation. Identification of resolved merodiploids which had retained the mutation was achieved by showing the absence of the *Bst*XI site in the *typA* gene. Colony boilates were made of each resolved merodiploid by taking 5-10 colonies, suspending them in 100 µl of sterile distilled water and boiling the suspension for 10 minutes. The *typA* gene was then amplified by PCR using primers MIR1A and MIR13 in the following recipe:

Boilate DNA	1µl
MIR1A	3µl
MIR13	3µl
7mM buffer	1µl
dNTPs	1µl
dH_2O	5µl
Taq polymerase	0.05µl

PCR conditions were 95°C for 30 seconds, 60°C for 30 seconds and 72°C for 1 minute. Each PCR reaction was performed in duplicate. The first product was separated using 2% gel electrophoresis to confirm that the reaction had worked whilst the second product was digested with *Bst*XI at 55°C for four hours to determine the presence or absence of a *Bst*XI site. The expected PCR product size was 1.2 kb. The *Bst*XI site in wild-type *typA* is present at 150 bp downstream of the start codon. Therefore, if the *typA* product is digested with *Bst*XI then no shift in product size would be seen in mutant *typA* whereas a decrease in size to 1.05 kb would be seen in wild-type *typA*. It was found that the PCR amplification resulted in two DNA fragments being produced. The first was approximately 1.2 kb as

expected whilst the second was a non-specific 400 bp fragment. This fragment, however, was not digested with *Bst*XI and did not affect analysis of the 1.2 kb fragment. Of 40 resolved merodiploids screened, one was found not to possess a *Bst*XI site. The low frequency of recombination of the mutated fragment was probably due to the mutation being slightly off-centre in the cloned fragment. This would result in a lower recombination efficiency. This resolved merodiploid was designated SCC1002.

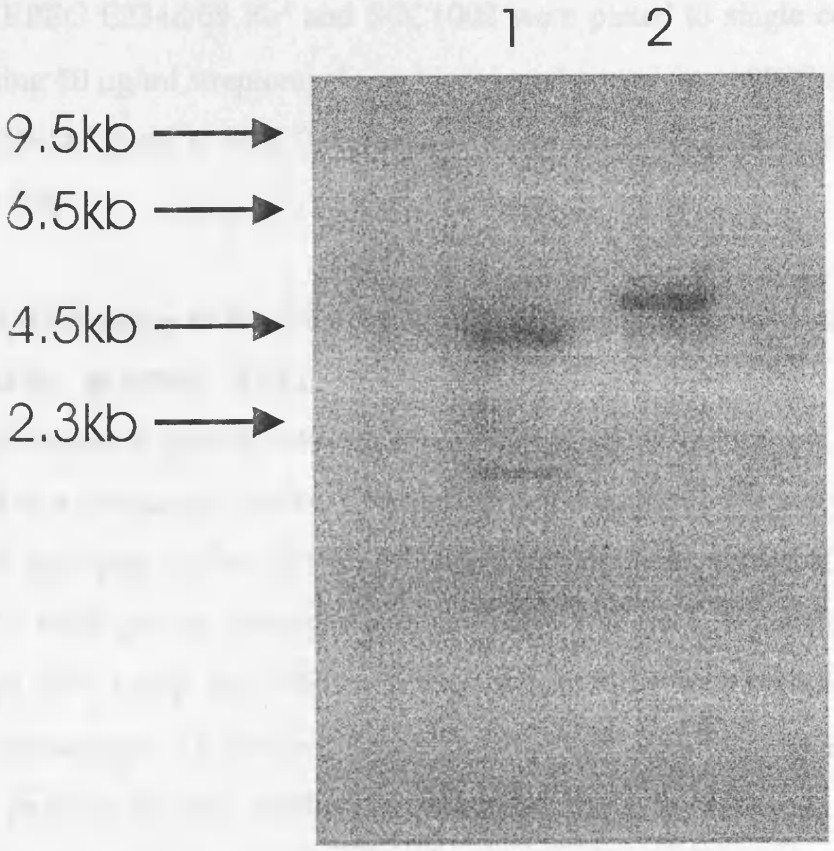
4.2.5 Confirmation of the introduction of the *typA* mutation into strain SCC1002 by Southern blot hybridisation.

The introduction of the *typA* mutation into the chromosome of SCC1002 was confirmed by Southern blot hybridisation analysis. This was achieved by preparing chromosomal DNA from strains EPEC E2348/69 Str^r and SCC1002. Approximately 3 μg of chromosomal DNA from each strain was digested with *Bam*HI for four hours at 37°C followed by digestion with *Bst*XI overnight at 55°C. The DNA was then separated by 1% gel electrophoresis and the DNA transferred to nylon membrane by the Southern blot method. The DNA was probed with *typA* which was amplified by PCR using primers MIR1A and MIR13 using the method in **section 3.2.1.1** and detected with the Amersham Life Science Fluorescein Gene Images™ labelling and detection system. Expected DNA fragments of 3.4 kb and 1.6 kb were visualised in the lane corresponding to DNA from strain EPEC E2348/69 Str^r thus indicating the presence of a *Bst*XI site within the *typA* gene (**figure 4.5**). A DNA fragment of 5 kb was visualised in the lane corresponding to DNA from strain SCC1002 indicating that no *Bst*XI site was present in *typA*. This result confirmed the presence of a *typA* mutation in strain SCC1002 at the region corresponding to the *Bst*XI site in wild-type strains.

4.2.6 Confirmation of the introduction of the *typA* mutation into strain SCC1002 by DNA sequencing.

Although the Southern blot hybridisation analysis performed above confirms the presence of a *typA* mutation in the chromosome of strain SCC1002, it does not confirm that the mutation is not spontaneous. Such a mutation may not result in a

Figure 4.5 Southern blot hybridisation analysis of chromosomal DNA from strains SCC1001 and SCC1002 probed with *typA*. Chromosomal DNA from strain SCC1001 (lane 1) and SCC1002 (lane 2) was prepared and digested with *Bam*HI and *Bst*XI (**see section 3.2.4**). The resultant DNA fragments were separated using 1% gel electrophoresis, transferred to nylon membrane by Southern blot and hybridised with *typA*. The blot demonstrates the presence of a *Bst*XI site in strain SCC1001 and the absence of a *Bst*XI site in strain SCC1002. This indicates the presence of a mutation at the *Bst*XI site of typA in strain SCC1002.



frameshift mutation and the translation of *typA* may not be affected. Sequencing of a *typA* PCR product was therefore performed. This was achieved by amplifying *typA* as described in **section 4.2.4.4** followed by automated sequencing as described in **section 2.28** using primer MIR1A. Sequencing confirmed the presence of a four base-pair deletion at the expected site which was identical to that introduced into *typA* in pSCC12 (**figure 4.2**).

4.2.7 Growth of strain SCC1002 at ambient temperature and at 37°C.

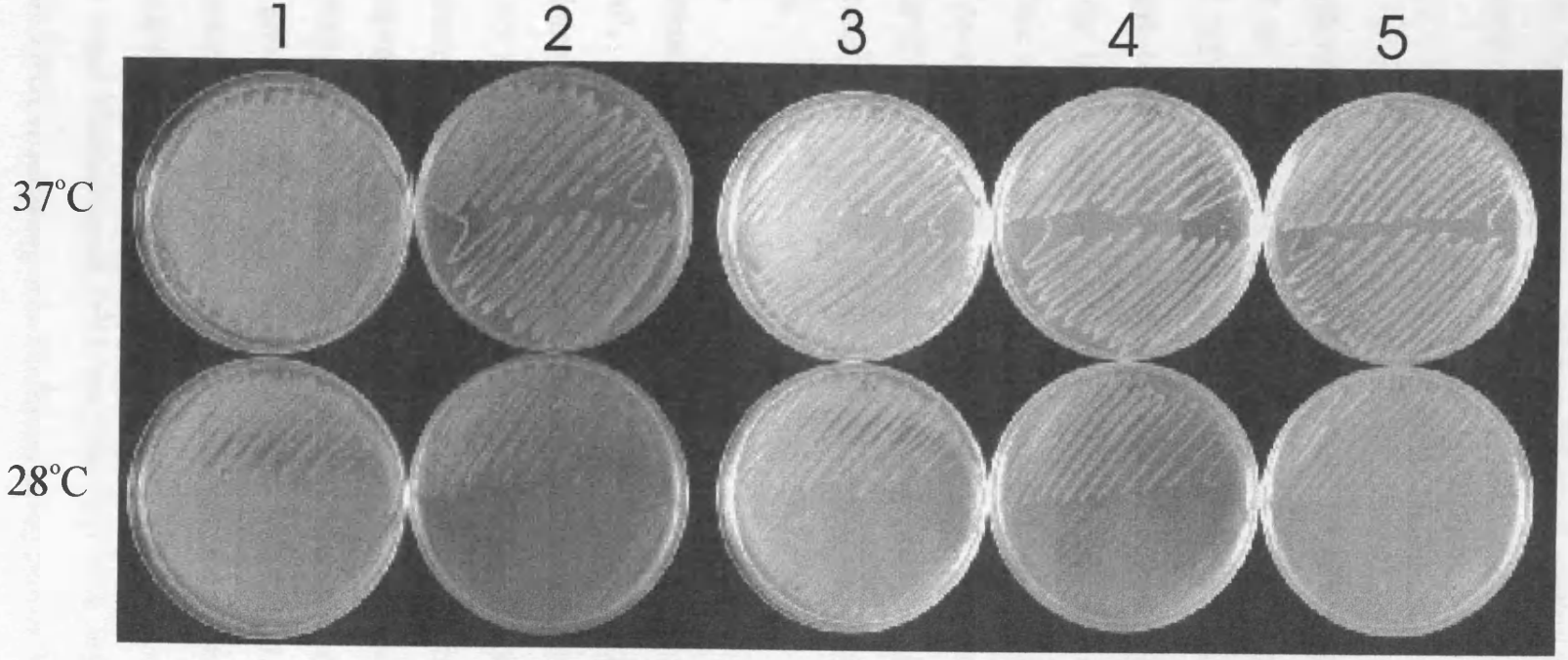
During the isolation of strain SCC1002 there was no known phenotype to enable simple screening for the mutant. However, during isolation it was noted that the strain did not grow at ambient temperatures (approximately 28°C). To confirm this, strains EPEC E2348/69 Str^r and SCC1002 were plated to single colonies on LBA containing 50 μg/ml streptomycin and incubated overnight at 28°C and 37°C. EPEC E2348/69 Str^r grew at both 28°C and 37°C whereas SCC1002 grew only at 37°C (**figure 4.6**).

4.2.8 Introduction of the *typA* frameshift mutation into strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1.

The differences in growth between strains EPEC E2348/69 Str^r and SCC1002 could be used as a phenotypic marker. Therefore, to determine if this had potential for use in other wild-type strains, EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 were grown overnight at 28°C and 37°C on LBA containing 50 μg/ml nalidixic acid, except for EPEC MAR001 which was grown on LBA containing 25 μg/ml kanamycin. All strains grew at both temperatures. As the above phenotype would provide an easy method of screening for *typA* resolved merodiploids, a decision was made to transfer the mutation into other EPEC strains and RDEC-1. The *typA* mutation was introduced into the chromosomes of the above strains using the method described above for the introduction of the mutation into EPEC E2348/69 Str^r. The rationale for this was that two of these derivatives, namely MAR001 and JPN15, do not possess the EAF virulence plasmid which possesses genes encoding bundle-forming pili. Furthermore, the wild-type strains of E2348/69

Figure 4.6 Photographs of SCC1002, SCC1003, SCC1004 and SCC1005 with their respective wild-type parents grown on LBA at 37°C and 28°C. Each strain was inoculated onto LBA half-plates so that the mutant was positioned opposite its wild-type parent. The plates were incubated at 37°C and 28°C to determine the effect of temperature on growth. The upper half of each plate is the wild-type strain whilst the lower half is the *typA* mutant strain.

Upper half	Lower half
1. SCC1001	SCC1002
2. E2348/69 Nal^r	SCC1003
3. MAR001	SCC1004
4. JPN-15	SCC1005
5. RDEC-1	SCC1006


1
2
3
4
5
37°C
28°C

possess different antibiotic resistance markers and it would be useful to check that different markers do not affect the virulence phenotype when a *typA* mutation is introduced. A *typA* mutation in the strain RDEC-1 would also be useful if experiments on fluid secretion were to be performed using a rabbit ileal loop test. The latter would provide *in vivo* evidence of the effects of a *typA* mutation on virulence.

4.2.8.1 Growth of strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 at ambient temperature and at 37°C. Resolved merodiploids of strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 were plated to single colonies on duplicate LBA plates containing 50 μg/ml nalidixic acid, except for EPEC MAR001 which was grown on LBA containing 25 μg/ml kanamycin. One of each LBA plate was incubated overnight at 28°C and 37°C respectively. An average of six resolved merodiploids for each strain grew at 37° but not at 28°C suggesting that the phenotypic selection procedure had been successful. These were selected for further analysis to confirm the presence of the *typA* mutation.

4.2.8.2 Confirmation of the introduction of the *typA* mutation into strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1. Attempts to confirm the introduction of the *typA* mutation into the chromosomes of strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 were made by PCR and Southern blot hybridisation. This was performed as for the confirmation of the *typA* mutation in strain SCC1002. According to PCR amplification of *typA* followed by *Bst*XI digestion of the DNA product, all resolved merodiploids gave an amplification product of 1.2 kb and did not digest with *Bst*XI. This suggested that the resolved merodiploids possessed the four base-pair deletion in *typA*. However, the DNA fragments visualised by Southern blot hybridisation were not of the expected size when chromosomal DNA was digested with *Bst*XI and probed with *typA*. Therefore DNA sequencing was attempted using primer MIR1A as before but no sequence data were obtained from any of the strains even after repeating. This could be due to the *typA* sequence in these strains being sufficiently different from

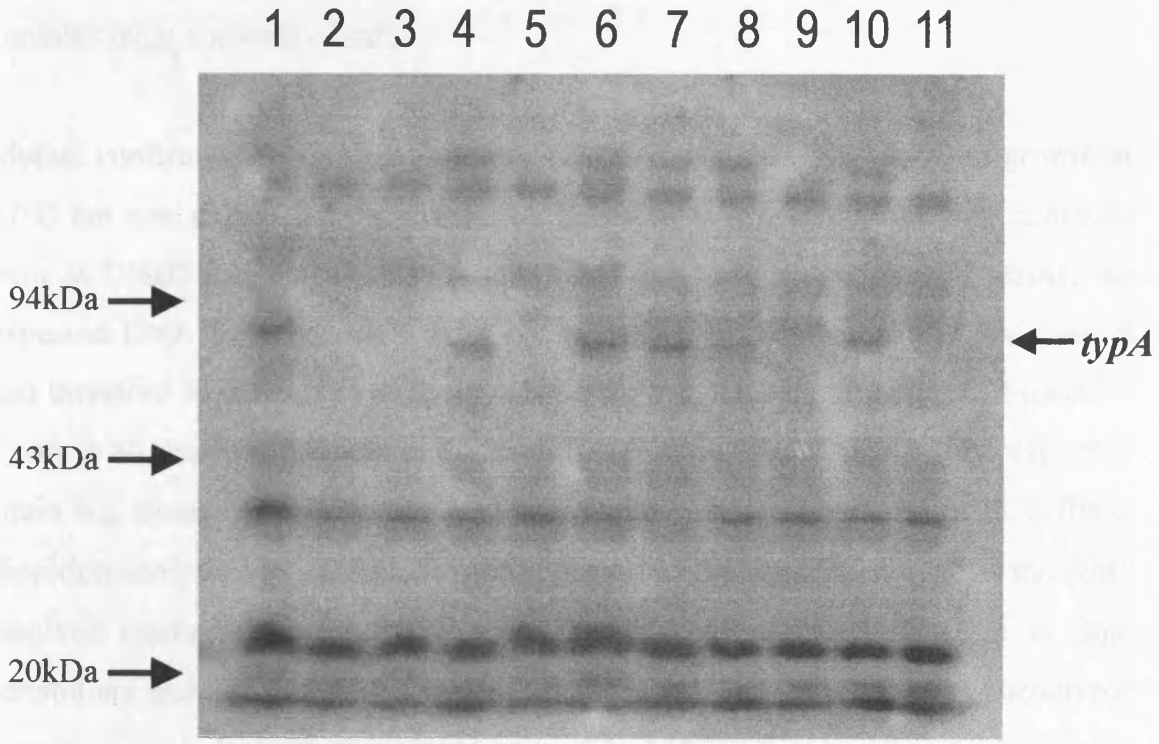
that of E2348/69 Str^r that primer MIR1A was unable to anneal. This would also explain the Southern blot hybridisation results where fragment sizes different from those expected were visualised. To confirm the *typA* mutations by another method, Western blot analysis was performed instead. Whole fractions of wild-type strains of EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 and their corresponding resolved merodiploids were used for this analysis. 5 ml LB cultures of each strain were grown at 37°C to an OD of 1.0 at 600nm. The bacterial growth was processed according to the method in **section 2.39** and the Western blot probed with *typA*-specific antibodies. These antibodies were rabbit-derived and produced at the University of Leicester (Freestone P, personal communication). A pre-immune bleed was performed to check that the rabbit did not previously possess antibodies to *typA*. The blot indicated that *typA* was present in all wild-type strains and absent from all resolved merodiploids except MAR001 (**figure 4.7**). Although the genotype of this last resolved merodiploid was dubious it was decided not to attempt making another merodiploid due to time constraints. The resolved merodiploids from strains EPEC E2348/69 Nal^r, EPEC MAR001, EPEC JPN15 and RDEC-1 were designated SCC1003, SCC1004, SCC1005 and SCC1006 respectively.

4.3 Discussion

The construction of a non-polar mutation in a putative virulence gene is necessary for proving its role in virulence. This chapter has described the construction of a non-polar mutation in the *typA* gene of various EPEC strains. The method used for introducing the mutation into the EPEC chromosome involved the use of a counterselectable marker, namely the *Bacillus subtilis sacB* gene. This has been used successfully in EPEC for proving the role of a number of EPEC genes in virulence including *eaeA* (Donnenberg and Kaper 1991), *espB* (Donnenberg *et al* 1993), *espA* (Kenny *et al* 1996), *bfpB* (Ramer *et al* 1996; Anantha *et al* 1998a), *espD* (Lai *et al* 1997), *bfpF* (Anantha *et al* 1998b), *espF* (McNamara and Donnenberg 1998) and most recently *bfpC*, *bfpH* and *bfpU* (Anantha *et al* 1998a). This method of mutant construction is therefore widely accepted.

Figure 4.7 Western blot hybridisation analysis of resolved merodiploids with anti-TypA antibodies. Whole protein fractions were prepared of EPEC E2348/69 Nal^r, MAR001, JPN15 and RDEC-1 and their corresponding resolved merodiploids from 5ml LB cultures grown to an OD of 1.0. The bacterial growth was processed according to the method in **section 2.39** and the Western blot probed with TypA-specific antibodies. The blot indicated that TypA was expressed in all wild-type strains and absent from all resolved merodiploids except MAR001.

Lane 1	SCC1001	Lane 2	SCC1002
Lane 3	SCC1002	Lane 4	EPEC E2348/69 Nal^r
Lane 5	SCC1003	Lane 6	MAR001
Lane 7	SCC1004	Lane 8	JPN15
Lane 9	SCC1005	Lane 10	RDEC-1
Lane 11	SCC1006		



Non-polar *typA* mutations were made in EPEC strains E2348/69 Str^r and E2348/69 Nal^r, in EPEC derivatives MAR001 and JPN15, and also in RDEC-1. These strains have been designated SCC1002, SCC1003, SCC1004, SCC1005 and SCC1006 respectively. The mutation introduced into the chromosome of these strains was a four base-pair *typA* deletion leading to premature translation termination and a predicted protein product of 87 amino acids (approximately 10 kDa). As the usual *typA* product is 592 amino acids (approximately 65 kDa), this was a major reduction in protein product. The *typA* mutation was confirmed in all of the above strains by PCR and subsequent digestion of the DNA product with *Bst*XI.

During mutant construction it was noted that the EPEC E2348/69 *typA* mutant did not grow at 28°C. Such a phenotype would not confer an advantage upon a bacterial strain and may be an indirect result of another environmental factor. For example, some bacteria are more sensitive to salt at lower temperatures and therefore could be the factor in this EPEC mutant. Further work on this strain, and the other *typA* mutants, at 28°C and 37°C would identify the factor(s) involved. Growth comparisons could easily be made in Luria broth and DMEM, the former of which contains large amounts of salt.

Mutant confirmation was therefore also achieved using the phenotype of growth at 37°C but not at 28°C. Confirmation by Southern blot hybridisation was achieved with SCC1002 and not with the other strains due to spurious fragment sizing; the expected DNA fragment sizes were not visualised by Southern blot hybridisation and therefore Western blot analysis was performed. This confirmed the absence of TypA in all resolved merodiploids except EPEC derivative MAR001. However, this strain had already been shown to possess a mutation at the *Bst*XI site by PCR/*Bst*XI digestion analysis. In addition, the strain possessed the same phenotype as the other resolved merodiploids in that it grew at 37°C but not at 28°C. Due to time constraints and the fact that it would be used for a limited number of phenotypic assays, it was decided to proceed with this MAR001 *typA* mutant, designated SCC1004, and not to construct a new merodiploid strain.

