Analysis of the *vlpA* genes and their roles in the pathogenesis of *Vibrio cholerae* O139.

Pranom Kaewrakon, M.Sc.

Microbial Pathogenesis Unit

Discipline of Microbiology and Immunology

Department of Molecular Biosciences

The University of Adelaide,

Adelaide, SA 5005

Australia

A Thesis Submitted for the Degree of Doctor of Philosophy,

December, 2000.

I dedicate this thesis to my father **Insorn Jitkum**,

and to the memory of my mother **Khummoon Jitkum**,

and dearest sister **Tassanee Jitkum**.

Acknowledgments

Firstly I am extremely grateful and deeply indebted to my supervisors, Dr Tony Focareta and Professor Paul Manning for their excellent supervision and constant encouragement, and also to the Thai Government for financial support. I have immensely appreciated the generosity of Dr Sue Williams who has helped in numerous ways with endless patience throughout this study. I acknowledge and thank Dr Stephen Attridge for handling mouse work.

My sincerest thanks go to all the members of the Microbial Pathogenesis Unit, past and present, for their support, friendship, and laughter during my time "in the lab". The past five years of research have been a fulfilling and enjoyable experience for me. I would like to thank Chris Vindurampulle, Chris Cusaro, Chris Clark, Luisa van den Bosch, James Botten, Angello Fallarino, Leanne Purins, Kate Dredge, Sally McCloud, Tom Duthy, Elizabeth Parker, Dr Renato Morona, Dr Michael Bawden, Dr Elena Voss, Dr Monica Ogierman, Dr Maria Mavris, Dr Craig Daniels, Dr Kathy Daniels, and Dr Uwe Stroeher.

I would also like to thank Dr Ruth Hall and Dr Christina Collis for useful discussion and helpful comments.

Thankyou to Athirad Jantharawong for being a great study partner and for your computer expertise.

Finally, I thank my family for their continuing love, and I especially thank my husband Wichian, and my son Wachiranunt for the concern and understanding they have shown me during this study.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Pranom Kaewrakon

Abbreviations

59-be	59-base element
A_{600}	absorbance at 600 nm
A	adenine
Ace	accessory cholera enterotoxin
ACF	accessory colonization factor
Ap	ampicillin
ATP	adenosine 5′-triphosphate
attI	attachment site
BIME	bacterial interspersed mosaic element
bp	base pair
BSA	bovine serum albumin
C	cytosine
Cm	chloramphenicol
Ctx	cholera-toxin
CTXϕ	cholera-toxin phage
DIG	digoxigenin
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
ddNTP	dideoxyribonucleoside triphosphate
DTT	dithiothreitol
ECL	enhanced chemiluminescence
E. coli	*Escherichia coli*
EDTA	ethylene-diamine-tetra-acetic-acid, disodium salt
ERIC	enterobacterial repetitive intergenic consensus sequence
EtBr	ethidium bromide
G	guanine
GM_1	Galactosyl-N-acetyl-galactosaminyl-Sialosyl-Lactosylceramide
GTP	guanosine 5′-triphosphate
HA	haemagglutinin

Hly	haemolysis
IM	inner membrane
IntI	integrase
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
Km	kanamycin
LB	Luria broth
LD_{50}	dose of organisms capable of killing 50% of infant mice within 48 hours
LPS	lipopolysaccharide
mg	milligram
MFRHA	mannose-fucose-resistant haemagglutinin
ml	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
MSHA	D-mannose-sensitive haemagglutinin
NA	Nutrient agar
NANase	neuraminidase
NB	Nutrient broth
nt	nucleotide
OD	optical density
OM	outer membrane
OMP	outer membrane protein
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol-6000
pmol	picomoles
Pmx	polymyxin B
PU	palindromic unit sequences
RBS	ribosome binding site
REP	repetitive extragenic palindromic sequence

Rif	rifampicin
RNA	ribonucleic acid
RNase	ribonuclease
RT-PCR	reverse transcriptase polymerase chain reaction
RT	room temperature
^{R}	resistant
^{S}	sensitive
sc	subcutaneous
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-poly acrylamide gel electrophoresis
T	thymine
TCP	toxin-coregulated pilus
TEMED	N,N,N′,N′-tetramethyl-ethylene-diamine
Tet	tetracycline
Tn	transposon
Tris	Tris-(hydroxymethyl)-aminomethane
TTBS	Tris-buffered saline with Tween-20 added
U	uracil
UV	ultraviolet
µF	microFarad
µg	microgram
µl	microlitre
V. cholerae	*Vibrio cholerae*
VCO	*V. cholerae* open reading frame
VCR	*Vibrio cholerae* repetitive DNA sequence
vlpA	*Vibrio cholerae* lipoprotein A
VPI	*Vibrio cholerae* pathogenicity island
VPIϕ	*Vibrio cholerae* pathogenicity island phage
v/v	volume per volume
WC	whole cell
WM	whole membrane
w/v	weight per volume

X-gal	5-bromo-4-chloro-3-indolyl-β-galactopyranoside
X-pho	5-bromo-4-chloro-3-indolyl phosphate
Zot	zonula occludens toxin

Abstract

The *vlpA* (*Vibrio cholerae* lipoprotein A) gene encodes the first identified bacterial member of the α_2-microglobulin (lipocalin) superfamily, which function as small hydrophobic molecule transporters. This gene is present in *V. anguillarum*, *V. cholerae* O1, and some strains of non-O1 serotypes in multiple copies and flanked by a repeat sequence of approximately 124-bp, termed VCR for *V. cholerae* repeat sequence. The VCR element has all the structural features associated with the 59-base element (59-be) (or *attC* site) required for site-specific recombination mediated by integron encoded integrases. Over 150 copies of this VCR element are present in the *V. cholerae* genome, interspersed with open reading frames or gene cassettes. This constitutes the *V. cholerae* mega-integron, a chromosomally located integron of approximately 120-kb. Both integrases IntI1 from *Pseudomanas aeruginosa* and IntI4 from *V. cholerae* are capable of mediating deletion of gene cassettes containing VCRs in plasmid clones in *Escherichia coli*, and in *V. cholerae* chromosomally located *vlpA* genes. The deletion event occurs at the VCR sites in the presence of an over-expressed clone of both IntI1 and IntI4. Interestingly, IntI4 specificity is not limited to the VCR elements of *V. cholerae*, but also recognizes the *attI* site of the *P. aeruginosa* integron. However, the deletions require the presence of the *attI*1 site in addition to the VCR elements for IntI1 mediated excision. This is also the case for IntI4 which requires the *attI*4 site as well as VCRs for activity. Site-specific recombination catalyzed by IntI1 occurs at both VCR and *attI*1 sites, whereas IntI4 was shown to capable of mediating deletion only at VCR sites. This is the first time IntI4 has been shown to function as an integrase mediated the deletion of gene cassettes in the *V. cholerae* chromosome.

A mutant lacking VlpA activity has been constructed in *V. cholerae* O139 which contains two copies of *vlpA*, by allelic exchange after inactivation via the insertion of kanamycin and tetracycline resistance gene cartridges. The VlpA-negative mutant has been assessed for virulence in the infant mouse cholera model. This mutant does not show a marked defect in its ability to persist in the infant mouse gut and competes equally well with the wild-type organism. The function of VlpA is still unknown, although the eukaryotic lipocalins are best known for their binding of small hydrophobic ligands.