Chapter 5

Phenotypic characterisation of non-polar *typA* mutations in different enteropathogenic *Escherichia coli* strains

5.1 Introduction

In prokaryotes, tyrosine phosphorylation has mostly been described in non-pathogens to date. Recent data suggest, however, that tyrosine phosphorylation of one particular bacterial protein, namely TypA, may be important in the virulence of EPEC by affecting its ability to cause the characteristic attaching and effacing lesion (Farris *et al* 1998; Freestone *et al* 1998b). In order to determine whether *typA* is an important virulence determinant, *typA* mutants of various EPEC strains have been constructed, as described in chapter 4. It is now necessary to perform various phenotypic assays on these mutant strains to show the effect of the *typA* mutation.

There are numerous phenotypic assays that could be performed on a bacterial strain but this thesis is concerned only with those that indicate the virulence potential of the strain. In other studies, these have included *in vitro* and *in vivo* assays depending on the characteristic virulence determinants of the organism, such as those for adhesion, invasion and the production of secreted proteins. These assays have previously been performed on various enteric pathogens, such as salmonella (Wood *et al* 1998), EPEC (Jerse *et al* 1990; Gómez-Duarte and Kaper 1995; Rosenshine *et al* 1996) and campylobacter (Doig *et al* 1996) to determine the effects of gene mutations on their virulence. In EPEC, mammalian cell culture and the alkaline phosphatase assay have been used for studying AE expression (Jerse *et al* 1990; Gómez-Duarte *et al* 1995), although other assays have also been used to study other aspects of EPEC virulence. Experiments have also been performed *in vitro* on human paediatric tissue (Hicks *et al* 1998) and *in vivo* on human volunteers (Donnenberg *et al* 1993; Donnenberg *et al* 1998). The phenotypic analysis of EPEC mutants has revealed a large amount of information in recent years (Nataro and

Kaper 1998). This has led to a much greater understanding of the mechanisms underlying the virulence of EPEC and therefore other gastrointestinal pathogens by revealing that similar virulence mechanisms are possessed by different bacterial genera and species. The locus of enterocyte effacement and bundle-forming pili are the two main EPEC virulence traits which have been studied in great detail and where mutant analysis has been performed. Although the actual role of Bfp in the virulence of EPEC remains controversial with the possibility that other adherence factors exist (Haigh R, personal communication), it is clear that a failure to produce or a reduced expression of Bfp leads to an inability or reduced ability of EPEC to adhere to HEp-2 cells (Donnenberg *et al* 1992).

This chapter will therefore be concerned with the phenotypic characterisation of the *typA* mutation in the EPEC strains chosen in chapter 4. Due to the fact that there were problems associated with creating the initial frameshift mutation of *typA* and therefore time limitations, a limited number of phenotypic assays will be chosen. The two main areas which may be exploited for determining the virulence potential of EPEC strains are adherence/invasion and secreted protein production. As work on *typA* has already suggested a role for *typA* in production of the AE lesion, either area could be chosen for analysis as both are required for full AE lesion production. However, the production of the lesion is initially dependent upon adherence of the EPEC bacterium to the host cell and it therefore made sense to study this area first. In addition, it was possible that further work could follow on from these assays in a logical manner if the adherence/invasion phenotype was found to be affected by the *typA* mutation.

Farris and colleagues (1998) have shown that a *typA* mutant of EPEC derivative MAR001 does not cause the classic attaching and effacing phenotype. However, MAR001 does not possess the EAF virulence plasmid and is of unknown parentage. The data gained from the phenotypic analysis of this mutant must therefore be interpreted with caution. Chapter 4 described the construction of five *typA* mutant strains. Two of these were in EPEC derivatives MAR001 and JPN15. These derivatives are thought to possess identical genotypes because they both lack the

EAF virulence plasmid and therefore do not express Bfp. As such they cannot adhere to cultured epithelial cells in a localised pattern. The parentage of MAR001 is unclear whereas it is known that EPEC strain JPN15 is derived from EPEC E2348/69 Nal^r. Therefore the phenotype of these derivative mutants can be compared both with the *typA* mutants of wild-type EPEC and with each other to demonstrate any differences in genotype. Any discrepancy between the phenotypes of the MAR001 and JPN15 *typA* mutants would be apparent and the effect of the absence of the EAF plasmid would also be noted. The results presented in this chapter therefore detail the phenotypic assays performed on the EPEC *typA* strains/mutants constructed in chapter 4. All of these assays were performed *in vitro* and concentrate on particular virulence determinants associated with EPEC.

5.2 Results

5.2.1 Growth of strains SCC1002, SCC1003, SCC1004, SCC1005 and SCC1006 in LB and DMEM at 37°C.

The phenotypes of different bacterial strains or species can be compared by performing various assays as discussed. In order to analyse and interpret these assays, however, it is necessary to determine the growth rate of each strain or species to check that the growth rates are similar. The phenotypic characterisation of the EPEC E2348/69 $Strep^r$ *typA* mutant, SCC1002, included in later sections assays for adherence, gentamicin protection (invasion), Bfp expression and AE lesion formation (FAS test). It was therefore imperative that the growth rate of the wild-type strain was compared with the mutant before any conclusions could be drawn about the results from these phenotypic assays. If any difference in growth rate existed then this would need to be taken into account when interpreting assay data.

EPEC *typA* mutant strains SCC1002, SCC1003, SCC1004, SCC1005 and SCC1006 and their respective parent strains were inoculated into 5 ml of LB containing 50 μg/ml streptomycin, 25 μg/ml kanamycin or 50 μg/ml nalidixic acid depending on the strain, and incubated overnight at 37°C with shaking. After incubation the OD of

the culture was measured at 600 nm using a Pharmacia Biotech Ultrospec 2000 spectrophotometer. Each culture was diluted approximately 1/50 in 50 ml of pre-warmed LB and DMEM to give similar starting ODs and incubated at 37°C, without the addition of antibiotics, with shaking at 270 rpm. OD readings were taken at time zero and every 30 minutes until three similar readings were taken indicating that the cultures had reached stationary phase. The experiments was repeated three times and average readings recorded.

Using the results obtained from the growth assay, growth curves were produced for each strain by using the values from each OD reading. For each medium, the growth curves of all strains were plotted together so that the growth rates could be easily compared (**figures 5.1 and 5.2**). The growth curves in LB indicate that each *typA* mutant strain grew at a similar rate to its parent, and shows that the *typA* mutation did not adversely affect the growth rate of the EPEC strains used in this study. The same was seen for the growth curves in DMEM except for MAR001; the *typA* mutant SCC1004 grew faster than its parent. Strain SCC1004 was therefore not used for further phenotypic assays as its growth rate differed from its parent. This was the mutant which could not be confirmed as possessing a *typA* mutation (see **section 4.2.8.2**). Except for strain SCC1004, further phenotypic assays may therefore be performed without the necessity of accounting for differences in growth rate. However, strain SCC1003 was also not used for phenotypic assays elsewhere in this chapter as its growth rate was identical to that of SCC1002, their parent was of the strain serotype and duplicated strain results are not required. Furthermore, recent data suggest that EPEC derivative E2348/69 Nal^r contains one or more mutations, in addition to the *gyrA* mutation responsible for nalidixic acid resistance, which appears to affect Bfp expression (Clarke SC, Goldberg M and Haigh R, unpublished observations).

5.2.2 Microscopic observation of bundle-forming pilus production in DMEM using strain SCC1002.

Bfp production can be visualised microscopically in culture medium by observing clumping in a localised manner, similar to that seen during adherence to cultured

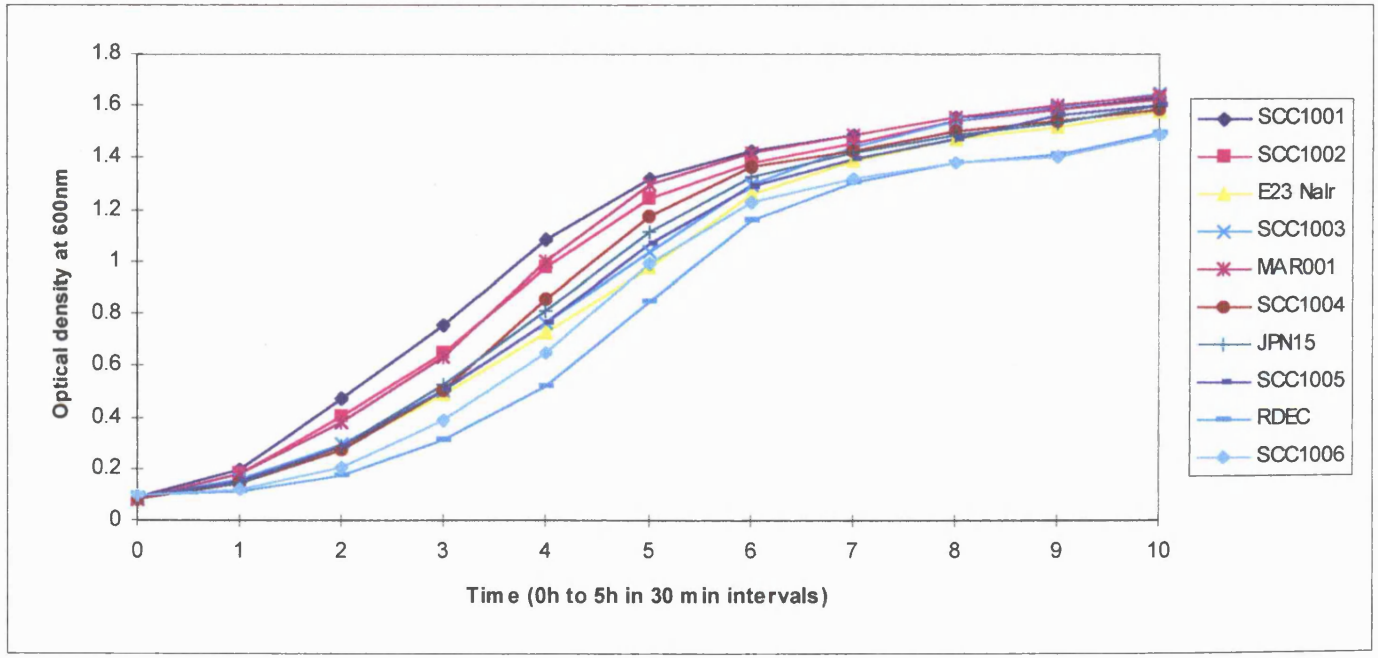


Figure 5.1 Growth curves of strains SCC1002, SCC1003, SCC1004, SCC1005, SCC1006 and their respective wild-types in Luria broth. Each strain was grown overnight at 37°C in LB in the presence of the appropriate antibiotic to stationary phase. Each strain was then diluted approximately 1/50 in 50 ml of pre-warmed LB to give similar starting ODs and incubated at 37°C without the addition of antibiotics with shaking at 270 rpm. OD readings at 600 nm were taken at time zero and every 30 minutes until three similar readings were noted indicating that the cultures had reached stationary phase.

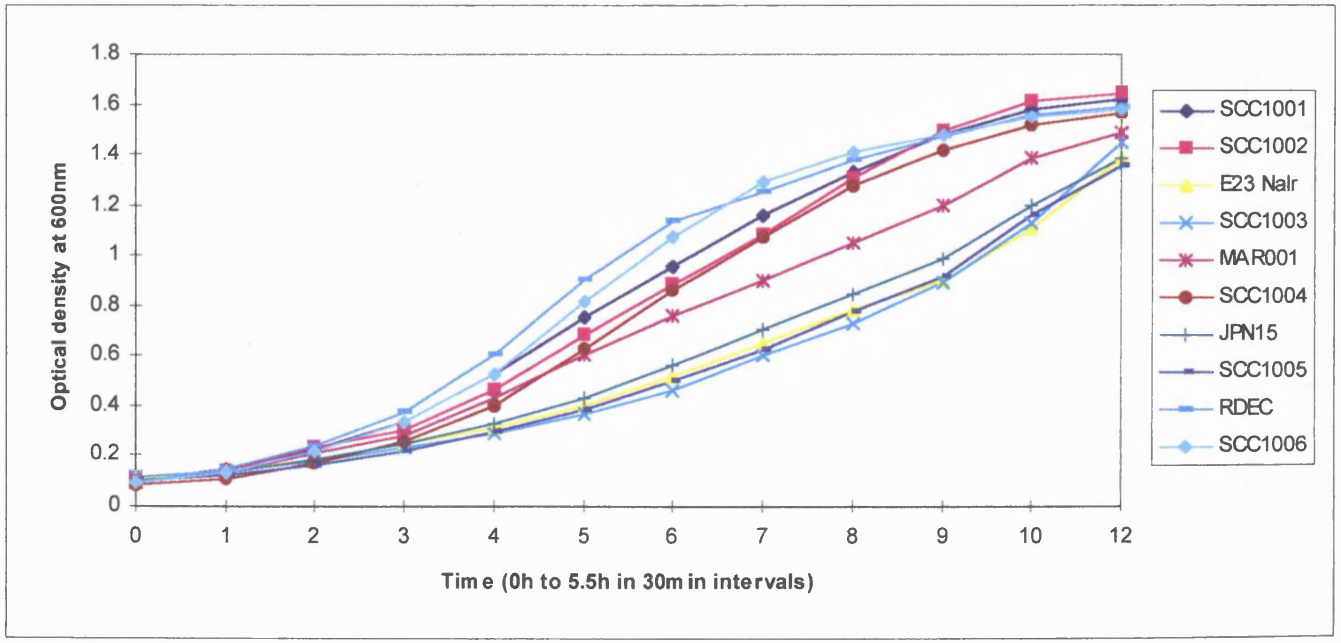


Figure 5.2 Growth curves of strains SCC1002, SCC1003, SCC1004, SCC1005, SCC1006 and their respective wild-types in DMEM. Each strain was grown overnight at 37°C in LB in the presence of the appropriate antibiotic to stationary phase. Each strain was then diluted approximately 1/50 in 50 ml of pre-warmed DMEM to give similar starting ODs and incubated at 37°C without the addition of antibiotics with shaking at 270 rpm. OD readings at 600 nm were taken at time zero and every 30 minutes until three similar readings were noted indicating that the cultures had reached stationary phase.

cells. However, microscopic observation of Bfp production is not specific and should be interpreted with care. Only two of the wild-type strains used in this study possessed Bfp, namely EPEC E2348/69 Str^r (SCC1001) and EPEC E2348/69 Nal^r. Derivatives MAR001 and JPN15 do not possess the EAF virulence plasmid and therefore do not produce Bfp. Strain RDEC-1 also does not produce Bfp. Since it was hypothesised that *typA* may be involved in virulence, the effect of the *typA* mutation on Bfp expression by strain SCC1002 was tested.

Single colonies from pure cultures of SCC1001 and SCC1002 were inoculated into 5 ml of LB and incubated at 37°C for one hour. 20 µl of each culture were inoculated into 1 ml of pre-warmed DMEM in a 6 well tissue culture plate and incubated for up to three hours at 37°C in 5% CO_2. Bfp production was observed using light microscopy at x400 magnification. EPEC cells that produced Bfp were visualised in large clumps where bacteria had adhered to each other (data not shown). By this criterion, both SCC1001 and SCC1002 produced Bfp although production was visibly reduced by approximately 50% in the *typA* mutant strain SCC1002 compared to SCC1001.

5.2.3 Determination of bundle-forming pilus production in strain SCC1002 using the alkaline phosphatase assay.

It was shown above that there appears to be a difference in Bfp expression between wild-type SCC1001 and *typA* mutant strain SCC1002. This could have consequences for the ability of EPEC to adhere to cultured epithelial cells and cause actin accumulation as shown by the FAS test. However, this difference was observed microscopically with no quantitative measurement. Therefore, a quantitative assay is required to determine the expression of Bfp so that it may be compared between strains. A Bfp reporter is available (Donnenberg *et al* 1990) which enables the quantitative expression of *bfpA* to be determined. This construct, namely p6-8-1 from strain EPEC 6-8-1, possesses a Tn*phoA* insertion approximately 330 bp into the *bfp*A gene in the EAF plasmid and was constructed during experiments aimed at identifying genes necessary for HEp-2 cell invasion in EPEC (Donnenberg *et al* 1990; Donnenberg *et al* 1992). Strain EPEC 6-8-1 adheres

to HEp-2 cells in a diffuse pattern after six hours and is FAS test positive, but is unable to invade HEp-2 cells (Donnenberg *et al* 1990).

The transposon Tn*phoA* was constructed by Manoil and Beckwith (1985) by inserting the bacterial alkaline phosphatase gene *phoA* within the upstream terminal sequence of the transposon Tn5 (**figure 5.3**). The modified *phoA* gene retains the catalytic domain but lacks the DNA sequences encoding the promoter and amino-terminal export signal. Tn*phoA* is therefore an ideal transposon for studying gene expression. Alkaline phosphatase is normally found in the periplasm of Gram-negative bacteria and must be located extracytoplasmically for it to be active (Brockman and Heppel 1968). If Tn*phoA* is introduced into a gene in the correct orientation and in-frame then a gene fusion is made. If the gene fusion encodes a protein which is normally located in the inner membrane, periplasm, or outer membrane, then the fusion protein will be exported and will possess alkaline phosphatase activity. Bacterial alkaline phosphatase activity can be demonstrated by growing bacteria on LBA containing the alkaline phosphatase indicator substrate 5-chloro-4-bromo-3-indoyl phosphate (XP); PhoA-positive colonies possess a blue coloration.

The p6-8-1 construct can be introduced into EPEC strains that do not not possess an EAF plasmid and selected for by plating onto LBA containing kanamycin. The p6-8-1 construct can also be introduced into strains already possessing an EAF plasmid by selective competition. The EAF plasmid exists only as a single copy in EPEC and therefore, when p6-8-1 is introduced, either it or the wild-type EAF plasmid remain in the cell. Again, the presence of p6-8-1 can be selected for by plating onto LBA containing kanamycin. This section describes the use of p6-8-1 as a reporter in a quantitative assay to determine the expression of Bfp in strain SCC1002 compared to its parent SCC1001.