Different promoters from plasmid vectors were used in *E. coli* and *V. cholerae* to express *vlpA*, but a gene product was detected in *E. coli* using the T7 phage expression system, with a protein size of ~19 kDa. No specific products with Western blotting analysis were present to show that VlpA was expressed in *V. cholerae*, even in iron limiting conditions. Northern blot analysis, primer extension, and RT-PCR were utilized to determine if transcription of *vlpA* occurred, but its expression from the *V. cholerae* chromosome is still unclear.

Table of Contents

Chapter 1

Introduction

1.1 Introduction

Cholera is a serious epidemic disease, responsible for the death of millions of people, and continues to be a major health problem worldwide. It was one of the first epidemic diseases to be controlled by public health measures in the nineteenth century (Schoenberg, 1974). Under the terms of the International Health Regulation of 1969, cholera is one of three diseases (small pox, yellow fever, cholera) for which it is mandatory to notify the World Health Organization (WHO) (World Health Organization, 1993).

Cholera has swept the world in seven major pandemics, the most recent in 1992, and now effects at least 98 countries (Kaysner and Hill, 1994). Cases of cholera have been reported in Europe, Asia, Africa, Australia, New Zealand, South and Central America, and the United States, and typically, it has been able to cross the globe (Barua, 1992). Cholera appears in many areas of the world even where it is not currently a health problem. Even though modern sanitation practices in developed countries have virtually eliminated epidemic cholera from public health concerns, those countries still contain areas of endemic toxigenic strains, for example, the United States Gulf Coast (Kaysner and Hill, 1994).

1.2 General background

The disease cholera is caused by *Vibrio cholerae*, a Gram-negative bacteria in the genus *Vibrio*, which is the type genus of the family *Vibrionaceae* (Rabbani, 1986; Baumann *et al.*, 1984; Finkelstein, 1988; Barua, 1992). The *Vibrio* group contains four genera, *Vibrio*, *Aeromonas*, *Photobacterium*, and *Plesiomonas* (Baumann & Schubert,

1984). Most of these organisms are facultative, anaerobic, oxidase positive, asporogenous, motile, straight or curved rods which possess a fermentative metabolism (Baumann & Schubert, 1984; Kay *et al.*, 1994). Although some members of the *Vibrio* group are peritrichously flagellated, most of them have a single polar flagellum. Most vibrios and related bacteria are aquatic, found either in fresh water or marine habitats. In addition, most of the cholera cases are also due to water contaminated with human faeces, and are spread within households in undeveloped countries by the faecal-oral route. Cholera vibrios attach firmly to the small intestine epithelium, where they multiply and release cholera toxin (Ctx). Ctx causes the loss of fluids and salts into the stool, of approximately 20 litres per day, catalysing a life-threatening diarrhoea resulting in dehydration, leading to shock, organ failure and death. The mortality rate can be as high as 60 percent if untreated.

1.3 History

The history of six cholera pandemics in the period 1817-1923 involving the seventh pandemic was described by Pollitzer (1959) and Barua (1992). The intervals between the individual pandemics are not always clear, with that between the second and the third being particularly obscure. The first pandemic spanned the period from 1817 to 1823, initiating in India and spreading through many surrounding countries in Asia (MacNamara, 1876). The second pandemic was believed to have occurred from 1829 to 1851 which reached Moscow by the autumn of 1830 and two years later arrived in Great Britain, also spreading to many other countries in Europe (Barua, 1992). Cholera existed in epidemic proportions in Arabia from 1828 to 1831, and rapidly spread among the pilgrims assembled in Mecca. In 1832 cholera appeared on the west coast of Australia and in the same year it was reported that this disease crossed the Atlantic to America spreading to Canada. In 1849

cholera was described as "America's greatest scourge" after its ravages in New York, and in this year the disease also spread to North Africa (Barua, 1992).

International cooperation in health began during the third pandemic in 1852-1859. Cholera was rampant in 1853-1854 in many countries in Europe, and also in the United States. When cholera occurred in Tuscany Italy in 1854, an Italian physician Filippo Pacini, observed large numbers of curved bacteria in the intestinal contents of some cadavers of cholera victims and called these bacteria *Vibrio cholera* (Sakazaki, 1992), but the aetiological relationship was not convincingly demonstrated. Cholera was recorded in various parts of Central America from 1856 to 1857, with more serious outbreaks occurring in Asia in 1857-1859.

The Mecca pilgrimage in 1865 was the scene of a major epidemic during the fourth pandemic in 1863-1879, resulting in approximately 30,000 deaths amongst the 90,000 pilgrims. The worst epidemics were recorded in 1866 when cholera raged from Caucasus to Russia and Europe with approximately 385,000 people perishing in many countries (Barua, 1992). Cholera ravaged Africa in 1865 and it had severe and protracted outbreaks in 1869 where 70,000 people were reported to have died in Zanzibar. The fifth pandemic was widespread in Europe in 1881-1896. During this time, this disease was serious in India in 1881 and amongst the Mecca pilgrims in 1881-1882. Importantly, by 1883 the disease was shown by Robert Koch to be caused by a bacterium. During 1890-1895, many countries in Asia were affected with a total of about 340,000 cases (Barua, 1992).

Although cholera had disappeared from the Americas and most of Africa and Europe by the turn of the century (Blake, 1994), it re-appeared during the sixth pandemic in 1899-1923. A severe outbreak cholera occurred again in India in 1899. In 1905 a new biotype, the El Tor vibrio, was identified and first described in Indonesia, and was termed "Paracholera" by de Moor (de Moor, 1949). In 1961 the El Tor vibrio spread out of

Indonesia to other countries in Asia, beginning the seventh cholera pandemic. The seventh pandemic arose approximately 36 years after the sixth pandemic subsided in 1925, occurring between early 1961 and late 1993, possibly still continuing today (Lan and Reeves, 1998). The cases of cholera were mostly limited to areas in Asia and the Indian subcontinent where cholera is endemic (Swerdlow & Isaäcson, 1994). About nine years after the beginning of this seventh pandemic, cholera invaded Africa and Europe, and dominated the world's public health problems. An unprecedented event occurred in the history of cholera in 1992 with the emergence of a novel causative serogroup classified as *V. cholerae* O139 Bengal (Bhattacharya *et al.*, 1993; Sharma *et al.*, 1997). This cholera serogroup was able to rapidly spread through many countries in Asia.

1.4 Classification

V. cholerae is classified into O1 and non-O1 serogroups (serovars) based on their ability to cause cholera epidemics (Bik *et al.*, 1995). Currently, more than 150 serogroups of *V. cholerae* have been identified. However, the aetiological agents of the first four pandemics are not known, although the last three cholera pandemics since 1881 were caused by *V. cholerae* serogroup O1 (Barua, 1992; Albert, 1994; Kay *et al.*, 1994). *V. cholerae* non-O1 serogroups are not known to cause epidemics of diarrhoea, although they have been known to cause sporadic cases and small outbreaks of diarrhoea and extra-intestinal infection (Albert, 1994). It was therefore assumed that only the *V. cholerae* O1 serogroup had epidemic potential (Bik *et al.*, 1995). An exception to this however, was the most recent cholera epidemic which started in 1992 in the region around the Bay of Bengal in India and quickly spread to other parts of Asia (Albert *et al.*, 1993; Bhattacharya *et al.*, 1993; Ramamurthy *et al.*, 1993). This new serogroup has been classified as *V. cholerae*

O139 Bengal and has been implicated as the responsible pandemic strain of cholera (Nair *et al.*, 1994).

Epidemic strains of *V. cholerae* O1 can be classified into two major biotypes, classical and El Tor, on the basis of several phenotypic characteristics including biochemical differences, for example susceptibility to polymyxin B and hemagglutination of chicken erythrocytes (Kaper *et al.*, 1995; DiRita *et al.*, 1996). In one study by Kaper *et al.* (1995), it was estimated that an infection with classical strains resulted in severe disease in 11% of cases, whereas only 2% of infections with El Tor strains resulted in severe disease. However, since the El Tor vibrio was first described in Indonesia in 1937, much of the world has been currently experiencing an El Tor pandemic, unlike the fifth and the sixth pandemics that were due to classical biotypes (Blake, 1994).

On the basis of the heat-stable somatic- or O-antigen, both classical and El Tor biotypes can be further subdivided into two major serotypes designated Ogawa and Inaba (Finkelstein, 1975; Burrows *et al.*, 1946; Sakazaki and Tamura, 1971; Redmond *et al.*,1973). No serotypes have been described for the O139 serogroup (Kay *et al.*, 1994). The predominant serological reactivity of the two serotypes, determined by agglutination and agglutinin absorption test is generally accepted to be explained simply on the basis of three antigens A, B, and C (Burrows *et al.*, 1946; Finkelstein, 1975). The strains of the Ogawa serotype express the A and B antigens and a small amount of C antigen whereas Inaba strains express only the A and C antigens (Sakazaki and Tamura, 1971; Redmond *et al.*, 1973). A third serotype, Hikojima, an intermediate between the Inaba and Ogawa serotypes, is rare and unstable. It possesses all three antigens A, B, and C. In addition, Hikojima has been suggested to represent strains that undergo conversion at an elevated frequency (Sakazaki and Tamura, 1971). *V. cholerae* O1 strains have been demonstrated to interconvert between the Ogawa and Inaba forms (Bhaskaran and Gorrill, 1957; Sheehy *et*

al., 1966; Gangarosa *et al.*, 1967). However, the frequency of conversion of Ogawa to Inaba is approximately 10^{-5} (Bhaskaran and Gorrill, 1957), whereas the conversion from Inaba to Ogawa appears to be less frequent and is strain dependent (Stroeher *et al.*, 1992). Moreover, it has been proposed that seroconversion from Inaba to Ogawa correlates well with the host immune response *in vivo*. Although the identification of the serotype-specific genetic determinant has implications for vaccine development, it has been shown that the serotype specificity has no effect on virulence (Stroeher *et al.*, 1992).

1.5 Transmission of cholera

Cholera is exclusively a human disease and no animal species has been found to be consistently infected (Beneson, 1991). It may be difficult in some outbreaks of cholera to determine whether food or water was the vehicle of transmission (Kaysner and Hill, 1994). Waterborne transmission plays an important role and can involve a wide variety of food and water vehicles. The primary source of infection in cholera epidemics is faeces from persons acutely infected with *V. cholerae* O1 (Mintz *et al.*, 1994) where large numbers of *V. cholerae* are discharged (Kaysner and Hill, 1994). The El Tor biotype is more viable in the environment than the classical biotype and can cause more asymptomatic infections. This biotype multiplies more rapidly and appears better adapted to foodborne transmission, since it survives in many foods longer than the classical biotype (Bik *et al.*, 1995).

Vehicles of transmission for *V. cholerae* O1 are the foods that are nearly neutral in pH (Roberts, 1992). Frequently transmission is via seafood, especially fish and shellfish as they are consumed raw in many parts of the world. *V. cholerae* is also transmitted via other foods under appropriate conditions. Grains and legumes support the growth of *V. cholerae* O1 quite well and have been repeatedly implicated in cholera transmission (Kolvin and Roberts, 1982). Fruits and vegetables also support the survival of *V. cholerae* O1 for a

number of days, but its survival is greatly decreased in a dry environment or an environment with a pH of less than 4.5 such as is seen in citrus fruits (Mintz *et al.*, 1994). Produce that grows close to the ground and is eaten raw, could easily be contaminated by irrigation water and might subsequently transmit cholera. Furthermore, there is evidence that cholera transmission can occur via meat or frozen food (Swaddiwudhipong *et al.*, 1990). Direct person-to-person transmission of cholera is not expected because the infectious dose is high. The contacts of cholera patients in the household do not have a significantly increased risk of infection with *V. cholera* O1 (Mintz *et al.*, 1994). Spread of cholera from one person to another by contact has never been demonstrated by rigorous scientific studies.

Control of cholera depends primarily on satisfactory sanitation measures, particularly in the treatment of sewage and the purification of drinking water. The efforts to protect, monitor, treat and prevent cholera require multiple strategies, including public education, adequate public health infrastructure as well as investigations into the ecology of the organisms and epidemiology of the disease. One of the more recent axioms that can be applied to the prevention of the spread of cholera by food and water is the "boil it, cook it, peel it, and forget it" philosophy (Kozicki *et al.*, 1985). Education is a critical factor for the prevention of cholera (Kaysner and Hill, 1994). This begins with encouragement of better hygiene in the kitchen. Investigation of the ecology of *V. cholerae* may help determine the natural reservoir of the organism. Sewage treatment would also greatly reduce the risk and spread of cholera. Many countries now provide safe drinking water and adequate sewage disposal. Drinking water is treated, usually by boiling or chlorination. The risk of outbreaks of cholera is greater, where the disease is endemic if large groups of people are crowded together without adequate sanitary facilities and adequate food-handling practices (Kaysner and Hill, 1994).