5.2.3.1 Introduction of the Bfp reporter p6-8-1 into SCC1002 by electroporation. To perform the Bfp::PhoA assay it was necessary to introduce p6-8-1 into strain SCC1002 and its parent SCC1001. This was achieved by preparing

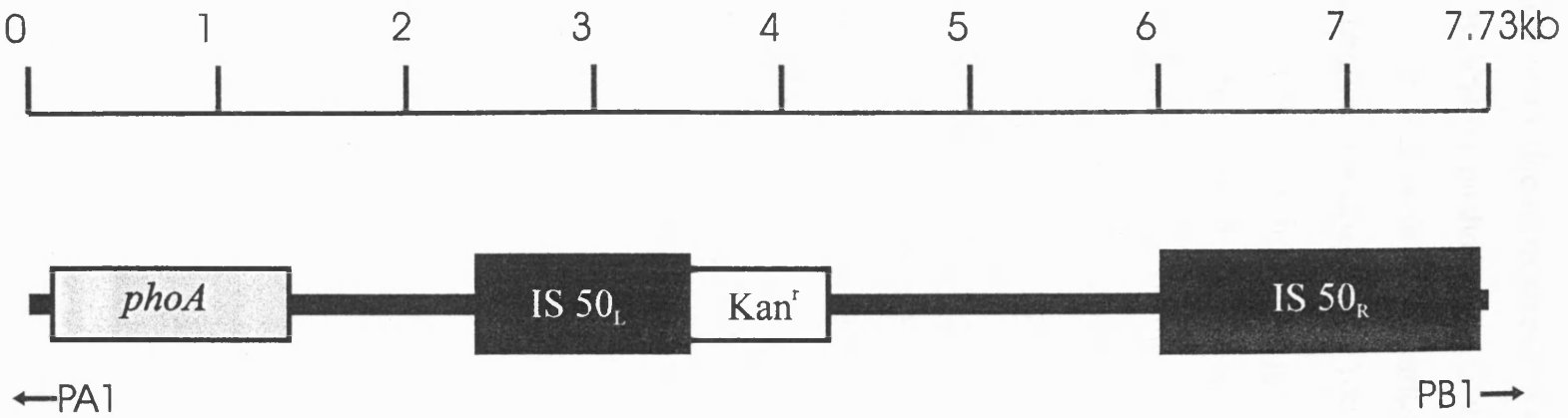
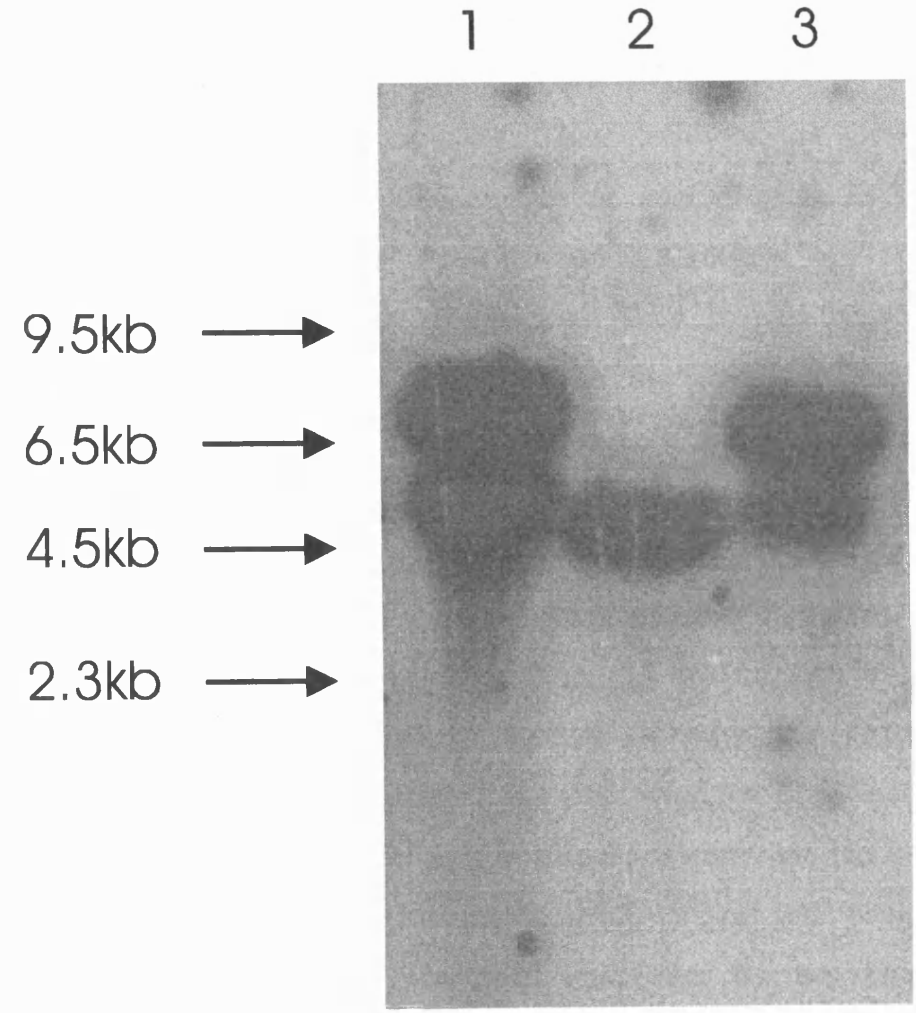


Figure 5.3 Diagrammatic representation of the transposon Tn*phoA*. Transposon Tn*phoA* is 7.73 kb in length as indicated by the values shown in the diagram (Manoil and Beckwith 1985). Tn*phoA* was used in construct EPEC 6-8-1 for introducing a mutation into the *bfpA* gene (Donnenberg *et al* 1990). PA1 and PB1 are primer annealing sites directed outwards of Tn*phoA* for sequencing DNA on either side of the transposon.

chromosomal DNA from *E. coli* DH5α p-6-8-1 using the CTAB chromosomal method and introducing it into each strain by electroporation. Chromosomal DNA was chosen since greater than 80% of the p6-8-1 would be expected to separate with the chromosomal DNA due to its size. It was too large to purify using the Qiagen® Tip100 midi-preparation method. Due to the expected inefficiency of the p6-8-1 transformation, the complete transformation mixture was centrifuged after appropriate incubation, resuspended in 100 µl of distilled water, plated onto LBA containing 25 µg/ml kanamycin and 20 µg/ml XP and incubated overnight at 37°C. As few colonies grew compared to a standard plasmid transformation, all blue colonies were replated to single colonies onto LBA containing 25 µg/ml kanamycin and 20 µg/ml XP to confirm that they were blue. One blue colony from each strain was taken and grown overnight at 37°C in 50 ml LB with the addition of 25 µg/ml kanamycin. After growth, chromosomal DNA was prepared using the CTAB method and 10 µl of the purified DNA was digested with *Bam*HI at 37°C for four hours. The restriction enzyme *Bam*HI was chosen because digestion with this enzyme would release the fragment containing *bfp*A and *pho*A. The digested DNA was then separated using 1% gel electrophoresis and the DNA transferred to nylon membrane using the Southern blot method. A 5.4 kb *bfp* DNA fragment was removed from pRDH7 (supplied by Haigh RD, Leicester University) by restriction digestion with *Bam*HI and purified using the polyallomer wool/QIAGEN QIAquick™ PCR purification kit method. The plasmid pRDH7 is a clone possessing a 5.4 kb *bfp* fragment with a *phoA* fusion in *bfpA*. The fragment was used as a probe according to the Amersham Life Science Fluorescein Gene Images™ labelling and detection system. By calculating the size of band present on the photographic film, the presence or absence of p6-8-1 could be determined. If p6-8-1 was present then 7.8 kb and 5.2 kb bands would be observed; if p6-8-1 was absent then a 5.3 kb band would be observed. This provided confirmation that p6-8-1 had been successfully introduced into SCC1001 and SCC1002 (**figure 5.4**); the strains were designated SCC1001 p6-8-1 and SCC1002 p6-8-1.

Figure 5.4 Southern blot hybridisation analysis of chromosomal DNA from strains EPEC 6-8-1, SCC1001 and SCC1002 p6-8-1 probed with *bfp*. Chromosomal DNA from strain EPEC E2348/69 p6-8-1 (lane 1), SCC1001 (lane 2) and SCC1002 p6-8-1 (lane 3) was prepared and digested with *Bam*HI. The resultant DNA fragments were separated using 1% gel electrophoresis, transferred to nylon membrane by Southern blot and hybridised with a 5.4 kb *bfp* DNA fragment from pRDH7 (supplied by Haigh RD, Leicester University). Bands of sizes 7.8 kb and 5.2 kb were present in lanes 1 and 3 indicating the presence of the *bfpA*-Tn*phoA* gene fusion in strains E2348/69 p6-8-1 and SCC1002 p6-8-1. A single band of 5.3kb was present in lane 2 indicating the absence of the *bfpA*-Tn*phoA* gene fusion in strain SCC1001.

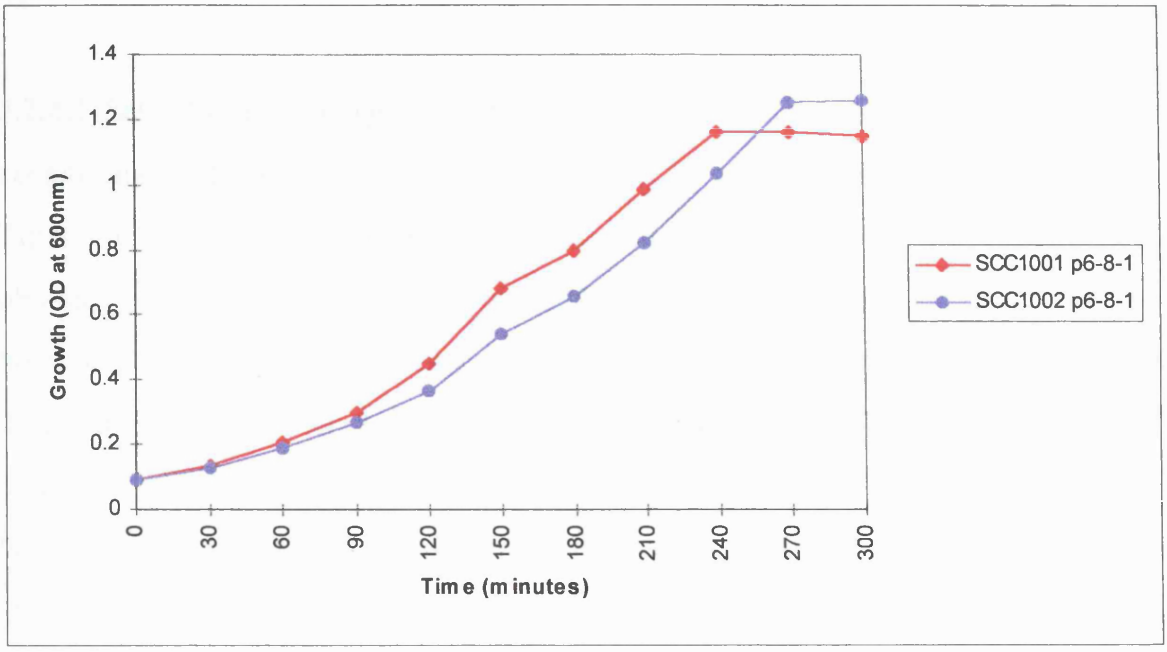


5.2.3.2 Bundle-forming pilus PhoA assay of SCC1001 p6-8-1 and SCC1002 p6-8-1. This experiment was performed in duplicate. Strains SCC1001 p6-8-1 and SCC1002 p6-8-1 were grown overnight in 5 ml of LB containing 50 μg/ml streptomycin and 25 μg/ml kanamycin. Each overnight culture was diluted 1/50 in colourless DMEM to a final volume of 100 ml and grown at 37°C with shaking at 270 rpm in a New Brunswick Scientific gyrotory water bath shaker (model G76D). The OD at 600 nm and alkaline phosphatase activity of each culture were measured at time zero and every 30 minutes until three similar readings were taken indicating that the cultures had reached stationary phase. Although the experiment was performed in duplicate there was little difference between the readings from each. Although the wild-type SCC1001 p6-8-1 grew very slightly faster and reached stationary phase before SCC1002 p6-8-1 (**figure 5.5a**), the alkaline phosphatase activities of the two strains were very similar (**figure 5.5b**).

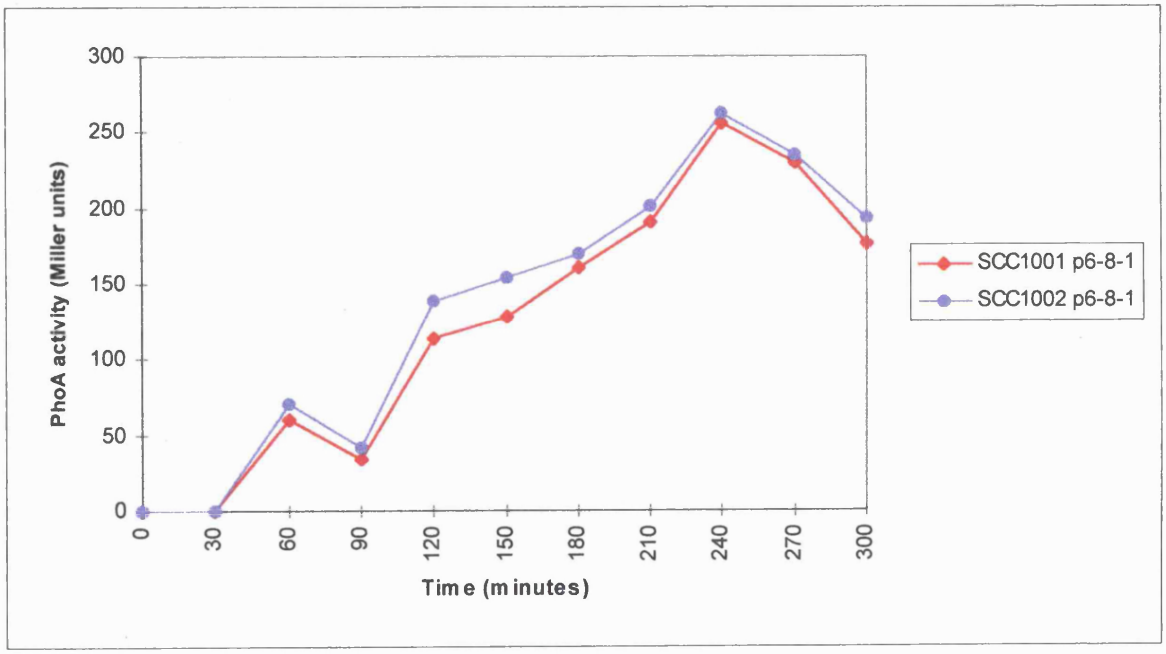
5.2.4 Introduction of *typA* into the expression vector pKK223-3 and its subsequent introduction into strains SCC1002 and SCC1005.

If a mutation affects the phenotype of the strain then it must be shown that complementation of the mutation restores the wild-type phenotype. Complementation may be achieved by the introduction of the wild-type gene on a single or low-copy plasmid. It was demonstrated above that the *typA* mutation in strain SCC1002 appeared to affect Bfp production when observed microscopically. The *typA* mutation therefore required complementation before further assays could be performed and conclusions drawn from them. As described in chapter 3, a plasmid containing wild-type *typA* was constructed. Although this gene originates from *E. coli* K-12 it has been demonstrated that there were few difference between the *E. coli* K-12 and EPEC E2348/69 *typA* sequences at either the DNA or amino acid levels. It was therefore possible that the *E. coli* K-12 *typA* gene could be used for complementation although not confirmed until after its use. This section therefore describes the transfer of *typA* from pSCC-T to the expression vector pKK223-3. This is a single-copy vector which has been used extensively in gene expression studies, including toxin expression experiments in enteroinvasive *E. coli*

Figure 5.5 Graph showing Bundle-forming pilus PhoA assay of SCC1001 and SCC1002 p6-8-1. This experiment was performed in duplicate. Strains SCC1001 p6-8-1 and SCC1002 p6-8-1 were grown overnight in 5 ml of LB containing 50 μg/ml streptomycin and 25 μg/ml kanamycin. Each overnight culture was diluted 1/50 in colourless DMEM to a final volume of 100 ml and grown at 37°C with shaking at 270 rpm. The OD at 600 nm and alkaline phosphatase activity of each culture were measured at time zero and every 30 minutes until three similar readings were taken indicating that the cultures had reached stationary phase. The assay indicated that the alkaline phosphatase activities of SCC1001 p6-8-1 and SCC1002 p6-8-1 were very similar



Graph A - growth curves



Graph B - PhoA activity

(Nataro *et al* 1995). Use of this single-copy vector for complementation ensures that *typA* is not over-expressed with potential problems such as lethality.

5.2.4.1 Introduction of *typA* into the expression vector pKK223-3 creating the recombinant plasmid pSCC-T2. Plasmid pSCC-T was prepared using the Qiagen® Tip100 midi-preparation method so that concentrated, good quality DNA was obtained. Approximately 1 µg of plasmid DNA was digested with *Eco*RI for one hour at 37°C in the appropriate reaction buffer followed by purification of the digested DNA by ethanol precipitation. The DNA was redissolved in 20 µl of distilled water and digested with *Hind*III for one hour at 37°C in the appropriate reaction buffer. The DNA digestion products were separated using 1% gel electrophoresis and the band corresponding to approximately 1.8 kb was excised from the gel. The DNA was purified using the polyallomer wool/QIAGEN QIAquick™ PCR purification kit method. The DNA fragment was then ligated with *Eco*RI-*Hind*III digested pKK223-3. The ligation product was purified using the Qiagen® QIAquick PCR purification method and transformed into DH5α by electroporation. Transformants were plated onto LBA containing 100 µg/ml ampicillin and incubated overnight at 37°C. After incubation, transformants were patch-plated onto LBA containing 100 µg/ml ampicillin and onto an Amersham Hybond™ N+ gridded nylon membrane placed onto LBA containing 100 µg/ml ampicillin. The former plate was stored at 4°C for future reference. The transformants were then screened by colony blot analysis using the Amersham Life Science Fluorescein Gene Images™ labelling and detection system as described in the Materials and Methods. The hybridisation probe was prepared from pSCC-T as described in **section 3.2.1.2**. Two colony patches were found to hybridise with pSCC-T and the corresponding duplicates were isolated from the stored LBA plate. Confirmation of the clone was achieved by preparing plasmid using the Qiagen® QIAprep Spin mini-preparation method, digesting 0.5 µg with *Bst*XI at 55°C for two hours and visualising the digestion product using 1% gel electrophoresis. The clone was visualised at the correct size of approximately 6.4 kb. This confirmation was performed to check that pSCC-T had not been selected in error; in which case it

would have been approximately 4.5 kb in size. The recombinant plasmid pKK223-3 containing *typA* was designated pSCC-T2.

5.2.4.2 Introduction of pSCC-T2 into strains SCC1002 and SCC1005. Plasmid pSCC-T2 was prepared using the Qiagen® Tip100 midi-preparation method so as to obtain concentrated, good quality DNA. Electrocompetent cells of strains SCC1002 and SCC1005 were electroporated with pSCC-T2 and plated onto LBA containing 50 μg/ml streptomycin and 100 μg/ml ampicillin, or 50 μg/ml nalidixic acid and 100 μg/ml ampicillin, respectively. All plates were incubated overnight at 37°C. Transfer of the plasmid to each strain was confirmed by Qiagen® QIAprep Spin mini-preparation of plasmid DNA followed by restriction enzyme digest with *Bst*XI. The strains were designated SCC1002 pSCC-T2 and SCC1005 pSCC-T2.

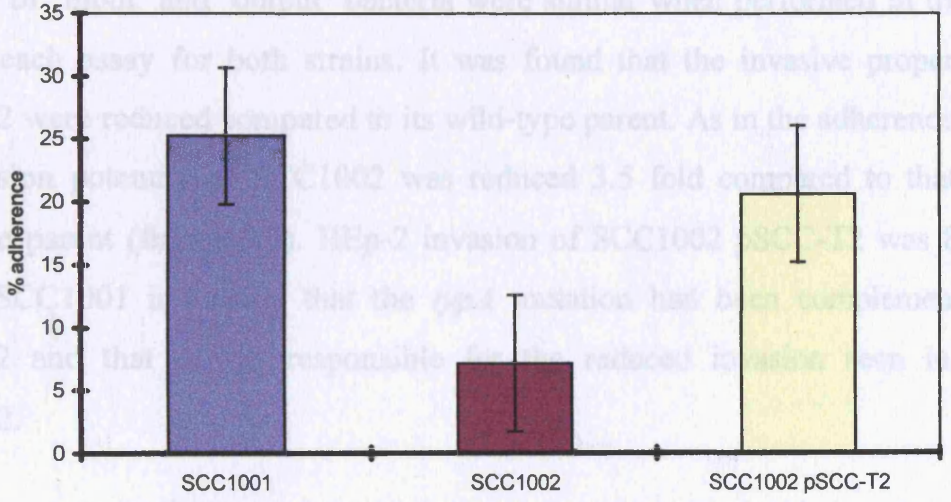
5.2.5 Determination of the adherence and invasion properties of SCC1001, SCC1002 and SCC1002 pSCC-T2 on HEp-2 cells.

The ability of EPEC, and indeed most other gastrointestinal pathogens, to adhere to the intestinal mucosa is a pre-requisite for pathogenicity (Nataro and Kaper 1998). Although EPEC is not generally regarded as an invasive organism (Levine *et al* 1978; Edelman and Levine 1983; Levine 1987; Robins-Browne 1987), intracellular bacteria have been observed in EPEC-infected tissue culture cells (Knutton *et al* 1987) and invasion assays have proved useful in the analysis of EPEC mutants because *eae* is required for invasion (Donnenberg and Kaper 1991). It has also been demonstrated that significant numbers of bacteria are internalised when EPEC are used to infect HEp-2 cells (Donnenberg *et al* 1989). Invasion assays are therefore a useful addition to adherence assays for determining virulence potential *in vivo*. This section describes adherence and invasion assays to determine the ability of E2348/69 Strep^r and SCC1002 to adhere to and invade cultured epithelial cells *in vitro*. It is expected from previous studies (Donnenberg *et al* 1990; Francis *et al* 1991) that approximately 50% of the input wild-type bacteria will adhere and that 1% will invade.

5.2.5.1 Adherence properties of SCC1001, SCC1002 and SCC1002 pSCC-T2 on HEp-2 cells. This adherence assay was performed three times with each strain being inoculated in triplicate in each assay. Strains SCC1001, SCC1002 and SCC1002 pSCC-T2 were grown overnight in 5 ml of LB containing 50 μg/ml streptomycin. Monolayers of HEp-2 cells were prepared in 24-well plates overnight as described. After incubation, the DMEM/10% foetal calf serum was replaced with fresh DMEM containing 10% mannose and incubated for one hour at 37°C in 5% CO_2 to equilibrate. 5 μl of the overnight cultures of each strain were inoculated into each well so that approximately $2x10^7$ bacteria were used. At this point serial dilution of 5 μl of each overnight culture was performed to determine the number of 'input' bacteria. In addition, one well was inoculated as a control and not handled throughout the experiment. This control was used to determine the number of 'output' bacteria. After incubation at 37°C in 5% CO_2 for 3 hours and subsequent washing, the adherent bacteria were released by solubilisation using Triton X-100 in PBS, serially diluted down to $1/10^6$ and plated onto LBA containing 50 μg/ml streptomycin. The contents of the control 'output' well were also serially diluted. The bacteria were enumerated and the number of adherent bacteria calculated by multiplying the number counted by the dilution factor. It was noted that the numbers of 'input' and 'output' bacteria were similar when performed in triplicate and for each assay for both strains. The adherence of SCC1002 was reduced 3.5-fold compared with that of its wild-type parent (**figure 5.6**). HEp-2 invasion of SCC1002 pSCC-T2 was 84% of that of SCC1001 indicating that the *typA* mutation had been complemented by pSCC-T2.