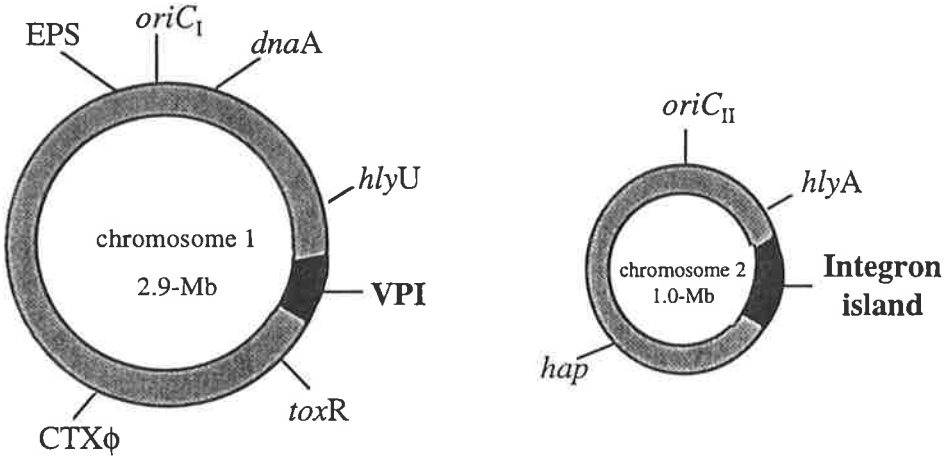
1.6 *Vibrio cholerae* pathogenicity island (VPI)

The *V. cholerae* genome contains two unique circular chromosomes (Trucksis *et al.*, 1998; Heidelberg *et al.*, 2000). The smaller chromosome is approximately 1.0-Mb in size and contains a pathogenicity island or locus with an integrated filamentous bacteriophage known as cholera-toxin phage (CTXϕ) which encodes the cholera toxin (Ctx) (Waldor and Mekalanos, 1996), and an integron-like gene capture system (Rowe-Magnus *et al.*, 1999; Heidelberg *et al.*, 2000; Waldor and RayChaudhuri, 2000). The larger chromosome is approximately 2.9-Mb in size containing the VPI encodeing a toxin-coregulated pilus (TCP) and represents another integrated filamentous bacteriophage. Remarkably, TCP functions both as a colonization factor and as a CTXϕ receptor, with the subunits of this pilus being identified. A second copy of CTXϕ is also present in *V. cholerae* classical biotype as the coat protein of the bacteriophage (Figure 1.1) (Trucksis *et al.*, 1998; Heidelberg *et al.*, 2000).

Karaolis *et al.* (1999) first identified this VPI as the genome of a filamentous bacteriophage, and termed it VPIϕ (for *V. cholerae* pathogenicity island phage). Virulent and epidemic strains of *V. cholerae* require both these two genetic elements to cause disease, VPIϕ and CTXϕ. The VPIϕ interacts with recipient *V. cholerae* cells, resulting in infection and insertion of VPIϕ DNA into the bacterial chromosome. This allows production of the TCP which permits infection by the CTXϕ. Such a *V. cholerae* strain now has a full complement of virulence genes. All strains of *V. cholerae* that can cause epidemic cholera contain this VPIϕ, but it is absent in non-toxigenic strains (Kovach *et al.*, 1996; Karaolis *et al.*, 1998, 1999; Taylor, 1999).

Figure 1.1 Distribution on the *V. cholerae* O1 El Tor chromosomes of key genetic loci required for virulence. The DNA sequences acquired from other sources, such as VPI, CTXϕ, and the integron island are shown (adapted from Waldor and RayChaudhuri, 2000).

Figure 1.1


EPS
oriC$_I$
dnaA
hlyU
chromosome 1
2.9-Mb
VPI
toxR
CTXϕ
oriC$_{II}$
hlyA
chromosome 2
1.0-Mb
Integron island
hap

1.7 Virulence factors in *V. cholerae*

Virulence factors can be loosely classified into two categories : those that promote bacterial colonization and those that cause damage to the host. The pathogenicity of toxigenic *V. cholerae* depends on a combination of virulence properties, including colonization factors, production of Ctx and the presence of associated outer membrane proteins (Rodrigue *et al.*, 1994).

1.7.1 Colonization factors

Successful colonization of the small intestine by *V. cholerae* depends on the adhesion of vibrios to the intestinal mucosal surface. This process appears to be complex, involving factors required for penetration and attachment, a co-ordinated expression of chemotactic and motility functions, adherence, and colonization modifiers.

1.7.1.1 Penetration and attachment

In order to be pathogenic, *V. cholerae* requires the expression of a number of coordinately regulated virulence factors to overcome the host's natural defense mechanisms (Kovach *et al.*, 1996). Prior to intestinal colonization the vibrios must survive exposure to the gastric acid of the stomach, localize in the small bowel, and penetrate the mucous gel protecting the intestinal microvilli. The mucous coat on the epithelium contributes to viscosity and reduces the ability of bacteria to penetrate the natural defense system. However, *V. cholerae* undergoes many changes as it adapts itself to the human body. Jenkin and Rowley (1959) revealed that the mucous layer of glycoprotein had to be penetrated before vibrios could be virulent. Burnet and Stone (1947) demonstrated that cholera vibrio produce neuraminidase (NANase) and mucinase to help penetration. Following penetration of the protective mucous gel, the vibrios colonize the epithelial cell

surface of the microvilli (Kovach *et al.*, 1996). At some point in the colonization process, the synthesis of Ctx, TCP and ACF (accessory colonization factor) occurs (Taylor *et al.*, 1987; Mekalanos, 1988).

1.7.1.2 Motility and chemotaxis

Most studies agree that motility and the associated phenomenon of chemotaxis are important for the virulence of *V. cholerae* enabling it to reach the enterocyte surface, but to date no mechanism has been defined (Richardson, 1994). *V. cholerae* has a single polar flagellum which is sheathed in what appears to be an extension of the outer membrane lipid bilayer, containing LPS and specific proteins. Evidence that flagellar motility may help *V. cholerae* reach the intestinal mucosa has been demonstrated by Guentzel and Berry (1975), and by Attridge and Rowley (1983). Non-motile mutants are less virulent than wildtypes, although both are still capable of inducing diarrhoea and can adhere to the intestinal epithelium. Although little is known about chemotaxis or the signals that guide *V. cholerae* to the mucosa, it is highly likely that there will be a close relationship between the chemotactic response to various molecules and gene regulation (Manning, 1994). The presence of the flagellum is essential for this response.

The role of chemotaxis in the ability of *V. cholerae* to penetrate the mucous gel has been investigated in germ-free and infant mice as well as rabbit ileal loops. Richardson (1991) assessed the contributions of flagellar structure and motility to *V. cholerae* pathogenesis in rabbit loops (RITARD), and infant mice. All three models revealed that motility was the major contributor to enhanced virulence. The RITARD model indicated that flagellar structure (or LPS) may play a subtle role in virulence.

1.7.1.3 Adherence

When cholera vibrios reach the small intestine, they must adhere to and colonize the host epithelial cells. Many factors have been implicated in adherence. Two types of bacterial adherence mechanisms are described, namely fimbrial (or pili), and afimbrial adhesins.

1.7.1.3.1 Fimbrial adhesins

To date, the best understood mechanism of adherence is attachment mediated by rod-shaped protein structures called pili. Pili are essential components of the infection strategy for a variety of pathogenic bacteria, as they mediate specific binding and make an important contribution to colonization of host tissue. The mechanism by which *V. cholerae* O1 adheres to the intestine is not entirely understood. However, significant progress has been made in elucidating this process (Rabbani and Greenough III, 1992). Fimbrial adhesins have long been suspected as being important in vibrio adherence, and it now seems clear that pili are the critical factor in colonization (Sack, 1992). Several pili in *V. cholerae* have been described, for example TCP, mannose-sensitive haemagglutinin (MSHA), and ACF (Kaufman and Taylor, 1994).

1.7.1.3.1.1 Toxin-coregulated pilus (TCP)

The best characterized pilus of *V. cholerae* is the toxin-coregulated pilus first described by Taylor *et al.* (1987). Tcp proteins show homology to proteins associated with type IV pilus biogenesis, but TCP probably more appropriately belongs to a subclass which includes the bundle-forming pilus of enteropathogenic *E. coli* (Giron *et al.*, 1991, 1994). The genes within this locus are organized and expressed as an operon (Brown and Taylor, 1995). Regulation of TCP is controlled by the ToxR regulon via ToxT with the possible

involvement of TcpP, TcpH, TcpI, AphA, AphB, and cyclic AMP, which are needed for maximal *toxT* transcription (Shaw *et al.*, 1988; Taylor *et al.*, 1988; DiRita, 1994, Thomas *et al.*, 1995; Carroll *et al.*, 1997; Skorupski and Taylor, 1997; Häse and Mekalanos., 1998; Murley *et al.*, 1999).

Little is known about the interaction between TCP and host cells, and the host cell receptor for these pili has not yet been identified. TCP is essential for colonization of classical vibrios in infant mice and the human gut (Taylor *et al.*, 1987; Herrington *et al.*, 1988). Mutants carrying Tn*phoA* fused to *tcpA* (the major pilin subunit) are avirulent in human volunteers. These mutants are greatly reduced in their ability to colonize in animal models and in volunteers (Herrington *et al.*, 1988; Tacket *et al.*, 1998). Studies in the infant mouse cholera model have shown that TCP is a critical virulence determinant in classical strains and that passive antibodies are protective. Antibodies to TcpA provide a high level of protection when administered with virulent vibrios to infant mice (Sharma *et al.*, 1989; Sun *et al.*, 1990).

It has been recently shown that the genes encoding the structural and assembly components of the TCP are located on VPIϕ (Waldor and Mekalanos, 1996; Karaolis *et al.*, 1998, 1999). TCP itself functions as the CTXϕ (cholera toxin carrying bacteriophage) receptor. In addition, the TcpA subunit of TCP also functions as the coat protein of VPIϕ (Karaolis *et al.*, 1998, 1999). It is now known that TCP serves as both a bacteriophage receptor and colonization factor. However, the mechanism by which TCP functions is not understood, particularly as it has not been shown to bind directly to intestinal tissue or to cultured epithelial cells (Waldor and Mekalanos, 1996; Karaolis *et al.*, 1999).

1.7.1.3.1.2 Mannose-sensitive haemagglutinin (MSHA)

Another putative colonization factor in *V. cholerae* O1 that may induce anticolonization immunity is the MSHA. This adhesin was discovered by Lankford in 1959 and has been named MSHA because its ability to adhere is blocked *in vitro* by high levels of mannose and mannose derivatives (Richardson, 1992). MSHA is a pilus that is expressed on the surface of El Tor, but rarely on classical vibrios (Hanne and Finkelstein, 1982; Jonson *et al.*, 1992; Svennerholm *et al.*, 1992). It was initially thought to be an El Tor equivalent to TCP and was proposed to be the major adhesin of El Tor vibrios (Finn *et al.* 1987; Jonson *et al.*, 1991a, b). MSHA is a member of the type IV pili family (Marsh and Taylor, 1999) which are found on the surface of a variety of Gram-negative bacteria. It has been demonstrated as important as host colonization factors, bacteriophage receptors, and mediators of DNA transfer. In general, it was believed that classical *V. cholerae* attach mainly by TCP and that El Tor biotype primarily use MSHA (Jonson *et al.*, 1991a). However, this is no longer believed to be the case (Attridge *et al.*, 1996).

The function of MSHA was previously regarded as an important biotype-specific colonization factor (Osek *et al.*, 1994). However, inactivation of the *mshA* gene did not significantly diminish intestinal colonization in suckling mice and humans (Attridge *et al.*, 1996; Thelin and Taylor., 1996; Tacket *et al.*, 1998). Jouravleva and colleagues (1998) demonstrated that the MSHA pilus serves as the receptor for the filamentous bacteriophage 493, which was isolated from *V. cholerae* O139 strain and has been suggested to have a role in the horizontal gene transfer in the evolution of *V. cholerae*.

1.7.1.3.1.3 Accessory colonization factor (ACF)

The *acf* locus is a set of genes which encode another possible adhesin or colonization factor. This locus contains a cluster of four genes which is located on the

VPIϕ and is under the control of the ToxT protein (Higgins and DiRita, 1996; Kovach *et al.*, 1996; Karaolis *et al.*, 1998). These genes were identified by screening Tn*phoA* insertions in the *V. cholerae* chromosome producing active PhoA under conditions known to induce synthesis of Ctx (DiRita, 1992). The products of *acf* are outer membrane proteins that enhance colonization in the infant mouse model, and mutants affecting ACF show a decreasing ability to colonize the intestinal tracts of animals (Peterson and Mekalanos, 1988). The mechanism by which the Acf proteins contribute to colonization is not known, although ACF is thought to represent a minor pilus type (DiRita, 1992; Manning, 1994).

Everiss and colleagues (1994 a, b) have shown that the *tcp* and *acf* gene clusters are physically linked in the *V. cholerae* genome. In addition, it has been shown by Kovach *et al.* (1996) that the *tcp* and *acf* clusters are located on the VPIϕ which also contain the *toxT* gene and also an integrase *int* gene mediating the integration or excision of the phage. This phage-like integrase gene is located at the distal end of the *tcp-acf* gene cluster and a putative attachment site (*attI*) has also been identified at the end of this region (Kovach *et al.*, 1996). It was suggested that the TCP and ACF factors may interact in some manner to promote successful colonization of the small bowel (Everiss *et al.*, 1994 a, b). Slot blot analysis of DNA isolated from *V. cholerae* O1 and non-O1 strains has revealed that the region of the *V. cholerae* chromosome encoding TCP, ACF is present in the Asiatic-cholera-causing strains but absent from the non-cholera vibrio strains tested by Kovach and colleagues (1996).

1.7.1.3.2 Afimbrial adhesins

Some bacteria have cell surface proteins that are clearly important for adherence but that do not assemble themselves into pili-like structures. Afimbrial adhesins mediate tight binding between bacteria and host cells which cannot be attained with pili alone. An

afimbrial adhesin in *V. cholerae*, the 38-kDa OmpU outer membrane protein, was characterized and its role in adhesion to mammalian cells was investigated (Sperandio *et al.*, 1995). The amino-terminal sequence of OmpU has similarity with the sequence of Fha (filamentous haemagglutinin), an important adhesin in *Bordetella pertussis*. Antibodies directed against OmpU or their $F(ab)_2$ fragments completely inhibit adhesion of several *V. cholerae* strains to HeLa, Hep-2, Caco-2 and Henle 407 epithelial cells and also inhibit intestinal colonization (Sperandio *et al.*, 1995).