5.2.5.2 Invasion assays of SCC1001, SCC1002 and SCC1002 pSCC-T2 on HEp-2 cells. As for the adherence assay, this assay was performed three times with each strain being inoculated in triplicate in each assay. Strains SCC1001, SCC1002 and SCC1002 pSCC-T2 were also grown overnight in 5 ml of LB containing 50 μg/ml streptomycin. Monolayers of HEp-2 cells were prepared in 24-well plates overnight as described. After incubation, the DMEM/10% foetal calf serum was replaced with fresh DMEM containing 10% mannose and incubated for one hour at 37°C in 5%

Figure 5.6 Adherence of strains SCC1001, SCC1002 and SCC1002 pSCC-T2 to HEp-2 cells. The assay was performed in triplicate and in three separate experiments. Approximately $2x10^7$ bacteria of strains SCC1001, SCC1002 or SCC1002 pSCC-T2 were added to monolayers of HEp-2 cells in 24-well plates. The cells were incubated at 37°C in 5% CO_2 for 3 hours and then washed with sterile PBS. The adherent bacteria were released by solubilisation, serially diluted down to $1:10^6$, plated onto LBA containing 50 µg/ml streptomycin and then enumerated. The bars indicate the number of bacteria recovered expressed as a percentage of the inoculum.

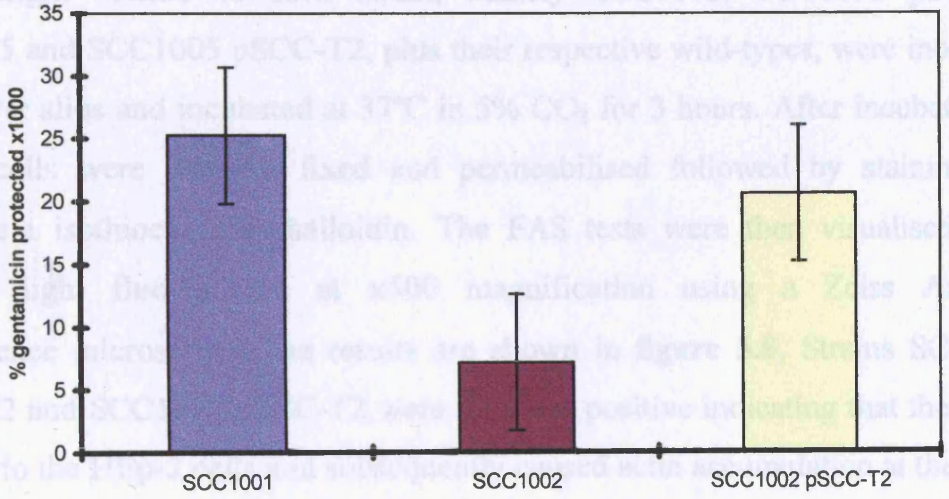


CO_2 to equilibrate. 5 µl of the overnight cultures of each strain were inoculated into each well so that approximately $2x10^7$ bacteria were used. At this point serial dilution of 5 µl of each overnight culture was performed to determine the number of 'input' bacteria. In addition, one well was inoculated as a control and not handled throughout the experiment. This control was used to determine the number of 'output' bacteria. After incubation at 37°C in 5% CO_2 for 3 hours and subsequent washing, 100 µg/ml gentamicin was added and incubation continued for a further 2 hours at 37°C in 5% CO_2. The invasive bacteria were released by solubilisation using Triton X-100 in PBS, serially diluted down to $1/10^5$ and plated onto LBA containing 50 µg/ml streptomycin. The contents of the control 'output' well were also serially diluted. The bacteria were enumerated and the number of invasive bacteria calculated by multiplying the number counted by the dilution factor. The numbers of 'input' and 'output' bacteria were similar when performed in triplicate and for each assay for both strains. It was found that the invasive properties of SCC1002 were reduced compared to its wild-type parent. As in the adherence assay, the invasion potential of SCC1002 was reduced 3.5 fold compared to that of its wild-type parent (**figure 5.7**). HEp-2 invasion of SCC1002 pSCC-T2 was 88% of that of SCC1001 indicating that the *typA* mutation had been complemented by pSCC-T2 and that it was responsible for the reduced invasion seen in strain SCC1002.

5.2.6 Use of the FAS test in demonstrating actin accumulation in SCC1002 and SCC1005.

The AE lesion is central to the pathogenesis of EPEC. One aspect of AE lesion formation is the accretion of actin at the site of bacterial adherence. Genes responsible for the processes leading up to and including formation of the AE lesion include the *per* genes, *eae* and a number of genes contained within the LEE (Nataro and Kaper 1998). If *typA* interacts with the expression of any of these genes then actin accumulation may be affected. As I have shown above, Bfp expression appears to be reduced when observed microscopically in the *typA* mutant strain SCC1002 although according to PhoA assay data expression is similar to the parent.

Figure 5.7 Invasion of Hep-2 cells by strains SCC1001, SCC1002 and SCC1002 pSCC-T2. The assay was performed in triplicate and in three separate experiments. Approximately $2x10^7$ bacteria of strains SCC1001, SCC1002 or SCC1002 pSCC-T2 were added to monolayers of HEp-2 cells in 24-well plates. Cells were incubated at 37°C in 5% CO_2 for 3 hours and then washed with sterile PBS. DMEM containing 100 µg/ml gentamicin was added to the cells and incubation was continued for a further 2 hours at 37°C in 5% CO_2. The invasive bacteria were released by solubilisation, serially diluted down to $1:10^5$, plated onto LBA containing 50 µg/ml streptomycin and then enumerated. The bars indicate the number of bacteria recovered expressed as a percentage of the inoculum x1000.

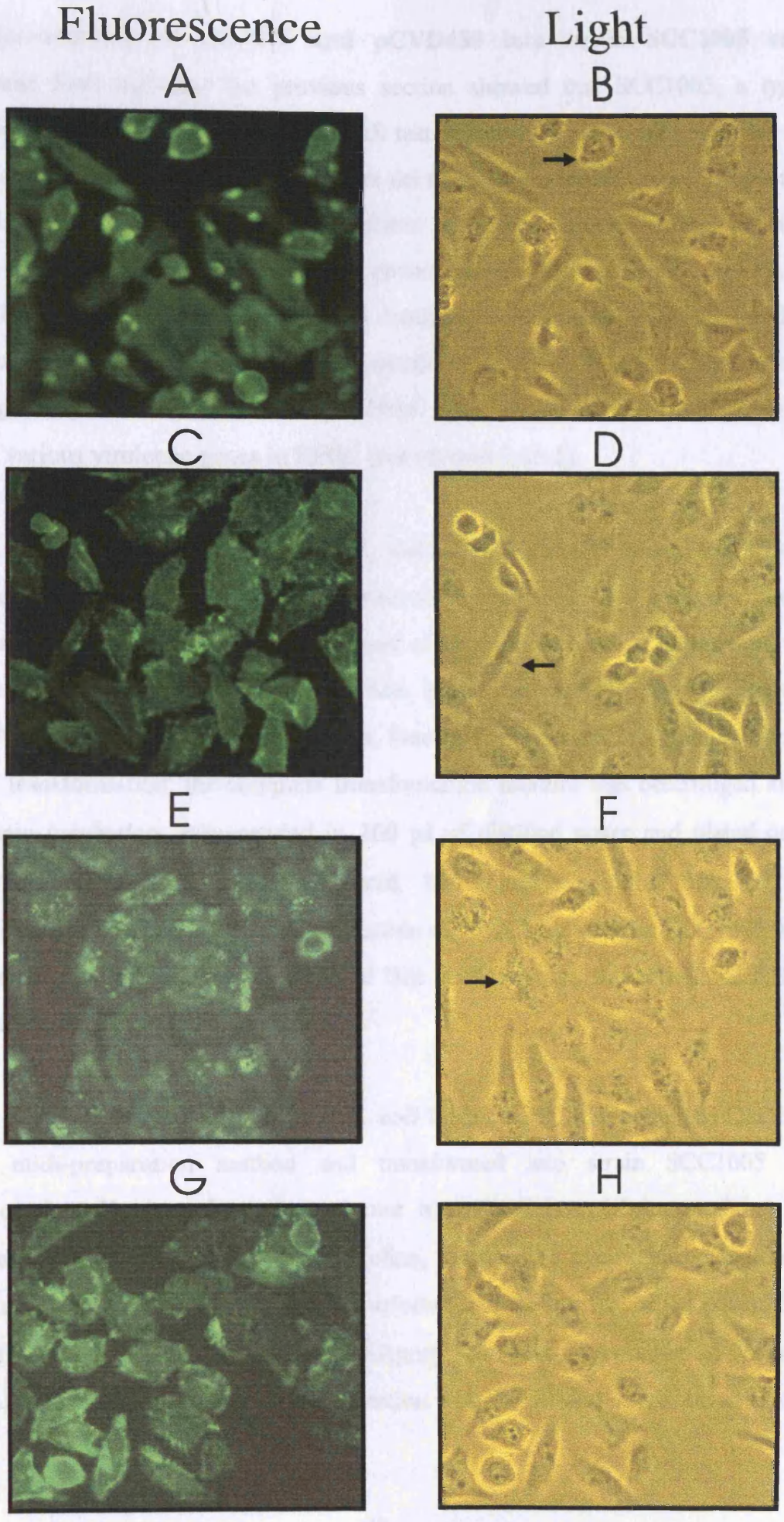


Nevertheless, invasion is reduced by 72% in SCC1002. It is therefore reasonable to suspect that the *typA* mutation may affect AE lesion formation and actin accumulation in some way. This section describes the use of the FAS test in demonstrating the ability of strains SCC1002 and SCC1005 and their respective parents to induce actin accumulation at the site of bacterial attachment. Mutational complementation using strains SCC1002 pSCC-T2 and SCC1005 pSCC-T2 was also performed.

5.2.6.1 FAS testing of SCC1002 and SCC1005. The FAS test provides a standard phenotypic assay for the pathogenicity traits of EPEC and other AE lesion-forming pathogens by determining the presence of intimate adherence, lesion formation and subsequent actin rearrangement (Knutton *et al* 1989). Monolayers of HEp-2 cells were prepared on sterile cover slips as described in Materials and Methods. 5 µl of an overnight culture of each strain, namely SCC1002, SCC1002 pSCC-T2, SCC1005 and SCC1005 pSCC-T2, plus their respective wild-types, were inoculated onto cover slips and incubated at 37°C in 5% CO_2 for 3 hours. After incubation the HEp-2 cells were washed, fixed and permeabilised followed by staining with fluorescein isothiocyanate-phalloidin. The FAS tests were then visualised at by incident light fluorescence at x500 magnification using a Zeiss Axiophot fluorescence microscope. The results are shown in **figure 5.8.** Strains SCC1001, SCC1002 and SCC1002 pSCC-T2 were FAS test positive indicating that the strains adhered to the HEp-2 cells and subsequently caused actin accumulation at the site of adherence. Wild-type derivative JPN15 was also FAS test positive albeit with a decreased frequency of adherence due to lack of the EAF plasmid and therefore its inability to produce Bfp. The *typA* mutant SCC1005, however, was FAS test negative. Although the strain was able to adhere to HEp-2 cells it was unable to cause actin accumulation at the site of adherence (**figure 5.8**). The FAS test-positive phenotype was restored in SCC1005 pSCC-T2 indicating that the *typA* mutation had been complemented by pSCC-T2.

Figure 5.8 FAS testing of strains SCC1002 and SCC1005. Monolayers of HEp-2 cells were prepared on sterile cover slips. 5 µl of an overnight culture of each strain, namely SCC1002, SCC1002 pSCC-T2, SCC1005 and SCC1005 pSCC-T2 plus their respective wild-types were inoculated onto a cover slip and incubated at 37°C in 5% CO_2 for 3 hours. After incubation the HEp-2 cells were washed, fixed and permeabilised followed by staining with fluorescein isothiocyanate-phalloidin. The FAS tests were then visualised by incident light fluorescence at x500 magnification using a Zeiss Axiophot fluorescence microscope. The full results are described in **section 5.2.5.1**; this figure shows representative positive and negative FAS test results in SCC1002 and SCC1005 and their wild-types.

Panel	Strain	Microscopy	Result
Panel A	SCC1001	Fluorescence	FAS test positive
Panel B	SCC1001	Light	Adherence positive
Panel C	SCC1002	Fluorescence	FAS test positive
Panel D	SCC1002	Light	Adherence test positive
Panel E	JPN15	Fluorescence	FAS test positive
Panel F	JPN15	Light	Adherence reduced
Panel G	SCC1005	Fluorescence	FAS test negative
Panel H	SCC1005	Light	Adherence reduced


Fluorescence
Light
A
B
C
D
E
F
G
H

5.2.6.2 Introduction of pMAR7 and pCVD450 into strain SCC1005 and subsequent FAS testing. The previous section showed that SCC1005, a *typA* mutant lacking the EAF plasmid, was FAS test negative. To determine whether an interaction exists between *typA* and genes on the EAF plasmid further assays are required. Two plasmids can be introduced into strain SCC1005 which may provide clues to the reason for the FAS negative phenotype. The first plasmid, pMAR7, is an ampicillin-tagged EAF plasmid which contains all the genes necessary for Bfp biogenesis as well as the *per* genes. The second plasmid, pCVD450, contains only the *per* genes (Gómez-Duarte and Kaper 1995). The *per* genes have been shown to regulate various virulence genes in EPEC (see **section 1.10.1**).

Plasmid pMAR7 was prepared from *E. coli* DH5α pMAR7 using the CTAB chromosomal method and introduced into strain SCC1005 by electroporation. Chromosomal DNA was chosen since most of the pMAR7 DNA would separate with the chromosomal DNA due to its size. It was too large to purify using the Qiagen® Tip100 midi-preparation method. Due to the expected inefficiency of the pMAR7 transformation, the complete transformation mixture was centrifuged after appropriate incubation, resuspended in 100 µl of distilled water and plated onto LBA containing 50 µg/ml nalidixic acid, 100 µg/ml ampicillin followed by incubation overnight at 37°C. The introduction of pMAR7 into strain SCC1005 was confirmed by microscopic observation of Bfp production as in **section 5.2.2**; the strain was designated SCC1005 pMAR7.

Plasmid pCVD450 was prepared from *E. coli* DH5α pCVD450 using the Qiagen® Tip100 midi-preparation method and transformed into strain SCC1005 by electroporation. The transformation mixture was plated onto LBA containing 50 µg/ml nalidixic acid and 20 µg/ml tetracycline, followed by overnight incubation at 37°C. Representative transformants were selected and the transfer of the plasmid to SCC1005 was confirmed by Qiagen® QIAprep Spin mini-preparation of plasmid DNA followed by restriction enzyme digestion with *Eco*RI and visualisation of the

3.5 kb *per* DNA fragment using 1% gel electrophoresis. The strain was designated SCC1005 pCVD450.

The FAS tests were performed as in **section 5.2.6.1**. It was found that strains SCC1005 pMAR7 and SCC1005 pCVD450 were able to adhere to HEp-2 cells and caused actin accumulation at the site of adherence in the same manner as wild-type EPEC E2348/69 (data not shown, but see **figure 5.8** panels A an B). This indicated that the FAS test phenotype was restored in these strains thereby suggesting that the *per* genes play a role, along with *typA*, in the expression of the FAS test phenotype.

5.3 Discussion

This chapter has described the effect of the *typA* mutation on the virulence of EPEC using a number of phenotypic assays. This analysis included assays for growth, bundle-forming pili production, cell adherence, cell invasion and actin accumulation. It was found that the *typA* mutation did not affect the virulence of EPEC E2348/69 using any of these phenotypic assays.

Although it was shown in chapter 4 that *typA* mutants grow poorly at 28°C, this phenotype was not chosen for further study. This is because all virulence assays are performed at 37°C to imitate the temperature found within the human gastrointestinal tract although further analysis of the effects of growth at 28°C would be interesting. However, other growth assays were performed. The growth rates of the mutant strains SCC1002, SCC1003, SCC1005 and SCC1006 in LB and DMEM at 37°C were very similar to those of their parents SCC1001, EPEC E2348/69 Nal^r, JPN15 and RDEC-1 respectively. This indicates that the *typA* mutation did not affect the growth characteristics of these strains/derivatives. This agrees with results from previous studies on *typA* in *E. coli* K-12 and EPEC derivative MAR001 (Freestone P, unpublished results). However, strain SCC1004 grew faster than its parent MAR001 and was therefore not used for further study. The *typA* mutation in this strain was not confirmed by any method although it displayed the 28°C growth phenotype exhibited by the other *typA* mutant strains.

However, further analysis could actually be performed on the strains to check that the cell size of each bacterial strain was the same for the wild-type and the mutant. If the cell size was different for either then this would affect the optical density and lead to false conclusions regarding growth. In addition, viable counts could be used instead to determine the number of bacteria present rather than optical density.

Microscopic observations indicated that bundle-forming pili production in strain SCC1002 incubated in DMEM at 37°C, 5% CO_2 appeared reduced compared to its wild-type parent. The growth conditions used for this experiment are thought to be similar in many respects to those in the host during infection. However, in the alkaline phosphatase assay, although the wild-type SCC1001 p6-8-1 grew very slightly faster and reached stationary phase before SCC1002 p6-8-1, the alkaline phosphatase activity of SCC1002 p6-8-1 was very similar to that of SCC1001 p6-8-1. The alkaline phosphatase assay is more reliable as an assay than visual observation of Bfp production. The results from this assay indicate that Bfp expression was similar in SCC1002 p6-8-1. Further analysis of Bfp expression could be performed using the Western blot method. Antibodies against Bfp are available and therefore the technique could be performed rapidly. Also it has recently been shown that the size of Bfp alters during EPEC growth, changing from thick to thin (Knutton *et al* 1999). This could therefore have an effect of increasing or decreasing the values recorded for Bfp expression according to the size of the Bfp at the time and the expression of *bfpA* related to complete Bfp expression.

The main phenotypic assays relating to the virulence of EPEC rely on interactions with cultured mammalian cells. Assays for adherence, invasion and actin accumulation were performed which are all major traits of EPEC virulence. Adherence to and invasion of mammalian epithethial cells are important for many pathogenic bacteria, particularly gastrointestinal pathogens. *Shigella* is one example of a pathogenic bacterium that adheres and invades intestinal epithelial cells. Although EPEC does not invade epithelial cells to such a degree *in vivo*, adherence and invasion assays are useful indicators of virulence potential when experiments are being performed *in vitro*. The adherence to, and invasion of, HEp-2 cells by

SCC1002 was found to be reduced by 3.5 fold compared with that of its wild-type parent. The reduction in adherence to HEp-2 cells by SCC1002 is not explained by the results of the PhoA assay as this indicated that Bfp expression of SCC1002 was similar to that of its parent SCC1001. However, the PhoA assay measured *bfpA* expression which indicates that BfpA is being produced but does not necessarily indicate pilin assembly. Therefore, the *typA* mutation may affect other genes within or external to the *bfp* gene cluster which are associated with pilin assembly.

The results of the FAS tests provided a surprising result in that SCC1002 was FAS test positive like its wild-type parent but SCC1005 was FAS test negative. As SCC1005 did not possess the EAF virulence plasmid this provided the first link between *typA* and virulence. Although previous work with a MAR001 *typA* mutant had shown a FAS test negative phenotype (Farris *et al* 1998), this was not expected to be due to an interaction between *typA* and genes on the EAF plasmid. It was suspected that the phenotype seen by Farris and colleagues was due either to the *typA* mutation directly or to another unknown mutation in EPEC derivative MAR001.

To prove that *typA* and the EAF plasmid were responsible for loss of the FAS test phenotype, further FAS tests were performed. Firstly, this involved introducing wild-type *typA* into SCC1005 on a single-copy expression vector (pSCC-T2). Strain SCC1005 pSCC-T2 was FAS test positive indicating that wild-type *typA* complemented the chromosomal *typA* mutation thereby restoring the wild-type FAS test phenotype. Secondly, plasmids pMAR7 and pCVD450 were introduced into SCC1005 separately. Plasmid pMAR7 is the EAF plasmid with ampicillin resistance and therefore possesses all the genes so far identified on the plasmid including the *bfp* and *per* genes. The latter have been shown to regulate various EPEC virulence genes and therefore would be prime candidates for an interaction with *typA*. Plasmid pCVD450 is pACYC184 with the *per* genes cloned into the multi-cloning site (Gómez-Duarte and Kaper 1995). Strain SCC1005 pMAR7 was FAS test positive indicating that genes on the EAF plasmid had suppressed the chromosomal *typA* mutation thereby restoring the wild-type FAS test phenotype. Furthermore, strain

SCC1005 pCVD450 was also FAS test positive thereby showing that the *per* genes on the EAF plasmid were responsible for complementation of the chromosomal *typA* mutation in strain SCC1005.

Complementation of the *typA* mutations was achieved using *typA* derived from *E. coli* K-12. It was suggested that this would be okay due to the *E. coli* K-12 and EPEC *typA* genes being very similar. Although this appeared to be sufficient and provided between 80 and 90% complementation, the use of EPEC *typA* for complementation may provide nearer 100% complementation.