It has been suggested that the flagellum of motile strains may possess an adhesive component which facilitates attachment to intestinal mucosa (Holt, 1982). Various haemagglutinins (HA) in *Vibrio spp.* have been identified and postulated that they may adhere to the surface of the small intestinal microvilli (Booth *et al.*, 1986). Finkelstein and Hanne (1982) described four distinct haemagglutinins and showed that all strains produced a soluble HA-protease. The D-mannose-, L-fucose-resistant HA (MFRHA) is one of the cell-associated HA's which is expressed by both *V. cholerae* biotypes. However, *in vivo* electron microscopic studies have shown that *V. cholerae* adheres to and colonizes the small intestinal epithelium, in the absence of the adherence of haemagglutinating activity and motility (Teppema *et al.*, 1987). In addition, Finkelstein *et al.* (1992) postulated that haemagglutinin mediates the detachment of the vibrios from host cell surfaces.

1.7.1.4 Colonization modifier

Although many colonization factors are involved in the virulence of *V. cholerae*, other proteins and hydrolytic enzymes also influence the colonization process. These factors include haemagglutinin/protease, neuraminidase and chitinase.

1.7.1.4.1 Haemagglutinin / protease (HA/P)

The haemagglutinin/protease (HA/P) of *V. cholerae* was originally described as a "mucinase" (Burnet, 1949). It was discovered as a secreted or "soluble" haemagglutinin and subsequently shown to be a zinc-and calcium-dependent protease (Finkelstein *et al.*, 1992). HA/P is a member of a large family of metalloproteases produced by both nonpathogenic and pathogenic *V. cholerae* strains (Finkelstein *et al.*, 1992). It has been shown to cleave several important substrates, including mucin, fibronectin and lactoferrin (Finkelstein *et al.*, 1983; Booth *et al.*, 1984). HA/P has also been demonstrated to be involved in the proteolytic activation of Ctx and activates the A subunit of Ctx by nicking it into the A_1 and A_2 fragments (Booth *et al.*, 1984).

The HA/P of *V. cholerae* is encoded by the structural gene, *hap*, which has been cloned from *V. cholerae* 3083 (El Tor) by Häse and Finkelstein (1991). Mutants in *hap* were found to be fully virulent in infant rabbits, suggesting that the HA/P is not a primary virulence factor for infant rabbits (Finkelstein *et al.*, 1992). Observations using cultured human intestinal cells by Finkelstein *et al.* (1992) indicate that HA/P prevents attachment of cholera vibrios at the intestinal epithelium and could therefore be responsible for detachment of the vibrios. The cytotoxicity of the cholera HA/P suggests that it might play some other roles in the pathogenesis of *V. cholerae*. However, since non-pathogenic *V. cholerae* strains also produce the HA/P, it is not regarded as a primary virulence determinant (Wu *et al.*, 1996). Benitez *et al.* (1997) demonstrated that inactivation of the HA/P increases binding of *V. cholerae* to the human intestinal cells. Thus, the HA/P may be a "detachase" which may act by destroying host cell receptors for several different putative *V. cholerae* adhesins, enabling release and transmission of the bacteria. Kimsey and Waldor (1998) suggested that the production of high levels of secreted HA/P may be a factor in preventing CTXϕ reinfection in natural environments and in the human host.

1.7.1.4.2 Neuraminidase (NANase) or sialidase

Neuraminidases have been implicated in the pathogenesis of many diseases. They seem to have a role during the initial stages of the infection, but are also produced by many non-pathogenic bacteria (Gashell *et al.*, 1995). *V. cholerae* NANase is part of a mucinase complex which may function in pathogenesis by degrading the mucin layer of the gastrointestinal tract (Crennell *et al.*, 1994). NANase can bind to cell surfaces in the small intestine (Crennell *et al.*, 1994), and remain active. Moreover, NANase can enhance the effect of Ctx by catalyzing the conversion of higher-order gangliosides to GM_1 (Holmgren *et al.*, 1992). The gene encoding NANase, *nanH*, has been cloned and expressed in *E. coli* by Galen and colleagues (1992), and DNA sequence data predicts a mature protein of 83.0-kDa for the extracellular enzyme. Isogenic *V. cholerae* strains with and without *nanH* deletions were also examined and it was concluded that NANase is not a primary virulence factor of *V. cholerae* (Galen *et al.*, 1992). It plays a subtle but definite role in the pathogenesis of *V. cholerae* by enhancing binding and uptake of Ctx by increasing the availability of its receptor.

1.7.1.4.3 Chitinase

Chitin is a homopolymer of N-acetylglucosamine (GlcNAc) which is found in the cell walls of fungi and in the integuments of insects and crustaceans. Chitinase production in *V. cholerae* presumably has relevance to its survival in the aquatic environment. *In vitro* experiments established that *V. cholerae* has the potential to use chitin as a sole source of carbon for growth (Nalin *et al.*, 1979; Garay *et al.*, 1985). Adherence of *V. cholerae* to chitin appears to be specific and the secretion of a chitinase may facilitate attachment and nutrient acquisition. Chitinase of *V. cholerae* is encoded by the *chiA* gene. The *chiA* gene encodes a polypeptide of 846 amino acids and has a predicted molecular mass of 88.7 kDa

(Connell *et al.*, 1998). All biotypes of *V. cholerae* have been shown by Colwell (1970) to produce chitinase. Chitinase from *V. cholerae* O1 and non-O1 may facilitate their adsorption and multiplication on different species of zooplankton and phytoplankton (Shukla *et al.*, 1995)

1.7.2 Toxins of *V. cholerae*

Although *V. cholerae* produces a variety of extracellular proteins that are cytotoxic to eukaryotic cells, Ctx is still acknowledged as the most important virulence factor. However, the occurrence of mild to moderate diarrhoea in human volunteers fed with candidate live oral cholera vaccines, or *V. cholerae* mutants incapable of producing active Ctx, suggests that other toxic proteins could also contribute to cholera virulence (Popovic *et al.*, 1994).

1.7.2.1 Cholera toxin (Ctx)

Ctx is responsible for the diarrhoeal disease caused by *V. cholerae*, as the strains or mutants that do not produce Ctx also do not cause significant diarrhoea in either animals or human volunteers (Sciortino, 1996). Ctx has been extensively studied at both the biochemical and genetic level, and it is currently one of the best understood of all bacterial toxins. Ctx is an A-B_5 type enterotoxin, with ADP-ribosylating activity. The toxin contains one A subunit and five identical B subunits. The five B subunits form a pentameric ring into which the A subunit is inserted (Holmgren *et al.*, 1991). In the mature toxin, the A subunit is proteolytically nicked to A_1 and A_2 proteins which are connected together covalently via a disulphide bridge. Both A and B subunits are required for Ctx activity, the A subunit having the ADP ribosylating activity while the B subunit is responsible for binding to the epithelial cell surface receptor, G_{M1} (King and van Heyningen, 1973;

Popovic *et al.*, 1994). The A subunit is cleaved into A_1 and A_2 fragments and the toxic activity of Ctx is found in the A_1 chain. The A_1 fragment ADP ribosylates a membrane protein called G_S. G_S regulates the activity of host adenylate cyclase, and thus determines the level of cyclic AMPin host cells. This alters the activities of sodium and chloride transporters, and induces the active secretion of water and salts from the mucosal cells into the intestinal lumen. Changes in the ionic balance causes the tremendous fluid loss associated with cholera (Popovic *et al.*, 1994).