The results clearly demonstrate that both *typA* and the *per* genes must be absent before the FAS test phenotype is lost. If only *typA* is absent then all the virulence traits of EPEC which were tested are essentially unaffected. Similarly, if the *per* genes are removed from EPEC then they are still able to adhere to HEp-2 cells as shown by the EPEC derivative JPN15. However, if the *per* genes are introduced into JPN15 via pCVD450 then adherence increases 100-fold. This increase in adherence is specific to an increased activity of *eaeA* (Gómez-Duarte and Kaper 1995). Therefore, *typA* and the *per* genes do not interact with each other as such but must each be responsible for interacting with other genes in EPEC that are responsible for the FAS test phenotype. Such genes that have been identified to date are located on the LEE (Elliot *et al* 1998b). There are four *per* genes, *perA-D*, so named because they are located on a plasmid and encode a regulatory function (Gómez-Duarte and Kaper 1995).

Although the MAR001 *typA* mutant strain SCC1004 was not used for further phenotypic analysis, the results gained from the JPN15 *typA* mutant SCC1005 suggest that there is no difference in the phenotypes of a MAR001 *typA* mutant and JPN15 *typA* mutant, as compared with the results of Farris *et al* (1998).

Chapter 6

General discussion

6.1 Recent advances in EPEC virulence

Molecular biology has made a major impact on EPEC research in the past decade as evidenced by the number of papers published on the subject and the increase in our knowledge of the virulence mechanisms of this gastrointestinal pathogen. Although virulence genes have been identified on both the chromosome and EAF plasmid of EPEC at various loci, there are probably other virulence genes which remain to be identified. For example, the model of EPEC pathogenesis remains controversial because of the role of bundle-forming pili. For a number of years Bfp were thought to be solely responsible for the localisation of EPEC bacteria on intestinal epithelial cells followed by intimate adherence mediated by the *eaeA* gene. However, recent work has shown that the ability to autoaggregate, known as the increased autoaggregation phenotype (IAP), exists in some strains of EPEC which has not been previously reported and cannot be discounted as having a role in either adhesion or the regulation of virulence genes (Haigh R, personal communication). In addition, the Tir protein was only recently fully described (Kenny *et al* 1997) although the protein was first identified in 1992 (Rosenshine *et al* 1992). Therefore the identification of further virulence genes in EPEC will probably continue for some time.

6.2 Further work

The reason for the initial interest in *typA* was the fact that its product was tyrosine phosphorylated in EPEC and not in *E. coli* K-12 strains. Unfortunately, the scope of this thesis has not enabled further work on this phosphorylation to be explored. It is therefore important that this is investigated via structural analyses of *typA* in both

pathogenic and non-pathogenic *E. coli* strains. It will be only after such analysis that the role, if any, of tyrosine phosphorylation will be understood and whether it is linked with the virulence potential of EPEC, and indeed other organisms that possess *typA* homologues.

The results of the phenotypic assays performed in this thesis have raised a multitude of questions about the interaction of *typA* with other virulence genes. Although it is unlikely that there is a direct interaction between *typA* and the *per* genes, interactions with other genes may be possible. The actual *per* genes involved with maintenance of the FAS test phenotype in the absence of *typA* must be determined. Therefore each *per* gene, *perA* to *perD*, must be cloned separately and then introduced into strain SCC1005. The effect on the FAS test phenotype can then be determined. In addition, apart from the *per* genes, this thesis has not explored the interaction of *typA* with other EPEC virulence genes. A *typA* mutation in *E. coli* K-12 has been shown to result in the alteration of expression of certain proteins (Freestone *et al* 1995). Such interactions may be important in gene regulation and therefore in EPEC virulence. One such protein whose expression is increased in the *E. coli* K-12 *typA* mutant is H-NS. This protein is a global regulator and may therefore play a role in the regulation of virulence genes. An EPEC E2348/69 strain possessing an arabinose-inducible *hns* gene is available (Goldberg M, personal communication) which could be used to determine such a role. The role of other genes in *typA* regulation could also be determined by constructing fusions of the chloramphenicol acetyl transferase (CATIII) gene with the *typA* promotor. Such constructs could then be used with a transposon mutagenesis approach to identify genes involved in the regulation of *typA*.

Also, the expression of bundle-forming pili in the E2348/69 *typA* mutant (SCC1002) requires further study. Results in chapter 5 showed that the quantitative expression of *bfpA* is similar between SCC1001 and SCC1002 although the adherence of SCC1002 to HEp-2 cells was reduced compared to its parent SCC1001. However, the PhoA assay measured *bfpA* expression which only provides an indication that BfpA is being produced but does not necessarily indicate pilin assembly. Further

work would therefore determine whether Bfp are being assembled completely or whether the *typA* mutation is having an effect on other genes involved in pilin assembly. Such analysis is difficult because it requires the quantitative measurement of bacteria not expressing Bfp; a reporter to show the expression of complete Bfp is not yet available.

The production of secreted proteins was not examined in this thesis and it is therefore essential that such analysis be performed so that further information can be gained on the reason for lack of the FAS test phenotype in *typA* and *per*-negative strains. The Tir protein (Kenny *et al* 1997), otherwise described as EspE (Deibel *et al* 1998), has recently been described. As Tir is a bacterial protein which is inserted into mammalian cells upon EPEC infection (Kenny *et al* 1997) and is required for full expression of the FAS test phenotype, the effect of the *typA* mutation on Tir expression and AE lesion formation would certainly be worth investigating. It is suggested, however, that a *typA* mutation would not have any effect on Tir expression because, if it did, an effect on the FAS test phenotype of SCC1002 would have been seen in the assays performed in this thesis. Also, mutation analysis of specific amino acids of the TypA protein would elucidate the regions within the protein which are necessary for its function. Furthermore, this may provide answers to the differences seen in tyrosine phosphorylation of TypA between EPEC and *E. coli* K-12.

To further assess the phenotype of *typA* and its role in EPEC virulence it is necessary to perform experiments *in vivo*. Although EPEC does not possess a suitable animal model and paediatric intestinal tissue is difficult to obtain, a realistic option is to transfer the *typA* frameshift mutation from EPEC E2348/69 to the rabbit EPEC strain RDEC-1, as described in chapter 4. Various *in vivo* experiments could be performed such as LD_{50} tests (Reed and Muench 1938) and the rabbit ileal loop test (RILT) (Gianella *et al* 1973). LD_{50} tests may be used to determine if the mutant organism retains its ability to cause the full symptoms associated with the disease. The rabbit ileal loop test, which has been modified to the rabbit ileal loop anastomosis test (RILAT) by Dr J. Ketley at the University of Leicester, may be

used to determine whether the mutant organisms retain their ability to cause fluid secretion. Although the RILAT is beyond the scope of this thesis, it is hoped that collaboration with Dr. Ketley in the near future will enable this test to be performed.

The *typA* mutation could also be introduced into other diarrhoeagenic *E. coli*, of both human and animal importance, to determine its effect on virulence phenotypes. In addition, it could be transferred to other attaching/effacing pathogens such as *Hafnia alvei* and *Citrobacter rodentium*. A plethora of phenotypic assays exist for the analysis of the effects of gene mutations on these bacteria. For *H. alvei*, similar assays to those performed on EPEC could be used. In addition, for *C. rodentium*, mice could be used as an animal model to study the interaction of a *typA* mutant with host gastrointestinal cells.

6.3 Is *typA* a virulence gene?

This thesis has investigated, to a limited extent, the role of *typA* in the virulence of EPEC. Initially the *typA* gene from EPEC strain E2348/69 was cloned and sequenced. This was in response to my hypothesis that the *typA* sequence of EPEC differed from that of *E. coli* K-12 thereby accounting for the differences seen in tyrosine phosphorylation of TypA between these two strains. However, DNA sequencing revealed no functional differences between *typA* of EPEC E2348/69 and *E. coli* K-12 and these results agree with those of Farris *et al* (1998) using EPEC derivative MAR001. After sequencing, a site-specific four base-pair deletion in the *typA* gene of EPEC was made resulting in a frameshift mutation. Standard phenotypic studies described for the analysis of EPEC virulence were performed on the mutant strains. My hypothesis that *typA* plays a role in the regulation of virulence genes on a global or non-global level was proven to be correct. Although *typA* was not responsible for any change in EPEC virulence phenotypes when deleted on its own, it was shown that the FAS test phenotype was lost when both *typA* and *per* were absent from EPEC. These results agree with those of Farris *et al* (1998) who used EPEC derivative MAR001, but goes further to show conclusively that loss of the FAS test phenotype is due to the absence of both *typA* and the *per*

locus. Farris *et al* (1998) did not suggest a role for the EAF plasmid in loss of the FAS test phenotype. It was not possible to clone the genes downstream of *typA* in EPEC but such cloning may provide further clues to the function of *typA*. A derivative of SCC1001 has now been constructed which possesses a kanamycin gene cassette towards the N-terminus of *typA* (Haigh R, personal communication). This construct will allow selective cloning of *typA* and the novel downstream regions. As this region is absent from *E. coli* K-12 it is not unreasonable to suspect that it may have a role in *typA* virulence function and *typA* tyrosine phosphorylation. The results presented in this thesis clearly strengthen the case for *typA* being a virulence gene. Two independent studies (Farris *et al* 1998; this thesis) have now shown that a *typA* mutation in EPEC, in the absence of the *per* locus, is responsible for loss of the FAS test phenotype. These results suggest that *typA* has a regulatory function. As other regulatory genes, such as those in the *per* locus, are considered to be virulence genes, *typA* may also now be considered a virulence gene. It is not yet associated with a phenotype specific to its gene product but this may come later as further work is performed.

6.4 The future of *typA*

Freestone and colleagues (Freestone *et al* 1995; 1998a; 1998b) have described differences in tyrosine phosphorylation of *typA* between various *E. coli* strains. TypA is tyrosine phosphorylated in EPEC MAR001 and an *E. coli* L-form, but not in *E. coli* K-12 (Freestone *et al* 1995; 1998a; 1998b). These differences have not been addressed in this thesis and therefore further work is required to determine if tyrosine phosphorylation of *typA* is important in EPEC or whether it is just an interesting observation that initiated the work in this thesis. Differences in tyrosine phosphorylation of proteins between bacterial species is not new. It has also been recently observed by Kenny (1999) that Tir is tyrosine phosphorylated in EPEC but not in EHEC O157:H7 even though both bacteria are able to induce cytoskeletal rearrangement in host epithelial cells upon infection (Nataro and Kaper 1998). However, tyrosine phosphorylation of Tir in EPEC is due to an additional tyrosine (tyrosine 474) which is absent in EHEC O157:H7 (Kenny 1999). No such

additional tyrosine residue is present in EPEC *typA*. Even so, Kenny suggests that tyrosine phosphorylation of Tir may be fortuitous as I have suggested for tyrosine phosphorylation of *typA* in EPEC. Alternatively, EHEC may have evolved an alternative, tyrosine phosphorylation-independent mechanism to perform the same function (Kenny 1999).

As discussed in chapter 4, it may be possible to transfer the *typA* mutation to other gastrointestinal pathogens. Furthermore, *typA* homologues have been shown in other pathogenic and non-pathogenic bacteria and therefore the role of *typA* in such bacteria should be determined. TypA may have a very important role in other pathogenic bacteria such as *Vibrio cholerae* or *Neisseria meningitidis*. Better vaccines are required for both these pathogens because current vaccines do not provide effective or long-term protection. The *typA* gene and TypA protein may therefore be a future target for vaccine development. Also, the threat of antibiotic resistance amongst bacteria is continually increasing so its spread must be slowed and new antimicrobial targets found (Anderson 1999). The *typA* gene or TypA may again be a target for a new class of antimicrobial agents. The potential of *typA* or TypA as a target for therapeutic or vaccine strategies cannot be ignored (Clarke *et al* 1998; Norris *et al*, submitted). Genome sequence data to date suggests that many bacteria possess *typA*, or homologues with high amino acid identity, and therefore a new "universal" antimicrobial agent or vaccine is an exciting proposition. However, such a vaccine may also affect non-virulent strains of similar species. For example, a *typA*/TypA vaccine for EHEC or EPEC may also affect comensal *E. coli* which would be detrimental to the gastrointestinal tract. Further studies would therefore need to determine whether a vaccine or therapeutic agent could "switch off" virulence determinants of a pathogen without affecting other non-virulent species.

Future work on *typA* will inevitably provide additional information on the role of *typA* in EPEC virulence, and indeed virulence in other pathogens. The most likely role for *typA* is as a regulatory gene which may mean it plays a role, if only small, in the regulation of many genes, some of which will be genes specifically associated

with virulence whilst others are not. At the same time, work on *typA* in non-pathogenic bacteria will shed further light on its role.

Appendix I

Genbank submission of the enteropathogenic *Escherichia coli* strain E2348/69 *typA* gene sequence

```
LOCUS       AF058333     1776 bp    DNA             BCT       15-APR-1998
DEFINITION  Escherichia coli tyrosine phosphorylated protein A (typA)
               gene,complete cds.
ACCESSION   AF058333
KEYWORDS    .
SOURCE      Escherichia coli.
  ORGANISM  Escherichia coli
            Eubacteria; Proteobacteria; gamma subdivision;
            Enterobacteriaceae; Escherichia.
REFERENCE   1  (bases 1 to 1776)
  AUTHORS   Freestone,P., Trinei,M., Clarke,S.C., Nystrom,T. and
            Norris,V.
  TITLE     Tyrosine phosphorylation in Escherichia coli
  JOURNAL   J. Mol. Biol. (1998) In press
REFERENCE   2  (bases 1 to 1776)
  AUTHORS   Clarke,S.C., Haigh,R.D. and Williams,P.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-1998) Microbiology and Immunology,
            University of Leicester, University Road, Leicester,
            Leicestershire LE1 9HN, UK
FEATURES             Location/Qualifiers
     source          1..1776
                     /organism="Escherichia coli"
                     /strain="enteropathogenic E2348/69"
                     /db_xref="taxon:562"
     gene            1..1776
                     /gene="typA"
     CDS             1..1767
                     /gene="typA"
                     /note="TypA; similar to Escherichia coli YihK"
                     /codon_start=1
                     /transl_table=11
                     /product="tyrosine phosphorylated protein A"

/translation="MIEKLRNIAIIAHVDHGKTTLVDKLLQQSGTFDSRAETQERVMD

SNDLEKERGITILAKNTAIKWNDYRINIVDTPGHADFGGEVERVMSMVDSVLLVVDAF

DGPMPQTRFVTKKAFAYGLKPIVVINKVDRPGARPDWVVDQVFDLFVNFDATDEQLDF

PIVYASALNGIAGLDHEDMAEDMTPLYQAIVDHVPAPDVDLDGPFQMQISQLDYNSYV

GVIGIGRIKRGKVKPNQQVTIIDSEGKTRNAKVGKVLGHLGLERIETDLAEAGDIVAI

TGLGELNISDTVCDTQNVEALPALSVDEPTVSMFFCVNTSPFCGKEGKFVTSRQILDR

LNKELVHNVALRVEETEDADAFRVSGRGELHLSVLIENMRREGFELAVSRPKVIFREI

DGRKQEPYENVTLDVEQQHQGSVMQALGERKGDLKNMNPDGKGRVRLDYVIPSRGLIG

FRSEFMTMTSGTGLLYSTFSHYDDVRPGEVGQRQNGVLISNGQGKAVAFALFGLQDRG

KLFLGHGAEVYEGQIIGIHSRSNDLTVNCLTGKKLTNMRASGTDEAVVLVPPIRMTLE

QALEFIHDDKLVEVTPISIRIR"
```

BASE COUNT 422 a 458 c 480 g 416 t

1 gtgatcgaaa aattgcgtaa tatcgccatc atcgcgcacg tagaccatgg taaaaccacc

61 ctggtagata agctgctcca acaatccggt acgttcgact ctcgtgccga aacccaagag

121 cgcgtgatgg actccaacga tttggagaaa gagcgtggga ttaccatcct cgcgaaaaac

181 accgctatca aatggaatga ttaccgtatc aacatcgttg ataccccggg gcacgccgac

241 ttcggtggtg aagttgaacg tgtaatgtcc atggtagact cagtgctgct ggtggttgac

301 gcatttgacg gcccgatgcc gcaaacgcgc ttcgtaacca aaaaagcgtt tgcttacggc

361 ctgaagccga ttgttgttat caacaaagtt gaccgccctg gcgcgcgtcc tgattgggtt

421 gtggatcagg tattcgatct gttcgttaac ttcgacgcga ccgacgagca gctggacttc

481 ccgatcgttt acgcttctgc gctgaacggt atcgcgggtc tggaccacga agatatggcg

541 gaagacatga ccccgctgta tcaggcgatt gttgaccacg ttcctgcgcc ggacgttgac

601 cttgacggtc cgttccagat gcagatttct cagctcgatt acaacagcta tgttggcgtt

661 atcggcattg gccgcatcaa gcgcggtaaa gtgaagccga accagcaggt tactatcatc

721 gatagcgaag gcaaaacccg caacgcgaaa gtcggtaaag tgctgggcca cctcggtctg

781 gaacgtatcg aaaccgatct ggcggaagct ggcgatatcg tggcgatcac gggtcttggc

841 gaactgaaca tttctgacac cgtttgcgac acgcaaaacg ttgaagcgct gccggcactc

901 tccgttgatg agccgaccgt ttctatgttc ttctgcgtta acacctcgcc gttctgcggt

961 aaagaaggta agttcgtaac gtctcgtcag atcctggatc gtctgaacaa agaactggta

1021 cacaacgttg cgctgcgcgt agaagaaacc gaagacgccg atgcgttccg cgtttctggt

1081 cgtggcgaac tgcacctgtc tgttcttatc gaaaacatgc gtcgtgaagg tttcgaactg

1141 gcggtatccc gtccgaaagt aatcttccgt gaaatcgacg gtcgtaaaca agagccgtat

1201 gaaaacgtga cgctggacgt tgaacaacag catcagggtt ctgtgatgca ggcgctgggc

1261 gaacgtaaag gcgacctgaa aaacatgaat ccagacggta aaggccgcgt acgtctcgac

1321 tacgtgatcc caagccgtgg tctgattggc ttccgttctg agttcatgac catgacttct

1381 ggtactggtc tgctgtactc caccttcagc cactacgacg acgtacgtcc gggtgaagtg

1441 ggtcagcgtc agaacggcgt actgatctct aacggtcagg gtaaagcggt agcgttcgcg

1501 ctgttcggtc tgcaggatcg cggtaagctg ttcctcggtc acggtgctga agtttacgaa

1561 ggtcagatta tcggtattca tagccgctct aacgacctga ctgtaaactg cctgaccggt

1621 aagaaactga ccaacatgcg tgcttccggt actgacgaag ccgttgttct ggttccgcct

1681 atccgcatga ctctggaaca agctctggaa ttcatccatg atgacaaact ggtagaagtg

1741 actccgatct ctatccgtat tcgttaacgt cactga//

Appendix II

Designation of parent and mutant *typA* strains

Designation	Relevant characteristics
SCC1001	E2348/69, Str^r
SCC1002	E2348/69 Str^r, *typA*
SCC1003	E2348/69 Nal^r, *typA*
SCC1004	MAR001, *typA*
SCC1005	JPN15, *typA*
SCC1006	RDEC-1, *typA*

Bibliography

Abe A, Kenny B, Stein M and Finlay BB (1997). Characterisation of two virulence proteins secreted by rabbit enteropathogenic *Escherichia coli*, EspA and EspB, whose maximal expression is sensitive to host body temperature. *Infection and Immunity* 65: 3547-3555.

Abe A, HeczkoU, Hegele RG and Finlay BB (1998). Two enteropathogenic *Escherichia coli* type III secreted proteins, EspA and EspB, are virulence factors. *Journal of Experimental Medicine* 188: 1907-1916.

Adams LM, Simmons CP, Rezmann L, Strugnell RA and Robins-Browne RM (1997). Identification and characterization of a K88- and CS31A-like operon of a rabbit enteropathogenic *Escherichia coli* strain which encodes fimbriae involved in the colonization of rabbit intestine. *Infection and Immunity* 65: 5222-5230.

Adu-Bobie J, Frankel G, Bain C, Goncalves AG, Trabulsi LR, Douce G, Knutton S and Dougan G (1998). Detection of intimins α, β, δ and γ, four intimin derivatives expressed by attaching and effacing microbial pathogens. *Journal of Clinical Microbiology* 36: 662-668.

Agin TS and Wolf MK (1997). Identification of a family of intimins common to *Escherichia coli* causing attaching-effacing lesions in rabbits, humans and swine. *Infection and Immunity* 65(1): 320-326.

Albert MJ, Faruque SM, Ansaruzzaman M, Islam MM, Haider K, Alam K, Kabir I and Robins-Browne R (1992). Sharing of virulence-associated properties at the phenotypic and genetic levels between enteropathogenic *Escherichia coli* and *Hafnia alvei*. *Journal of Medical Microbiology* 37: 310-314.

Albert MJ, Faruque SM, Faruque ASG, Ansaruzzaman M, Bhuiyan NA, Alan K and Akbar MS (1995). Controlled study of *Escherichia coli* diarrheal infections in Bangladeshi children. *Journal of Clinical Microbiology* 33: 973-977.

Albert MJ, Faruque SM, Faruque ASG, Bettelheim KA, Neogi PKB, Bhuiyan NA and Kaper JB (1996). Controlled study of cytolethal distending toxin-producing *Escherichia coli* infections in Bangladeshi children. *Journal of Clinical Microbiology* 34: 717-719.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 125: 3389-3402.

Anantha RP, Stone KD and Donnenberg MS (1998a). Non-polar mutagenesis of the *bfp* gene cluster from enteropathogenic *Escherichia coli*. *Abstracts of the 98th General Meeting of the American Society of Microbiology*, abstract B-155.

Anantha RP, Stone KD and Donnenberg MS (1998b). Role of BfpF, a member of the PilT family of putative nucleotide-binding proteins, in type IV pilus biogenesis and in interactions between enteropathogenic *Escherichia coli* and host cells. *Infection and Immunity* 66: 122-131.

Anderson R (1999). The pandemic of antibiotic resistance. *Nature Medicine* 147-149.

Anon (1996a). Outbreak of *Escherichia coli* O157:H7 infection - Georgia and Tennessee, June 1995. *Morbidity and Mortality Weekly Report* 45: 249-251.

Anon (1996b). Lake associated outbreak of *Escherichia coli* O157:H7 - Illinois, 1995. *Morbidity and Mortality Weekly Report* 45: 437-439.

Anon (1996c). Outbreak of *Escherichia coli* infection in Lanarkshire. *SCIEH Weekly Report* 30: 249.

Anon (1997). Outbreaks of *Escherichia coli* 0157:H7 infection and cryptosporidiosis associated with drinking unpasteurised apple cider - Connecticut and New York, October 1996. *Morbidity and Mortality Weekly Report* 46: 4-8.

Anon (1998). The QIAGEN guide template purification and DNA sequencing (2nd edn).

Ansorge W, Sproat B, Stegemann J, Schwager C and Zenke M (1987). Automated DNA sequencing - ultrasensitive detection of fluorescent bands during electrophoresis. *Nucleic Acids Research* 15: 4593-4602.

Atkinson M, Allen C and Sequeira L (1992). Tyrosine phosphorylation of a membrane protein from *Pseudomonas solanacearum*. *Journal of Bacteriology* 174: 4356-4360.

Atlung T and Ingmer H (1997). H-NS: a modulator of environmentally regulated gene expression. *Molecular Microbiology* 24: 7-17.

Ausubel F, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K eds. (1995). Short protocols in molecular biology, third edition. John Wiley and Sons.

Bachmann BJ (1996). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In: Neidhart FC (ed). *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington DC.

Bahrani FK, Sansonetti PJ and Parsot C (1997). Secretion of Ipa proteins by *Shigella flexneri*: inducer molecules and kinetics of activation. *Infection and Immunity* 65: 4005-4010.

Bain C, Keller R, Collington GK, Trabulsi LR and Knutton S (1998). Increased levels of intracellular calcium are not required for the formation of attaching and effacing lesions by enteropathogenic and enterhemorrhagic *Escherichia coli*. *Infection and Immunity* 66: 3900-3908.

Baldini MM, Kaper JB, Levine MM, Candy DCA and Moon HW (1983). Plasmid- mediated adhesion in enteropathogenic *Escherichia coli*. *Journal of Pediatric Gastroenterology and Nutrition* 2: 534-538.

Baldwin TJ (1998). Pathogenicity of enteropathogenic *Escherichia coli*. *Journal of Medical Microbiology* 47: 283-293.

Baldwin TJ, Brooks SF, Knutton S, Manjarrez-Hernandez HA, Aitken A and Williams PH (1990). Protein phosphorylation by protein kinase C in HEp-2 cells infected with enteropathogenic *Escherichia coli*. *Infection and Immunity* 58: 761-765.

Baldwin TJ, Knutton S, Sellers L, Manjarrez Hernandez HA, Aitken A and Williams PH (1992). Enteroaggregative *Escherichia coli* strains secrete a heat-labile toxin antigenically related to *E. coli* hemolysin. *Infection and Immunity* 60: 2092-2095.

Baldwin TJ, Ward W, Aitken A, Knutton S and Williams PH (1991). Elevation of intracellular free calcium levels in HEp-2 cells infected with enteropathogenic *Escherichia coli*. *Infection and Immunity* 59: 1599-1604.

Bäumler AJ (1997). The record of horizontal gene transfer in Salmonella. *Trends in Microbiology* 5: 318-322.

Beaudry M, Zhu C, Fairbrother JM and Harel J (1996). Genotypic and phenotypic characterization of *Escherichia coli* isolates from dogs manifesting attaching and effacing lesions. *Journal of Clinical Microbiology* 34: 144-148.

Beck S, O'Keeffe T, Coull JM and Koster H (1989). Chemi-luminescent detection of DNA - application for DNA sequencing and hybridisation. *Nucleic Acids Research* 17: 5155-5123.

Berlyn MKB (1998). Linkage map of *Escherichia coli* K-12, edition 10: the traditional map. *Microbiology and Molecular Biology Reviews* 62: 814-984.

Bieber D, Ramer SW, Wu CY, Murrary WJ, Tobe T, Fernandez R and Schoolnik GK (1998). Type IV pili, transient bacterial aggregates and virulence of enteropathogenic *Escherichia coli*. *Science* 280: 2114-2118.

Blanco JE, Blanco M, Blanco J, Mora A, Balaguer L, Cuervo L, Blasalobre C and Muñoa F (1997). Prevalence and characteristics of enteropathogenic *Escherichia coli* with the *eae* gene in diarrhoeic rabbits. *Microbiology and Immunology* 41: 77-82.

Blattner FR, Plunkett III G, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B and Shao Y (1997). The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1462.

Blomfield IC, Vaughn V, Rest RF and Eisenstein BI (1991). Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. *Molecular Microbiology* 5: 1447-1457.

Bokete TN, Whittam TS, Wilson RA, Clausen CK, O'Callahan CM, Moseley SL, Fritsche TR and Tarr PI (1997). Genetic and phenotypic analysis of *Escherichia coli* with enteropathogenic characteristics from Seattle children. *Journal of Infectious Diseases* 175: 1382-1389.

Bolton FJ and Aird H (1998). Verocytotoxin-producing *Escherichia coli* O157: public health and microbiological significance. *British Journal of Biomedical Science* 55: 127-135.

Bouzari S and Varghese A (1990). Cytolethal distending toxin (CLDT) production by enteropathogenic *Escherichia coli* (EPEC). *FEMS Microbiology Letters* 71: 193-198.

Bouzari S, Vatsala BR and Varghese A (1992). *In vitro* adherence property of cytolethal distending toxin (CLDT) producing EPEC strains and effect of the toxin on rabbit intestine. *Microbial Pathogenesis* 12: 153-157.

Bray J (1945). Isolation of antigenically homogenous strains of *Bact. coli neapolitanum* from summer diarrhoea of infants. *Journal of Pathology* 57: 239-247.

Bray J and Bevan TED (1948). Slide agglutination of *Bacterium coli* var. *neopolitanum* in summer diarrhoea. *Journal of Pathology and Bacteriology* 60: 395-401.

Brickman E and Beckwith J (1975). Analysis of the regulation of *Escherichia coli* alkaline phosphatase synthesis using deletions and phi80 transducing phages. *Journal of Molecular Biology* 96: 307-316.

Brockman RW and Heppel LA (1968). On the localisation of alkaline phosphatase and cyclic phosphodiesterase in *Escherichia coli*. *Biochemistry* 7: 2554-2562.

Buchrieser C, Prentice M and Carniel E (1998). The 102-Kilobase unstable region of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement. *Journal of Bacteriology* 180: 2321-2329.

Canil C, Rosenshine I, Ruschkowski S, Donnenberg MS, Kaper JB and Finlay BB (1993). Enteropathogenic *Escherichia coli* decreases the transepithelial electrical resistance of polarized epithelial monolayers. *Infection and Immunity* 61: 2755-2762.

Cantey JR and Blake RK (1977). Diarrhea due to *Escherichia coli* in the rabbit: a novel mechanism. *Journal of Infectious Diseases* 135: 454-462.

Chan M-S (1997). The global burden of intestinal nematode infections - fifty years on. *Parasitology Today* 13: 438-443.

Chen W and Kuo T (1993). A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. *Nucleic Acids Research* 21: 2260.

Chiang TM, Reizer J and Beachey EH (1989). Serine and tyrosine protein kinase activities in *Streptococcus pyogenes*. *Journal of Biological Chemistry* 264: 2957-2962.

Chu G (1990). Pulsed-field gel electrophoresis: theory and practice. *Methods: A companion to Methods in Enzymology* 1: 129-142.

Clark A, Morton S, Wright P, Corkish J, Bolton FJ and Russell J (1997). A community outbreak of Vero cytotoxin producing *Escherichia coli* O157 infection linked to a small farm dairy. *Communicable Disease Report* 7: R206-R211.

Clarke SC, Haigh R and Williams PH (1998). The complete sequence of the *typA* gene from enteropathogenic *Escherichia coli* E2348/69. *Abstracts of the American Society for Microbiology 38th Interscience Conference on Antimicrobial Agents and Chemotherapy*, abstract B-11.

Clausen CR and Christie DL (1982). Chronic diarrhea in infants caused by adherent enteropathogenic *Escherichia coli*. *Journal of Pediatrics* 100: 358-361.

Coia JE (1998). Clinical, microbiological and epidemiological aspects of *Escherichia coli* O157 infection. *FEMS Immunology and Medical Microbiology* 20: 1-9.

Collazo CM and Galán J (1997). The invasion-associated type III system of *Salmonella typhimurium* directs the translocation of Sip proteins into the host cell. *Molecular Microbiology* 24: 747-756.

Collington GK, Booth IW and Knutton S (1998a). Rapid modulation of electrolyte transport in Caco-2 cell monolayers by enteropathogenic *Escherichia coli* (EPEC) infection. *Gut* 42: 200-207.

Collington GK, Booth IW, Donnenberg MS, Kaper JB and Knutton S (1998b). Enteropathogenic *Escherichia coli* virulence genes encoding secreted signalling proteins are essential for modulation of Caco-2 cell electrolyte transport. *Infection and Immunity* 66: 6049-6053.

Connell C, Fung S, Heiner C, Bridgham J, Chakerian V, Heron E, Jones B, Menchen S, Mordan W, Raff M, Recknor M, Smith L, Springer J, Woo S and Hunkapillar M (1987). Automated DNA sequence analysis. *Biotechniques* 5: 342.

Cook GC (1996). Manson's Tropical Diseases, 20th edn. WB Saunders Co. Ltd, London.

Copass M, Grandi G and Rappuoli R (1997). Introduction of unmarked mutations in the *Helicobacter pylori vac*A gene with a sucrose sensitivity marker. *Infection and Immunity* 65: 1949-1952.

Cornelis GR (1998). The *Yersinia* deadly kiss. *Journal of Bacteriology* 180(21): 5495-5504.

Cornelis GR and Wolf-Watz H (1997). The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. *Molecular Microbiology* 23(5): 861-867.

Cortay JC, Duclos B and Cozzone AJ (1986a). Phosphorylation of an *Escherichia coli* protein at tyrosine. *Journal of Molecular Biology* 187: 305-308.

Cortay JC, Rieul C, Duclos B and Cozzone AJ (1986b). Characterisation of the phosphoproteins of *Escherichia coli* cells by electrophoretic analysis. *European Journal of Biochemistry* 159: 227-237.

Cozzone AJ (1988). Protein phosphorylation in prokaryotes. *Annual Review of Microbiology* 42: 97-125.

Cravioto A, Gross RJ, Scotland SM and Rowe B (1979). An adhesive factor found in strains of *Escherichia coli* belonging to the traditional infantile enteropathogenic serotypes. *Current Microbiology* 3: 95-99.

Crowley DS, Ryan MJ and Wall PG (1997). Gastroenteritis in children under 5 years of age in England and Wales. *Communicable Disease Report* 7: R82-R86.

Czeczulin JR, Balepur S, Hicks S, Phillips A, Hall R, Kothary MH, Navarro-Garcia F and Nataro JP (1997). Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infection and Immunity* 65: 4135-4145.

Dadssi M and Cozzone AJ (1990). Evidence of protein-tyrosine kinase activity in the bacterium *Acinetobacter calcoaceticus*. *Journal of Biological Chemistry* 265: 20996-20999.

Daniels DL (1996). Physical mapping of bacterial genomes. In: Neidhart FC (eds). *Escherichia coli* and Salmonella - cellular and molecular biology, 2nd edn. ASM Press, Washington D.C.

Dean-Nystrom EA, Bosworth BT, Cray WC and Moon HW (1997). Pathogenicity of *Escherichia coli* O157:H7 in the intestines of neonatal calves. *Infection and Immunity* 65: 1842-1848.

Dehio C, Prevost MC and Sansonetti PJ (1995). Invasion of epithelial cells by *Shigella flexneri* induces tyrosine phosphorylation of cortactin by a pp60c-src-mediated signalling pathway. *EMBO Journal* 14: 2471-2482

Deibel C, Krämer S, Chakraborty T and Ebel F (1998). EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90kDa protein. *Molecular Microbiology* 28: 463-474.

deLorenzo V and Timmis (1994). Analysis and construction of stable phenotypes in Gram-negative bacteria with Tn*5*- and Tn*10*-derived minitransposons. *Methods in Enzymology* 235: 386-405.

Doig P, Yao R, Burr DH, Guerry P and Trust TJ (1996). An environmentally regulated pilus-like appendage involved in *Campylobacter* pathogenesis. *Molecular Microbiology* 20: 885-894.

Donnenberg MS, Calderwood SB, Donohue-Rolfe A, Keusch GT and Kaper JB (1990). Construction and analysis of Tn*phoA* mutants of enteropathogenic *Escherichia coli* unable to invade HEp-2 cells. *Infection and Immunity* 58: 1565-1571.

Donnenberg MS, Donohue-Rolfe A and Jeusch GT (1989). Epithelial cell invasion: an overlooked property of enteropathogenic *Escherichia coli* (EPEC) associated with the EPEC adherence factor. *Journal of Infectious Diseases* 160: 452-459.

Donnenberg MS, Girón JA, Nataro JP and Kaper JB (1992). A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localised adherence. *Molecular Microbiology* 6:3427-3437.

Donnenberg MS and Kaper JB (1991). Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infection and Immunity* 59: 4310-4317.

Donnenberg MS, Tacket CO, Losonsky, Frankel G, Nataro JP, Dougan G and Levine MM (1998). Effect of prior experimental human enteropathogenic *Escherichia coli* infection on illness following homologous and heterologous rechallenge. *Infection and Immunity* 66: 52-58.

Donnenberg MS, Yu J and Kaper JB (1993). A second chromosomal gene necessary for intimate attachment of enteropathogenic *Escherichia coli* to epithelial cells. *Journal of Bacteriology* 175: 4670-4680.

Douce GR, Amin II and Stephen J (1991). Invasion of HEp-2 cells by strains of *Salmonella typhimurium* of different virulence in relation to gastroenteritis. *Journal of Medical Microbiology* 35: 349-403.

Dower WJ, Miller JF and Ragsdale CW (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Research* 16: 6127-6145.

Drolet R, Fairbrother JM and Vaillancourt D (1994). Attaching and effacing *Escherichia coli* in a goat with diarrhea. *Canadian Veterinary Journal* 35: 122-123.

Duclos B, Grangeasse C, Vaganay E, Riberty M and Cozzone AJ (1996). Auto-phosphorylation of a bacterial protein at tyrosine. *Journal of Molecular Biology* 259: 8910895.

Duncan HE and Edberg SC (1995). Host-microbe interaction in the gastrointestinal tract. *Critical Reviews in Microbiology* 21: 85-100.

DuPont HL and Ericsson CD (1993). Prevention and treatment of traveller's diarrhoea. *The New England Journal of Medicine* 328: 1821-1827.

DuPont HL (1995). Diarrheal diseases in the developing world. *Infectious Disease Clinics of North America* 9: 313-324.

Dytoc M, Gold B, Louie M, Huesca M, Fedorko L, Crowe S, Lingwood C, Brunton J and Sherman P (1993). Comparison of *Helicobacter pylori* and attaching-effacing *Escherichia coli* adhesion to eukaryotic cells. *Infection and Immunity* 61: 448-456.

Dytoc MT, Ismaili A, Pilpott DJ, Soni R, Brunton JL and Sherman PM (1994): Distinct binding properties of *eaeA*-negative verocytotoxin-producing *Escherichia coli* of serotype O113:H21. *Infection and Immunity* 62: 3494-3505.

Edelman R and Levine MM (1983). Summary of a workshop of enteropathogenic *Escherichia coli*. *Journal of Infectious Diseases* 147: 1108-1118.

Edwards RA and Puente JL (1998). Fimbrial expression in enteric bacteria: a critical step in intestinal pathogenesis. *Trends in Microbiology* 6: 282-287.

Ellinghorst EA (1994). Measurement of invasion by gentamicin resistance. In: *Methods in Enzymology*, Volume 236. Bacterial pathogenesis, part B - Interaction of pathogenic bacteria with host cells: 405-420.

Elliot SJ and Nataro JP (1995). Enteroaggregative and diffusely adherent *Escherichia coli*. *Reviews in Medical Microbiology* 6: 196-206.

Elliot SJ, Srinivas S, Albert MJ, Alam K, Robins-Browne RM, Gunzburg ST, Mee BJ and Chang BJ (1998a). Characterisation of the roles of hemolysin and other

toxins in enteropathy caused by alpha-hemolytic *Escherichia coli* linked to human diarrhoea. *Infection and Immunity* 66: 2040-2051.

Elliot SJ, Wainwright LA, McDaniel TK, Jarvis KG, Deng Y, Lai L, McNamara BP, Donnenberg MS and Kaper JB (1998b). The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Molecular Microbiology* 28: 1-4.

Embaye H, Batt RM, Farrant J, Fletcher J, Getty B, Saunders JR and Hart CA (1993). Evidence that enteropathogenic *Escherichia coli* may not cause seceretory diarrhea. *Gastroenterology* 104: abstract.

Escherich T (1885). Die Darmbakterien des Neugeborenen und Sauglings. *Fortschr Med* 3: 515-522, 547-554.

Eslava C, Navarro-García F, Czeczulin JR, Henderon IR, Cravioto A and Nataro JP (1998). Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infection and Immunity* 66: 3155-3163.

Evans MR and Read CA (1992). ^{32}P, ^{33}P and ^{35}S - selecting a label for nucleic acid analysis. *Nature* 358: 520-521.

Ewing WH, Davis BR and Montague TS (1963). Studies on the occurrence of *Escherichia coli* serotypes associated with diarrheal disease. Atlanta, Ga: US Department of Health, Education and Welfare.

Fagundes-Neto U, Freymuller E, Schmitz LG and Scaletsky I (1996). Nutritional impact and ultrastructural intestinal alteration in severe infections due to enteropathogenic *Escherichia coli* strains in infants. *Journal of the American College of Nutrition* 15(2): 180-185.

Fällman M, Persson C and Wolf-Watz H (1997). *Yersinia* proteins that target host cell signalling pathways. *Journal of Clinical Investigation* 99: 1153-1157.

Farthing MJG (1994). Travellers' diarrhoea. *Medicine International* 266-271.

Farris M, Grant A, Richardson TB and O'Connor CD (1998). BipA: a tyrosine-phosphorylated GTPase that mediates interaction between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells. *Molecular Microbiology* 28: 265-279.

Fiederling F, Boury M, Petit C and Milon A (1997). Adhesive factor/rabbit 2, a new fimbrial adhesin and a virulence factor from *Escherichia coli* O103, a serogroup enteropathogenic for rabbits. *Infection and Immunity* 65: 847-851.

Finlay BB and Falkow S (1997). Common themes in microbial pathogenicity revisited. *Microbiology and Molecular Biology Reviews* 61: 136-169.

Foster G, Ross HM, Pennycott TW, Hopkins GF and McLaren (1998). Isolation of *Escherichia coli* O86:K61 producing cyto-lethal distending toxin from wild birds of the finch family. *Letters in Applied Microbiology* 26: 395-398.

Foster PL (1992). *Escherichia coli* and *Salmonella typhimurium* mutagenesis. In: Lederberg (ed). *Encyclopedia of Microbiology*, volume 2: 107-114. Academic Press, London.

Foster R, Thorner J and Martin GS (1989). Nucleotidylation, not phosphorylation, is the major source of the phosphotyrosine detected in enteric bacteria. *Journal of Bacteriology* 171: 272-279.

Foubister V, Rosenshine I, Donnenberg MS and Finlay BB (1994a). The *eae*B gene of enteropathogenic *Escherichia coli* is necessary for signal transduction in epithelial cells. *Infection and Immunity* 64: 3038-3040.

Foubister V, Rosenshine I and Finlay BB (1994b). A diarrheal pathogen, enteropathogenic *Escherichia coli*, triggers a flux of inositol phosphates in infected epithelial cells. *Journal of Experimental Medicine* 179: 993-998.