Ctx is encoded by two genes, *ctxA* and *ctxB* which are part of the same transcriptional unit, the *ctxAB* operon (Miller and Mekalanos, 1985). Many strains of *V. cholerae* O1 contain multiple copies of the *ctx* operon (Otteman and Mekalanos, 1994). Classical strains contain two separate copies on the chromosome and about 30% of El Tor strains contain two or more adjacent copies (Mekalanos, 1983; Mekalanos, 1985). The duplicated *ctx* genes within each classical or El Tor strain appear to be identical (Ølsvik *et al.*, 1993). The *ctxA* and *ctxB* genes are carried on a filamentous lysogenic bacteriophage CTXϕ (Waldor and Mekalanos, 1996; Faruque *et al.*, 1998b). Karaolis *et al.* (1999) reported that the *ctxB* gene was identified on CTXϕ derived from VPIϕ. The expression of *ctx* and *tcp* genes is coregulated by the ToxR regulatory system, consisting of the proteins ToxR, ToxS, ToxT (DiRita *et al.*, 1991; DiRita, 1992; Faruque *et al.*, 1998a). The ToxR regulon is composed of a set of over 20 genes, and others have also been recently discovered including *aphA*, *aphB*, *tcpP* and *tcpH* which encode products that function to promote intestinal colonization, toxin production and survival within the host (Skorupski and Taylor, 1997; Faruque *et al.*, 1998a). ToxR is a transmembrane DNA-binding protein that is required for the transcription of ToxT and can also directly activate transcription of the *ctxAB* operon (Higgins and DiRita, 1996).

The *toxS* gene is located downstream of *toxR* and encodes a 19-kDa protein required by ToxR for activation of the *ctx* promoter (DiRita and Mekalanos, 1991). ToxS is also suggested by DiRita and Mekalanos (1991) to be a membrane-associated protein that resides largely within the cytoplasm. The ToxR, S system in *V. cholerae* plays a central role in modulation of virulence gene expression in response to environmental stimuli. The integration of multiple signaling inputs mediated by ToxR, -S, and -T controls virulence gene expression, leading to Ctx production (Wong *et al.*, 1998). The ToxT protein directly activates the transcription of virulence factors in *V. cholerae*, including Ctx and TCP (Schuhmacher and Klose, 1999). It has been shown that ToxT is primarily responsible for activating the expression of *ctxA* and *ctxB* (Champion *et al.*, 1997).

1.7.2.2 Other toxins produced by *V. cholerae*

Although Ctx is the most important toxin produced by *V. cholerae*, some strains are capable of producing other toxins. Mutant strains which cannot produce Ctx still cause mild diarrhoea in human volunteers suggesting the existence of other enterotoxin(s). Zonula occludens toxin (Zot), accessory cholera enterotoxin (Ace), and haemolysin/cytolysin are examples of other possible candidate toxins. The *zot* and *ace* genes along with *ctxA*, *ctxB*, *cep* (encoding core-encoded pilus) and *orfU* (encoding a product of unknown function) genes are known to be located in the core of the Ctx element, (initially called the virulence cassette) (Mekalanos, 1983; Trucksis *et al.*, 1993; Waldor and Mekalanos, 1996) which is present as the prophage state of the CTXϕ. Moreover, the product of *cep*, *ace*, *orfU*, and *zot* have been proposed to be involved in phage morphogenesis (Waldor and Mekalanos, 1996).

1.7.2.2.1 Zonula occludens toxin (Zot)

Zot was described by Fasano *et al.* (1991) as a putative enterotoxin of toxigenic strains of *V. cholerae* O1 which is distinct from Ctx. This toxin increases the permeability of the small intestinal mucosa by affecting the structure of the intercellular tight junction, or zonula occludens, allowing the passage of macromolecules through the mucosal barrier (Marinaro *et al.*, 1999). The tight junctions that bind mucosal cells together are normally so effective as a barrier that even ions do not diffuse readily between mucosal cells. Zot is thought to disrupt these junctions thereby disrupting the ion balance and causing diarrhoea. Marinaro *et al.* (1999) demonstrated that Zot is a novel potent mucosal adjuvant of microbial origin. A comparison with *E. coli* heat-labile enterotoxin (LT) revealed that the adjuvant activity of Zot is only seven-fold lower than that of LT. Moreover, Zot and LT induce similar patterns of Ova-specific IgG subclass antibody (Marinaro *et al.*, 1999).

The *zot* gene is now known to reside on part of the CTXϕ genome and is located immediately upstream of the *ctx* operon (Waldor and Mekalanos, 1996; Karaolis *et al.*, 1999). This gene consists of a 1.3-kb open reading frame that could potentially encode a 44.8-kDa polypeptide (Kaper *et al.*, 1994). The distribution of *zot* gene sequences includes both O1 and non-O1 strains of *V. cholerae* (Johnson *et al.*, 1993; Karasawa *et al.*, 1993). It has been shown that the *zot* gene is present only in strains that are *ctx*-positive and all *ctx*-positive strains carry *zot* (Johnson *et al.*, 1993). Koonin (1992) described the *zot* gene product to be homologous to a family of gene I products required for filamentous phage assembly such as the coliphage M13. Waldor and Mekalanos (1996) suggested that the *zot* gene product may have a role in CTXϕ morphogenesis. Fasano and colleagues (1998) regarded Zot as an important tool to gain more insight on the pathophysiology of the regulation of intestinal permeability through the paracellular pathway, and to develop alternative approaches for the oral delivery of drugs and macromolecules normally not

absorbed through the intestine. However, Waldor and Mekalanos (1996) argued that the biological activity previously designated "zonula occludens toxin" (Fasano *et al.*, 1991) is probably not directly associated with the *zot* gene product unless it has dual functions.

1.7.2.2.2 Accessory cholera enterotoxin (Ace)

Ace is the third enterotoxin in *V. cholerae* and has been identified by Trucksis and colleagues (1993). It has been shown to be enterotoxic in ligated rabbit ileal loops, and would consequently be expected to be diarrhoeagenic. Ace toxin is unrelated to Ctx and Zot. It causes diarrhoea in animals, but its role in human disease is unknown .

The *ace* gene is also a part of CTXϕ genome located immediately upstream of the *zot* gene (Trucksis *et al.*, 1993; Waldor and Mekalanos, 1996). The last codon of the *ace* open-reading frame overlaps with the start codon of *zot*. The predicted amino acid sequence of Ace shows striking similarity to eukaryotic ion-transporting ATPases, including the product of the cystic fibrosis gene in humans (Trucksis *et al.*, 1993). Interestingly, the *ace* gene product is also homologous to a small hydrophobic protein product of gene *VI* of *Pseudomonas* filamentous phage Pf1 assembled into virion particles (Waldor and Mekalanos, 1996). Ace increases the short-circuit current in Ussing chambers and causes fluid secretion in ligated rabbit ileal loops (Trucksis *et al.*, 1993). When crude toxin extracts were used in animal models, it was also shown that Ace increased transcellular ion transport, which is proposed to contribute to diarrhoea in cholera. However, the lack of purified protein has hindered the elucidation of the mechanism of action of Ace (Trucksis *et al.*, 1993). Trucksis and colleagues (1997) suggested that Ace appears to be bifunctional, having both toxin activity and functions in phage assembly.