Fraenkel DG (1996). Glycolysis. In: Neidhart FC (ed). *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington DC.

Francis CL, Jerse AE, Kaper JB and Falkow S (1991). Characterisation of interactions of enteropathogenic *Escherichia coli* O127:H6 with mammalian cells *in vitro*. *Journal of Infectious Diseases* 164: 693-703.

Frant S and Abramson H (1938). Epidemic diarrhea of the new-born. III. Epidemiology of outbreaks of highly fatal diarrhea among new-born babies in hospital nurseries. *American Journal of Public Health* 28:36-43.

Freestone P, Grant S, Toth I and Norris V (1995). Identification of phosphoproteins in *Escherichia coli*. *Molecular Microbiology* 15: 573-580.

Freestone P, Grant S, Trinei M, Onada T and Norris V (1998a). Protein phosphorylation in *Escherichia coli* L. form NC-7. *Microbiology* 144: 3289-3295.

Freestone P, Nyström T, Trinei M and Norris V (1997). The universal stress protein, UspA, of *Escherichia coli* is phosphorylated in response to stasis. *Journal of Molecular Biology* 274: 318-324.

Freestone P, Trinei M, Clarke S and Norris V (1998b). Tyrosine phosphorylation in *Escherichia coli*. *Journal of Molecular Biology* 279: 1045-1051.

Fukui H, Sueyoshi M, Haritani M, Nakazawa M, Naitoh S, Tani H and Uda Y (1995). Natural infection with attaching and effacing *Escherichia coli* (O103:H$^-$) in chicks. *Avian Diseases* 39: 912-918.

Gabastou JM, Kernéis S, Bernet-Carnard MF, Coconnier MM, Kaper JB and Servin AL (1995). Two stages of enteropathogenic *Escherichia coli* intestinal pathogenicity are up and down regulated by epithelial cell differentiation. *Differentiation* 59: 127-134.

Garrod LP, Lambert HP and O'Grady F (1981). Antibiotic and chemotherapy. 5th edn. Churchill Livingstone, London.

Gascon J, Vargas M, Quinto L, Corachan M, Jimenez de Anta MT and Vila J (1998). Enteroaggregative *Escherichia coli* strains as a cause of traveler's diarrhea: a case-control study. *Journal of Infectious Diseases* 177: 1409-1412.

Gautom RK (1997). Rapid pulsed-field gel electrophoresis protocol for typing *Escherichia coli* O157:H7 and other Gram-negative organisms in one day. *Journal of Clinical Microbiology* 35: 2977-2980.

Geissler S and Drummond M (1993). A counterselectable pACYC184-based *lazZ* alpha-complementing plasmid vector with novel multiple cloning sites: construction of chromosomal deletions in *Klebsiella pneumoniae*. *Gene* 136: 253-255.

Giammanco A, Maggio M, Giammanco G, Morelli R, Minelli F, Scheutz F and Caprioli A (1996). Characteristics of *Escherichia coli* serogroups isolated in Italy from children with diarrhea. *Journal of Clinical Microbiology* 34: 689-694.

Gianella RA, Formal SB, Dammin GJ and Collins H (1973). Pathogenesis of salmonellosis. Studies of fluid secretion, mucosal invasion, and morphologic reaction in the rabbit ileum. *Journal of Clinical Investigation* 52: 441-453.

Giesecke H, Obermaier B, Domdey H and Neubert WJ (1992). Rapid sequencing of the Sendai virus 6.8kb large (L) gene through primer walking with an automated DNA sequencer. *Journal of Virological Methods* 38: 47-60.

Girón JA, Gómez-Duarte OG, Jarvis KG and Kaper JB (1997). Longus pilus of enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili - a minireview. *Gene* 192: 39-43.

Girón JA, Ho ASY and Schoolnik GK (1991). An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. *Science* 254:710-713.

Girón JA, Ho ASY and Schoolnik GK (1993). Characterization of fimbriae produced by enteropathogenic *Escherichia coli*. *Journal of Bacteriology* 175: 7391-7403.

Girón JA, Jones T, Millán-Velasco F, Castro-Muñoz, Zárate L, Fry J, Frankel G, Moseley SL, Baudry B, Kaper JB, Schoolnik GK and Riley LW (1991). Diffuse-adhering *Escherichia coli* (DAEC) as a putative cause of diarrhea in Mayan children in Mexico. *Journal of Infectious Diseases* 163: 507-513.

Girón JA, Qadri F, Azim T, Jarvis KJ, Kaper JB and Albert MJ (1995). Monoclonal antibodies specific for the bundle-forming pilus of enteropathogenic *Escherichia coli*. *Infection and Immunity* 63: 4949-4952.

Goldschmidt MC and DuPont HL (1976). Enteropathogenic *Escherichia coli*: lack of correlation of serotype with pathogenicity. *Journal of Infectious Diseases* 133: 153-156.

Gómez-Duarte OG and Kaper JB (1995). A plasmid-encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. *Infection and Immunity* 63: 1767-1776.

Grangeasse C, Doublet P, Vaganay E, Vincent C, Deleage G, Duclos B and Cozzone AJ (1997). Characterisation of bacterial gene encoding an auto-phosphorylating protein tyrosine kinase. *Gene* 204: 259-265.

Groisman EA and Ochman H (1997). How Salmonella became a pathogen. *Trends in Microbiology* 5: 343-348.

Gross RJ (1990). *Escherichia coli* diarrhoea, chapter 3.25 In: Topley and Wilson's Principles of Bacteriology, Virology and Immunity. Volume 3: Bacterial diseases. Edward Arnold, London.

Guerrant RL and Bobak DA (1991). Bacterial and protozoal gastroenteritis. *New England Journal of Medicine* 325: 327-340.

Guerrant RL, Shields DS, Thorson SM, Schorling JB and Groschel DH (1985). Evaluation and diagnosis of acute infectious diarrhea. *American Journal of Medicine* 78: 91-98.

Gunzburg ST, Chang BJ, Elliot SJ, Burke V and Gracey M (1993). Diffuse and enteroaggregative patterns of adherence of enteric *Escherichia coli* isolated from Aboriginal children from the Kimberley region of Western Australia. *Journal of Infectious Diseases* 167: 755-758.

Gunzberg ST, Tornieporth NG and Riley LW (1995). Identification of enteropathogenic *Escherichia coli* by PCR-based detection of the bundle-forming pilus gene. *Journal of Clinical Microbiology* 33: 1375-1377.

Haigh R, Baldwin T, Knutton S and Williams PH (1995). Carbon dioxide regulated secretion of the EaeB protein of enteropathogenic *Escherichia coli. FEMS Microbiology Letters* 129: 63-68.

Hacker J (1992). Role of fimbrial adhesins in the pathogenesis of *Escherichia coli* infections. *Canadian Journal of Microbiology* 38: 720-727.

Hacker J, Blum-Oehler G, Mühldorfer I and Tschäpe H (1997). Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Molecular Microbiology* 23: 1089-1097.

Hanahan D (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* 166: 557-580.

Hardie DG, ed. (1993). Protein phosphorylation - a practical approach. IRL Press, Oxford.

Hart CA, Batt RM and Saunders JR (1993). Diarrhoea caused by *Escherichia coli. Annals of Tropical Paediatrics* 13: 121-131.

Heithoff DM, Conner CP and Mahan MJ (1997). Dissecting the biology of a pathogen during infection. *Trends in Microbiology* 5: 509-513.

Henikoff S (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 26: 351.

Hensel M and Holden DW (1996). Molecular genetic approaches for the study of virulence in both pathogenic bacteria and fungi. *Microbiology* 142: 1049-1058.

Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E and Holden DW (1995). Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269: 400-403.

Hicks S, Frankel G, Kaper JB, Dougan G and Phillips AD (1998). Role of intimin and bundle-forming pili in enteropathogenic *Escherichia coli* adhesion to pediatric intestinal tissue *in vitro*. *Infection and Immunity* 66: 1570-1578.

Hill SM, Phillips AD and Walker-Smith JA (1991). Enteropathogenic *Escherichia coli* and life threatening chronic diarrhoea. *Gut* 3: 154-158.

Hueck C (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and Molecular Biology Reviews* 62: 379-433.

Hulan S, Zhen LG, Mathan MM, Mathew MM, Olarte J, Espejo R, Khin Maung U, Ghafoor MA, Khan MA and Sami Z (1991). Etiology of acute diarrhoea among children in developing countries: a multicentre study in five countries. *Bulletin of the World Health Organisation* 69: 549-555.

Hultgren SJ, Jones HC and Normack S (1996). Bacterial adhesins and their assembly. In: Neidhart FC (eds). *Escherichia coli* and Salmonella - cellular and molecular biology, 2nd edn. ASM Press, Washington D.C.

Jarvis KG, Girón JA, Jerse AE, McDaniel TK, Donnenberg MS and Kaper JB (1995). Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proceedings of the National Academy of Science USA* 92: 7996-8000.

Jerse AE and Kaper JB (1991). The *eae* gene of enteropathogenic *Escherichia coli* encodes a 94-kilodalton membrane protein, the expression of which is influenced by the EAF plasmid. *Infection and Immunity* 59: 4302-4309.

Jerse AE, Yu J, Tall BD and Kaper JB (1990). A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. *Proceedings of the National Academy of Science USA* 87: 7839-7843.

Kaniga K, Delor I and Cornelis GR (1991). A wide-host-range suicide vector for improving reverse genetics in Gram-negative bacteria: inactivation of the *bla*A gene of *Yersinia enterocolitica*. *Gene* 109: 137-141.

Kaper JB (1994). Molecular Pathogenesis of enteropathogenic *Escherichia coli*. Chapter 12, 173-195. In: Miller VL, Kaper JB, Portnoy DA and Isberg RR (eds.) Molecular Genetics of Bacterial Pathogenesis. American Society of Microbiology.

Kaper and Gómez-Duarte OG (1997). The *per* regulator of enteropathogenic *Escherichia coli*. *Molecular Microbiology* 23(1): 179-181.

Kaper JB, McDaniel TK, Jarvis KG and Gómez-Duarte OG (1997). Genetics of virulence of enteropathogenic *Escherichia coli*. In: Mechanisms in the Pathogenesis of Enteric Diseases, edited by Paul *et al*. Plenum Press, New York.

Kauffman F (1947). The serology of the *coli* group. *Journal of Immunology* 57: 71-100.

Keenleyside WJ and Whitfield C (1995). Lateral transfer of *rfb* genes: a mobilizable ColE1-type plasmid carries the *rfb*0;54 (0;54 antigen biosynthesis) Gene cluster from *Salmonella enterica* serovar *borreze*. *Journal of Bacteriology* 177: 5247-5253.

Kelly-Wintenberg K, Anderson T and Montie TC (1990). Phosphorylated tyrosine in the flagellum filament protein of *Pseudomonas aeruginosa*. *Journal of Bacteriology* 172: 5135-5139.

Kennelly PJ and Potts M (1996). Fancy meeting you here! a fresh look at "prokaryotic" protein phosphorylation. *Journal of Bacteriology* 178: 4759-4764.

Kenny B (1999). Phosphorylation of tyrosine 474 of the enteropathogenic *Escherichia coli* (EPEC) Tir receptor molecule is essential for actin nucleating activity and is preceded by additional host modifications. *Molecular Microbiology* 31: 1229-1241.

Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA and Finlay BB (1997). Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell* 91: 511-520.

Kenny B, Lai LC, Finlay BB and Donnenberg MS (1996). EspA, a protein secreted by enteropathogenic *Escherichia coli*, is required to induce signals in epithelial cells. *Molecular Microbiology* 20: 313-323.

Kleckner N, Roth J and Botstein D (1977). Genetic engineering *in vivo* using translocatable drug-resistance elements. *Journal of Molecular Biology* 116: 125-159.

Knutton S, Baldini MM, Kaper JB and McNeish AS (1987) Role of plasmid encoded adherence factors in adhesion of enteropathogenic *Escherichia coli* to HEp-2 cells. *Infection and Immunity* 55: 78-85.

Knutton S, Baldwin T, Williams PH and McNeish AS (1989). Actin accumulation at sites of bacterial adhesion to tissue culture cells: basis of a new diagnostic test for enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infection and Immunity* 57: 1290-1298.

Knutton S, Phillips AD, Smith HR, Gross RJ, Shaw R, Watson P and Price E (1991). Screening for enteropathogenic *Escherichia coli* in infants with diarrhea by the fluorescent-actin staining test. *Infection and Immunity* 59: 365-371.

Knutton S, Rosenshine I, Pallen MJ, Nisan I, Neves BC, Bain C, Wolff C, Dougan G and Frankel G (1998). A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *EMBO Journal* 17: 2166-2176.

Knutton S, Shaw RK, Anantha RP, Donnenberg MS and Zorgani AA (1999). The type IV bundle-forming pilus of enteropathogenic Escherichia coli undergoes dramatic alterations in structure associated with bacterial adherence, aggregation and dispersal. Molecular Microbiology 33(3): 499-500.

Kovalenko SA, Tanaka M and Ozawa T (1994). Simple methods for preparation of plasmid DNA yielding long and accurate sequence data. *Nucleic Acids Research* 22: 5771-5772.

Kresse AU, Schulze K, Deibel C, Ebel F, Rohde M, Chakraborty T and Guzman CA (1998). Pas, a novel protein required for protein secretion and attaching and effacing activities of enterohemorrhagic *Escherichia coli*. *Journal of Bacteriology* 180: 4370-4379.

Kumate J and Isibasi A (1986). Pediatric diarrheal diseases: a global perspective. *Pediatric Infectious Diseases* 5: S21-S28.

Kuo JF and Greengard P (1969). An adenosine 3', 5'-monophosphate-dependent protein kinase from *Escherichia coli*. *Journal of Biological Chemistry* 224: 3417-3419.

Kyte J and Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157: 105-132.

Lai LC, Wainwright LA, Stone KD and Donnenberg MS (1997). A third secreted protein that is encoded by the enteropathogenic *Escherichia coli* pathogenicity island is required for transduction of signals and for attaching and effacing activities in host cells. *Infection and Immunity* 65: 2211-2217.

Laurelle L (1889). Etude bacteriologique sur les peritonites par perforation. *La Cellule* 5: 60-123.

Law D (1994). Adhesion and its role in the virulence of enteropathogenic *Escherichia coli*. *Clinical Microbiological Reviews* 7: 152-173.

Law D and Chart H (1998). Enteroaggregative *Escherichia coli*. *Journal of Applied Microbiology* 84: 685-697.

Lee CA (1996). Pathogenicity islands and the evolution of bacterial pathogens. *Infectious Agents and Disease* 5: 1-7.

Lee CA (1997). Type III secretion systems: machines to deliver bacterial proteins into eukaryotic cells? *Trends in Microbiology* 5: 148-156.

Lesage AA (1897). Contribution á l'étude des enterites infantiles - sero-diagnostic des races de *Bacterium coli*. *Comptes Rendues des Societe Biologique (Paris)* 49: 900-901.

Levine MM (1987). *Escherichia coli* that cause diarrhea - enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *Journal of Infectious Diseases* 155: 377-389.

Levine MM, Bergquist EJ, Nalin DR, Waterman DH, Hornick BB, Young CR and Sotman S (1978). *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* May 27: 1119-1122.

Levine MM, Losonsky G, Herrington D, Kaper JB, Tacket C, Bennels MB and Morris JG (1986). Pediatric diarrhea: the challenge of prevention. *Pediatric Infectious Diseases* 5: S29-S43.

Levine MM, Nataro JP, Karch MM, Baldini MM, Kaper JB, Black RE, Clements ML and O'Brien AD (1985). The diarrheal response of man to some classical serotype

enteropathogenic *Escherichia coli* is dependent on a plasmid encoding an enteroadhesiveness factor. *Journal of Infectious Diseases* 152: 550-559.

Lewis T and Matsui (1993). Viral gastroenteritis agents: Norwalk virus and astrovirus. *Virus and Life*, March 2-5.

Lim YS, Ngam CCL and Tay L (1992). Enteropathogenic *Escherichia coli* as a cause of diarrhoea among children in Singapore. *Journal of Tropical Medicine and Hygiene* 95: 338-342.

Link AJ, Phillips D and Church GM (1997). Methods for generating precise deletion and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterisation. *Journal of Bacteriology* 179: 6228-6237.

Lolkema JS and Slotboom DJ (1998). Hydropathy profile alignment: a tool to search for structural homologues of membrane proteins. *FEMS Microbiology Reviews* 22: 305-322.

Londesborough J. Phosphorylation of proteins in *Clostridium thermohydrosulfuricum*. *Journal of Bacteriology* 165: 595-601.

Mahan MJ (1996). Environmental regulation of virulence gene expression in *Escherichia*, *Salmonella* and *Shigella* spp. In: Neidhart FC (eds). *Escherichia coli* and Salmonella- cellular and molecular biology, 2nd edn. ASM Press, Washington.

Mahan MJ, Slauch JM and Makalanos JJ (1993). Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* 259: 666-668.

Mahan MJ, Tobias JW, Slauch JM, Collier RJ and Mekalanos JJ (1995). Antibiotic-based selection for bacterial genes that are specifically induced during infection of a host. *Proceedings of the National Academy of Sciences USA* 92: 669-673.

Manjarrez-Hernandez HA, Amess B, Sellers L, Baldwin TJ, Knutton S, Williams PH and Aitken A (1991). Purification of a 20kDa phosphoprotein from epithelial

cells and identification as a myosin light chain. *FEBS Microbiology Letters* 292: 121-127.

Manjarrez-Hernandez HA, Baldwin TJ, Aitken A, Knutton S and Williams PH (1992). Intestinal epithelial cell protein phosphorylation in enteropathogenic *Escherichia coli* diarrhoea. *Lancet* 339: 521-523.

Manoil C and Beckwith J (1985). Tn*phoA*: a transposon probe for protein export signals. *Proceedings National Academy of Science USA* 82: 8129-8133.

Maxam AM and Gilbert W (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Science* 74(2): 560-564.

Mayatepek E, Seebass E, Hingst V, Kroeger A and Sonntag HG (1993). Prevalence of enteropathogenic and enterotoxigenic *Escherichia coli* in children with and without diarrhoea in Esteli, Nicaragua. *Journal of Diarrhoeal Disease Research* 11: 169-171.

McCartney B, Howell LD, Kennelly PJ and Potts M (1997). Protein tyrosine phosphorylation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Journal of Bacteriology* 179: 2314-2318.

McDaniel TK, Jarvis KG, Donnenberg MS and Kaper JB (1995). A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proceedings of the National Academy of Science USA* 92: 1664-1668.

McDaniel TK and Kaper JB (1997). A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Molecular Microbiology* 23: 399-407.

McDonnell RJ, Rampling A, Crook S, Cockcroft PM, Wilshaw GA, Cheasty T and Stuart J (1997). An outbreak of Vero cytotoxin producing *Escherichia coli* O157 infection associated with takeaway sandwiches. *Communicable Disease Report* 7: R201-R205.

McNamara BP and Donnenberg MS (1998). A novel proline-rich protein, EspF, is secreted from enteropathogenic *Escherichia coli* via the type III export pathway. *FEMS Microbiology Letters* 166: 71-78.

Merryweather A, Bernander R and Nordström (1987). Direct selection for the exchange of alleles between a plasmid and the *Escherichia coli* chromosome. *Molecular and General Genetics* 210: 381-384.

Middendorf LR, Bruce JC, Bruce RC, Eckles RD, Grone DL, Roemer SC, Sloniker GD, Steffens DL, Sutter SL, Brumbaugh JA and Patonay G (1992). Continuous, online DNA sequencing using a versatile infrared laser scanner electrophoresis apparatus. *Electrophoresis* 13(8): 487-494.

Miller JF (1994). Bacterial transformation by electroporation. *Methods in Enzymology* 235: 375-385.

Miller VL and Mekalanos JJ (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence in *Vibrio cholerae* requires *toxR*. *Journal of Bacteriology* 170: 2575-2583.

Mitchell C, Morris PW and Vary JC (1992). Identification of proteins phosphorylated by ATP during sporulation of *Bacillus subtilis*. *Journal of Bacteriology* 174: 2474-2477.

Moffet HL, Shulenberge HK and Burkholder ER (1968). Epidemiology and etiology of severe infantile diarrhea. *Journal of Pediatrics* 72: 1-14.

Montiero-Neto V, Campos LC, Ferreira AJP, Gomes TAT and Trabulsi LR (1997). Virulence properties of *Escherichia coli* O111:H12 strains. *FEMS Microbiology Letters* 146: 123-128.

Moon HW, Whipp SC, Argenzio RA, Levine MM and Gianella RA (1983). Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infection and Immunity* 41: 1340-1251.

Morabito S, Karch H, Mariani-Kurkdjian P, Schmidt H, Mirelli F, Bingen E and Caprioli A (1998). Enteroaggregative, Shiga-toxin-producing *Escherichia coli* O111:H2 associated with an outbreak of haemolytic-uraemic syndrome. *Journal of Clinical Microbiology* 36: 840-842.

Morris KJ and Rao GG (1992). Conventional screening for enteropathogenic *Escherichia coli* in the UK. Is it appropriate or necessary? *Journal of Hospital Infection* 21: 163-167.

Nataro JP and Kaper JB (1998). Diarrheagenic *Escherichia coli. Clinical Microbiology Reviews* 11: 142-201.

Nataro JP, Maher KO, Mackie P and Kaper JB (1987). Characterization of plasmids encoding the adherence factor of enteropathogenic *Escherichia coli. Infection and Immunity* 55: 2370-2377.

Nataro JP, Scaletsky ICA, Kaper JB, Levine MM and Trabulsi LR (1985). Plasmid-mediated factors conferring diffuse and localised adherence of enteropathogenic *Escherichia coli. Infection and Immunity* 48: 378-383.

Nataro JP, Seriwatana J, Fasano A, Maneval DR, Guers LD, Noriega F, Dubovsky F, Levine MM and Morris JG Jr (1995). Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and Shigella strains. *Infection and Immunity* 63: 4721-4728.

Nataro JP, Steiner T and Guerrant RL (1998). Enteroaggregative *Escherichia coli* (EAEC) in diarrhoea and malnutrition. *Emerging Infectious Diseases* 4: 1-13.

Navarro-García F, Eslava C, Villaseca JM, López-Revilla R, Czeczulin JR, Srinivas S, Nataro JP and Cravioto A (1998). In vitro effects of a high-molecular-weight heat-labile enterotoxin from enteroaggregative *Escherichia coli. Infection and Immunity* 66: 3149-3154.

Neter E (1959). Enteritis due to enteropathogenic *Escherichia coli*: present-day status and unsolved problems. *Journal of Pediatrics* 55: 223-239.

Neter E, Westphal O, Luderitz R, Gino M and Gorzynski EA (1955). Demonstration of antibodies against enteropathogenic *Escherichia coli* in sera of children of various ages. *Paediatrics* 16: 801-807.

Norazaeh A, Rahizan I, Zainuldin T, Rohani MY and Kamel AG (1998). Enteropathogenic *Escherichia coli* in raw and cooked food. *Southeast Asian Journal of Tropical Medicine and Public Health* 29: 91-93.

Norris V, Baldwin TJ, Sweeney ST, Williams PH and Leach KL (1991). A protein kinase C-like activity in *Escherichia coli*. *Molecular Microbiology* 5: 2977-2981.

Norris V, Freestone P, Grant S, Toth I, Trinei M, Clarke SC, Nyström T and Onada T. Tyrosine phosphorylation in *E. coli* has implications for the manipulation of intra-cellular signalling and structure in anti-bacterial therapies. *Molecular Microbiology*, submitted.

Nyström T and Neidhardt FC (1996). Effects of overproducing the universal stress protein, UspA, in *Escherichia coli* K-12. *Journal of Bacteriology* 178: 927-930.

O'Gorman LE, Krejany EO, Bennett-Wood VR and Robins-Browne RM (1996). Transfer of attaching and effacing capacity from a strain of enteropathogenic *Escherichia coli* to *E. coli* K-12. *Microbiological Research* 151: 379-385.

Olukoya DK, Daini O, Alabi SA, Coker AD, Odugbemi T and Akinrimisi EO (1988). Antimicrobial resistance patterns and plasmids of enteropathogenic *Escherichia coli* isolated in Nigeria. *European Journal of Epidemiology* 4: 306-309.

Ostrovsky PC and Maloy S (1995). Protein phosphorylation on serine, threonine and tyrosine residues modulated membrane-protein interactions and transcriptional regulation in *Salmonella typhimurium*. *Genes and Development* 9: 2034-2041.

Paton JC and Paton AW (1998). Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. Clinical Microbiology Reviews 11: 45-479.

Pickett CL, Cottle DL, Pesci EC and Bikah G (1994). Cloning, sequencing and expression of the *Escherichia coli* cytolethal distending toxin genes. *Infection and Immunity* 62: 1046-1051.

Pickett CL, Pesci EC, Cottle DL, Russell G, Erdem AN and Zeytin H (1996). Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* species *cdt*B genes. *Infection and Immunity* 64: 2070-2078.

Pillien F, Chalareng C, Boury M, Tasca C, de Rycke J and Milon A (1996). Role of adhesive factor/rabbit 2 in experimental enteropathogenic *Escherichia coli* O103 diarrhea of weaned rabbit. *Veterinary Microbiology* 50: 105-115.

Plunkett G, Burland V, Daniels DL and Blattner FR (1993). Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. *Nucleic Acids Research* 21: 3391-3398.

Polotsky YE, Dragunskaya EM, Selverstova VG, Avdeeva TA, Chakhutinskay MG, Kétyi I, Verténvi A, Ralovich B, Emódy L, Málovics I, SafonovaWV, Snigirevshaya ES and Karyagina EI (1977). Pathogenic effect of enterotoxigenic *Escherichia coli* and *Escherichia coli* causing infantile diarrhoea. *Acta Microbiologie Acadamie Science Hungary* 24: 221-236.

Pospischil A, Mainil JG, Baljer G and Moon HW (1987). Attaching and effacing bacteria in the intestines of calves and cats with diarrhoea. *Veterinary Pathology* 24: 330-334.

Puente JL, Bieber D, Ramer SW and Schoolnik GK (1996). The bundle-forming pili of enteropathogenic *Escherichia coli*: transcriptional regulation by environmental signals. *Molecular Microbiology* 20: 87-100.

Pupo GM, Karadis DKR, Lan R and Reeves PR (1997). Evolutionary relationships among pathogenic and non-pathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infection and Immunity* 64: 2685-2692.

Qi SY, Li Y, Szyroki A, Giles IG, Moir A and O'Connor CD (1995). *Salmonella typhimurium* responses to a bactericidal protein from human neutrophils. *Molecular Microbiology* 17: 523-531.

Rahmsdorf HJ, Pai SH, Ponta H, Herrlich and Roskoski R. Protein kinase induction in *Escherichia coli* by bacteriophage T7. *Proceedings of the National Academy of Science USA* 71: 586-589.

Ramer SW, Bieber D and Schoolnik GK (1996). BfpB, an outer membrane lipoprotein required for the biogenesis of bundle-forming pili in enteropathogenic *Escherichia coli*. *Journal of Bacteriology* 178: 6555-6563.

Ray MK, Kumar GS and Shivaji S (1994). Tyrosine phosphorylation of a cytoplasmic protein from the antarctic psychrotrophic bacterium *Pseudomonas syringae*. *FEMS Microbiology Letters* 122: 49-54.

Reed LJ and Muench H (1938). A simple method for estimating fifty per cent endpoints. *The American Journal of Hygiene* 27: 493-497.

Ren L, Rahman MS and Humayun MZ (1999). *Escherichia coli* cells exposed to streptomycin display a mutator phenotype. *Journal of Bacteriology* 181: 1043-1044.

Reyrat J, Pelicic V, Gicquel B and Rappuoli R (1998). Counterselectable markers: untapped tools for bacterial genetics and pathogenesis. *Infection and Immunity* 66: 4011-4017.

Robins-Browne RM (1987). Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. *Reviews of Infectious Diseases* 9: 28-52.

Robins-Browne RM, Tokhi AM, Adams LM and Bennett-Wood V (1994). Host specificity of enteropathogenic *Escherichia coli* from rabbits: lack of correlation between adherence *in vitro* and pathogenicity for laboratory animals. *Infection and Immunity* 62: 3329-3336.

Robins-Browne RM, Yam WC, O'Gorman LE and Bettelheim KA (1993). Examination of archetypal strains of enteropathogenic *Escherichia coli* for properties associated with bacterial virulence. *Journal of Medical Microbiology* 38: 222-226.

Rosa AC, Mariano AT, Pereira AM, Tibana A, Gomes TA and Andrade JR (1998). Enteropathogenicity markers in *Escherichia coli* isolated from infants with acute diarrhoea and healthy controls in Rio de Janeiro, Brazil. *Journal of Medical Microbiology* 47: 781-790.

Rosenshine I, Donnenberg MS, Kaper JB and Finlay BB (1992). Signal transduction between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells: EPEC induces tyrosine phosphorylation of host cell proteins to initiate cytoskeletal rearrangement and bacterial uptake. *EMBO Journal* 11: 3551-3560.

Rosenshine I, Ruschkowski S and Finlay BB (1996a). Expression of attaching/effacing activity by enteropathogenic *Escherichia coli* depends on growth phase, temperature, and protein synthesis upon contact with epithelial cells. *Infection and Immunity* 64: 966-973.

Rosenshine I, Ruschkowski S, Stein M, Reinscheid DJ, Mills SD and Finlay BB (1996b). A pathogenic bacterium triggers epithelial signals to form a functional bacterial receptor that mediates actin pseudopod formation. *EMBO Journal* 15: 2613-2624.

Rothbaum R, McAdams AJ, Gianella R and Partin JC (1982). A clinicopathologic study of enterocyte-adherent *Escherichia coli*: a cause of protracted diarrhoea in infants. *Gastroenterology* 83: 441-454.

Ruczkowska J, Gosciniak G and Dolna I (1990). Antibiotic sensitivity of enteropathogenic strains of *Escherichia coli* isolated from children with diarrhea. *Przegl Epidemiol* 44: 199-201.

Rudd KE (1998). Linkage map of *Escherichia coli* K-12, edition 10: the physical map. *Microbiology and Molecular Biology Reviews* 62: 985-1019.

Russel M (1998). Macromolecular assembly and secretion across the bacterial cell envelope: type II protein secretion systems. *Journal of Molecular Biology* 279: 485-499.

Sambrook J, Fritsch EF and Maniatis T (1989). Molecular Cloning - a laboratory manual. Second edition. Cold Spring Harbour Laboratory Press.

Sanger F, Nicklen S and Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science USA* 74: 5463-5467.

Sanger JM, Ashton F, Zhukarev V, Chang R and Kaper JB (1995). Enteropathogenic *Escherichia coli* move on the surfaces of infected cultured cells. *Molecular Biology of the Cell* 66: abstract 665.

Scaletsky ICA, Gatti MSV, da Silveira JF, DeLuca IMS, Freymuller E and Travassos LR (1995). Plasmid encoding for drug resistance and invasion of epithelial cells in enteropathogenic *Escherichia coli* O11:H⁻. *Microbial Pathogenesis* 18: 387-399.

Scaletsky ICA, Pedroso MZ and Fagundes-Neto U (1996). Attaching and effacing enteropathogenic *Escherichia coli* O18ab invades epithelial cells and causes persistent diarrhea. *Infection and Immunity* 64: 4876-4881.

Schauer DB and Falkow S (1993). The *eae* gene of *Citrobacter freundii* biotype 4280 is necessary for colonization in transmissible murine colonic hyperplasia. *Infection and Immunity* 61: 4654-4661.

Schmidt H, Henkel B and Karch H (1997). A gene cluster closely related to type II secretion pathway operons of Gram-negative bacteria is located on the large plasmid of enterohemorrhagic *Escherichia coli* O157 strains. *FEMS Microbiology Letters* 148: 265-272.

Scotland S (1983). In: Edelman and Levine. *Journal of Infectious Diseases* 147: 1108-1118.

Schweizer HP (1992). Allelic exchange in *Pseudomonas aeruginosa* using novel ColE1-type vectors and a family of cassettes containing a portable *ori*T and the counter selectable *Bacillus subtilis sacB* marker. *Molecular Microbiology* 6: 1195-1204.

Segal ED, Falkow S and Tompkins LS (1996). *Helicobacter pylori* attachment to gastric cells indices cytoskeletal rearrangements and tyrosine phosphorylation of host cell proteins. *Proceedings of the National Academy of Science USA* 93: 1259-1264.

Selbitschka W, Niemann S and Puhler A (1993). Construction of gene replacement vectors for Gram- bacteria using a genetically modified *sacRB* gene as a positive selection marker. *Applied Microbiology and Biotechnology* 38: 615-618.

Senerwa D, Mutanda LN, Gathuma JM and Olsvik O (1991). Antimicrobial resistance of enteropathogenic *Escherichia coli* strains from a nosocomial outbreak in Kenya. *APMIS* 99: 728-734.

Shepherd NS and Smoller D (1994). The P1 vector system for the preparation and screening of genomic libraries. *Genetic Engineering (New York)* 16:213-228.

Simon R, Priefer U and Pühler A (1983). A broad host range mobilzation system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. *Biotechnology* November: 784-791.

Smith SC, Kennelly PJ and Potts M (1997). Protein-tyrosine phosphorylation in the Archaea. *Journal of Bacteriology* 1997; 179: 2418-20.

Sohel I, Puente JL, Ramer, Bieber D, Wu C and Schoolnik GK (1996). Enteropathogenic *Escherichia coli*: identification of a gene cluster encoding for bundle-forming pilus morphogenesis. *Journal of Bacteriology* 178: 2613-2628.

Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98: 503-511.

Stein M (1996). Characterisation of EspC, a 110kDa protein secreted by enteropathogenic *E. coli*. Seventh biennial UA-UC International Conference on Infectious Diseases 16A (abstract).

Stibitz S (1994). Use of conditionally counterselectable suicide vectors for allelic exchange. *Methods in Enzymology* 235: 458-464.

Stone KD, Zhang H, Carlson LK and Donnenberg MS (1996). A cluster of fourteen genes from enteropathogenic *Escherichia coli* is sufficient for the biogenesis of a type IV pilus. *Molecular Microbiology* 20: 325-337.

Strauss EJ and Falkow S (1997). Microbial pathogenesis: genomics and beyond. *Science* 276: 707-712.

Taylor J (1970). Infectious infantile enteritis, yesterday and today. *Proceedings of the Royal Society of Medicine* 63:1297-1301.

Taylor J, Hart A, Batt RM, McDougall C and McLean L (1986). Ultrastructural and biochemical changes in human jejunal mucosa associated with enteropathogenic *Escherichia coli* (O111) infection. *Journal of Pediatric Gastroenterology and Nutrition* 5(1): 70-73.

Taylor KA, Luther PW and Donnenberg MS (1999). Expression of the EspB protein of enteropathogenic *Escherichia coli* within HeLa cells affects stress fibers and cellular morphology. *Infection and Immunity* 67(1): 120-125.

Taylor KA, O'Connell CB, Luther PW and Donnenberg MS (1998). The EspB protein of enteropathogenic *Escherichia coli* is targeted to the cytoplasm of infected HeLa cells. *Infection and Immunity* 66: 5501-5507.

Taylor RK, Manoil C and Mekalanos JJ (1989). Broad-host-range vectors for delivery of Tn*phoA*: use in genetic analyisis of secreted virulence determinants of *Vibrio cholerae*. *Journal of Bacteriology* 171(4): 1870-1878.

Thomson S (1956). Is infantile gastro-enteritis fundamentally a milk-borne infection? *Journal of Hygiene* 54: 311-314.

Thoren A (1983). The role of enteropathogenic *Escherichia coli* in infantile diarrhoea. *Scandinavian Journal of Infectious Diseases* Supplement 37: 1-51.

Tizard R, Cate RL, Ramachandran KL, Wysek M, Voyta JC, Murphy OJ and Bronstein I (1990). Imaging of DNA sequences with chemiluminescence. *Proceedings of the National Academy of Science USA* 87: 4514-4518.

Tobe T, Schoolnik GK, Sohel I, Bustamante VH and Puente JL (1996). Cloning and characterisation of *bfpTVW*, genes required for the transcriptional activation of *bfpA* in enteropathogenic *Escherichia coli*. *Molecular Microbiology* 21: 963-975.

Toolan HW (1954).Transplantable human neoplasms maintained in cortisone-treated laboratory animals: H.S.#1; H.Ep.#2; H.Ep.#3; and H.Emb.Rh.#1. *Cancer Research* 14: 660-666.

Von Moll LK and Cantey JR (1997). Peyer's patch adherence of enteropathogenic *Escherichia coli* strains in rabbits. *Infection and Immunity* 65: 3788-3793.

Vuopio-Varkila J and Schoolnik GK (1991). Localised adherence by enteropathogenic *Escherichia coli* is an inducible phenotype associated with the expression of new outer membrane proteins. *Journal of Experimental Medicine* 174: 1167-1177.

Wainwright LA and Kaper JB (1998). EspB and EspD require a specific chaperone for proper secretion from enteropathogenic *Escherichia coli*. *Molecular Microbiology* 27: 1247-1260.

Wall PG, McDonnell RJ, Adak GK, Cheasty T, Smith HR and Rowe B (1996). General outbreaks of Vero cytotoxin producing *Escherichia coli* O157 in England and Wales from 1992 to 1994. *Communicable Disease Review* 6: R26-R33.

Wassenaar TM, Geilhausen B and Newell DG (1998). Evidence of genomic instability in *Campylobacter jejuni* isolated from poultry. *Applied and Environmental Microbiology* 64: 1816-1821.

Whittam TS, Wolfe ML, Wacksmith IK, Orskov F, Orskov I and Wilson RA (1993). Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhoea. *Infection and Immunity* 61: 1619-1629.

Wieler LH, McDaniel TK, Whittam TS and Kaper JB (1997). Insertion site of enterocyte effacement in enteropathogenic and enterohaemorrhagic *Escherichia coli* differs in relation to the clonal phylogeny of the strains. *FEMS Microbiology Letters* 156: 49-53.

Wieler LH, Schwanitz A, Vieler E, Busse B, Steinrück H, Kaper JB and Baljer G (1998). Virulence properties of Shiga toxin-producing *Escherichia coli* O118, a major group of STEC pathogens in calves. *Journal of Clinical Microbiology* 36: 1604-1607.

Wolf MK, Abdrews GP, Fritz DL, Sjogren RW and Boedeker EC (1988). Characterization of the plasmid from *Escherichia coli* RDEC-1 that mediates expression of adhesin AF/R1 and evidence that AF/R1 pili promote but are not essential for enteropathogenic disease. *Infection and Immunity* 56: 1846-1857.

Wolff C, Nisan I, Hanski E, Frankel G and Rosenshine I (1998). Protein translocation into host epithelial cells by infecting enteropathogenic *Escherichia coli*. *Molecular Microbiology* 28: 143-155.

Wong KK, McClelland M, Stillwell LC, Sisk EC, Thurston SJ and Saffer JD (1998). Identification and sequence analysis of a 27-kilobase chromosomal fragment containing a *Salmonella* pathogenicity island located at 92 minutes on the chromosome map of *Salmonella enterica* serovar Typhimurium LT2. *Infection and Immunity* 66: 3365-3371.

Wood MW, Jones MA, Watson PR, Hedges S, Wallis TS and Galyov EE (1998). Identification of a pathogenicity island required for *Salmonella* entero-pathogenicity. *Molecular Microbiology* 29: 883-891.

World Health Organisation (1987). Programme for control of diarrhoeal diseases. In: Manual for laboratory investigations of acute enteric infection, page 27. World Health Organisation, Geneva.

World Health Organisation (1992). WHO/PAHO Informal Consultation on Intestinal Protozoal Infections, Mexico 1991.

World Health Organisation (1994). Programme for control of diarrhoeal diseases: Ninth programme report 1992-1993. World Health Organisation, Geneva.

World Health Organisation (1998). World Health Report 1998. Life in the 21st century - a vision for all. World Health Organisation, Geneva.

Wren BW, Henderson J and Ketley JM (1994). A PCR based strategy for the rapid construction of defined bacterial mutants. *Biotechniques* 16:7-8.

Wu LF, Reizer A, Reizer J, Cai B, Tomich JM and Saier MH (1991). Nucleotide sequence of the *Rhodobacter capsulatus* FruK gene, which encodes fructose-1-phosphate kinase - evidence for a kinase superfamily including both phosphofructokinases of *Escherichia coli*. *Journal of Bacteriology* 173: 3117-3127.

Yamanaka K, Fang L and Inouye M (1998). The CspA family in *Escherichia coli*: multiple gene duplication for stress adaptation. *Molecular Microbiology* 27: 247-255.

Yannisch-Peron C, Vieira J and Messing J (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103-119.

Yuhan R, Koutsouris A, Savkovic SD and Hecht G (1997). Enteropathogenic *Escherichia coli*-induced myosin light chain phosphorylation alters intestinal epithelial permeability. *Gastroenterology* 113: 1873-1882.

Zhang HZ and Donnenberg MS (1996). DsbA is required for stability of the type IV pilin of enteropathogenic *Escherichia coli*. *Molecular Microbiology* 21: 787-797.

Zhang H, Lory S and Donnenberg MS (1994). A plasmid-encoded prepilin peptidase gene from enteropathogenic *Escherichia coli*. *Journal of Bacteriology* 176: 6885-6891.

Zhao S, Mitchell SE, Meng J, Doyle MP and Kresovich S (1995). Cloning and nucleotide sequence of a gene upstream of the *eae*A gene of enterohaemorrhagic *Escherichia coli* O157:H7. *FEMS Microbiology Letters* 133: 35-39.

Zhu C, Harel J, Dumas F and Fairbrother JM (1995). Identification of EaeA protein in the outer membrane of attaching and effacing *Escherichia coli* O45 from pigs. *FEMS Microbiology Letters* 129: 237-242.

Zhu C, Ménard S, Dubreuil JD and Fairbrother JM (1996). Detection and localisation of the EaeA protein of attaching and effacing *Escherichia coli* O45 from pigs using a monoclonal antibody. Microbial Pathogenesis 21: 205-213.