1.7.2.2.3 Haemolysin/Cytolysin

A *V. cholerae* haemolysin was initially purified by Honda and Finkelstein (1979) from an El Tor biotype and characterized by its ability to haemolyze sheep erythrocytes and mammalial cells in culture. This cytolytic protein is commonly called El Tor haemolysin (ETH) or HlyA. ETH activity was traditionally a biochemical characteristic used to differentiate El Tor from classical strains. However, nearly all El Tor isolates worldwide are now non-hemolytic, except for El Tor isolates from the US Gulf Coast and from Australia (Barrette and Blake, 1981; Kay *et al.*, 1994). The haemolytic status of El Tor strains now seems to be quite variable making it less reliable as a biotype character (Manning, 1994).

ETH is encoded by the *hlyA* gene and it has been demonstrated that *hlyA* expression is upregulated by *hlyU* (Williams and Manning, 1991; Williams *et al.*, 1993). Iron stress is also known to increase HlyA activity (Stoebner and Payne, 1988). Nagamune and colleagues (1996) reported that HA/P is involved in the generation of the 65-kDa mature ETH from the 79-kDa precursor or pro-ETH. In classical strains the *hlyA* gene has an 11 bp deletion, resulting in the production of a 27-kDa truncated protein (Alm *et al.*, 1988; Alm and Manning, 1990).

In general, bacterial haemolysins have been suggested to be important virulence factors of the *Vibrionaceae*, causing haemorrhagic septicemia and diarrhoea (Hirono *et al.*, 1996). HlyA is both cytotoxic and enterotoxic in the rabbit ileal loop assay (Honda and Finkelstein, 1979) and has been reported to be a virulence determinant in the infant mouse cholera model (Alm *et al.*, 1991). ETH damages the target cells by the formation of oligomeric transmembrane channels, which cause osmotic cytolysis. The haemolysin channel frequently flickered in the presence of divalent cations, suggesting that the channel spontaneously opened and closed (Ikigai *et al.*, 1996, 1997). Sathyamoorthy and colleagues

(1997) showed that haemolysis is inhibited by sucrose, an osmotic protectant and suggested that the initial action of HlyA on erythrocytes is to raise the basal cation permeability of the cell membrane.

1.8 Chromosomal location of virulence genes on mobile genetic elements

It is now known that the *V. cholerae* genome consists of two circular chromosomes, designated chromosome 1 for the larger and chromosome 2 for the smaller of the two (Trucksis *et al.*, 1998; Heidelberg *et al.*, 2000). The virulence properties of *V. cholerae* are due to two important mobile genetic elements, a CTXϕ encoding the Ctx, and a VPI, which encodes the TCP. The CTXϕ genetic element is present on both chromosomes of *V. cholerae* O1 classical biotypes and toxigenic strains (Trucksis *et al.*, 1998). The VPIϕ is located on the larger chromosome.

1.8.1 Mobile genetic elements on the larger chromosome (chromosome 1)

The large chromosome of *V. cholerae* strain N16961 is 2,961,146-bp in size with an average G + C % content of 46.9%. Most genes required for growth, in particular the metabolic and biosynthetic pathways, as well these required for bacterial pathogenicity (VPIϕ and CTXϕ) are located on this chromosome (Heidelberg *et al.*, 2000; Waldor and RayChaudhuri, 2000). It also contains the genes encoding DNA replication and repair, transcription, translation, and cell-wall biosynthesis (Heidelberg *et al.*, 2000). Approximately 42% of chromosome 1 contains hypothetical genes or genes of unknown function.

1.8.1.1 CTXϕ genetic element

The CTXϕ genetic element varies in size in individual strains of *V. cholerae* from a 7- to 9.7-kb segment of DNA. It has the structure of a compound transposon such as Tn*9* (Mekalanos, 1983; Pearson *et al.*, 1993). The CTXϕ genome has two regions, the 'core' and the direct repetitive sequence termed RS2 (Goldberg and Mekalanos, 1986; Pearson *et al.*, 1993; Waldor *et al.*, 1997). The core region is a 4.6-kb segment, contains *ctxAB* as well as four other genes (*zot*, *ace*, *cep* and *orfU*), and is flanked by two or more copies of a 2.7-kb repetitive sequence RS1 which is often duplicated immediately flanking the CTXϕ (Pearson *et al.*, 1993; Waldor and Mekalanos, 1996). The border between the core and RS2 region of the CTX element is located at the point where restriction mapping and DNA sequencing show that RS2 and RS1 diverge (Waldor *et al.*, 1997). The RS2 sequences encode at least four open reading frames which together determine expression of a site-specific recombination system that catalyzes the integration of plasmids carrying portions of the CTXϕ into the chromosomes of non-toxigenic *V. cholerae* strains at a particular attachment site called *attRS*1 (Waldor and Mekalanos, 1996). The copy number, location and arrangement of CTX prophages in the chromosomes differ in El Tor and classical biotype strains. The El Tor strains contain either a single copy or multiple copies of CTX prophage(s) which are tandemly arranged on the same chromosome. The classical strains always contain two copies of CTX prophages (Mekalanos, 1983; Sharma *et al.*, 1997a; Sharma *et al.*, 1997b; Trucksis *et al.*, 1998).

1.8.1.2 VPIϕ genetic element

The VPIϕ genetic element is approximately 40-kb in size. It contains genes associated with virulence, regulation and mobility. VPIϕ has a low G+C % content (35%),

contains putative integrase and transposase genes, is flanked by *att* sites, and inserts near a 10S RNA gene (*ssrA*). These *attI* sites presumably function as specific attachment sites for recombination between the VPIϕ element and the host bacterial chromosome (Karaolis *et al.*, 1998, 1999). Kovach *et al.* (1995) have found that the region contains an integrase gene (*int*) and an *att* site that marks the right end of a unique locus in pathogenic strains. In addition, two ToxR-regulated genes (*aldA* and *tagA*), as well as several ORFs with homologies to sequences not previously known to occur in *V. cholerae*, are also located on the VPIϕ (Karaolis *et al.*, 1998). VPIϕ is present in epidemic and pandemic strains but absent from non-pathogenic strains. Karaolis *et al.* (1998) suggested that it can be transferred within *V. cholerae* and this has been indeed shown to be the case.

1.8.2 Mobile genetic elements on the smaller chromosome (chromosome 2)

The smaller chromosome is 1,072,314-bp in size with an average G + C % content of 47.7%. A few of the genes required for growth, viability and essential for normal cell function are found in this chromosome, but it contains a larger fraction (59%) of hypothetical genes and genes of unknown function (Heidelberg *et al.*, 2000). Chromosome 2 contains many intermediaries of metabolic pathways, and also the CTXϕ element in *V. cholerae* classical biotype and the toxigenic strains (Trucksis *et al.*, 1998). In addition, it carries a 125-kb integron island which is discussed in **Section 1.10**.

1.9 Bacterial repetitive sequences

Repetitive sequences have been identified in a wide variety of eubacterial chromosomes (Lupski and Weinstock, 1992), and the best characterized of these are the repetitive extragenic palindromic (REP) sequences or palindromic units (PU) sequences, identified in members of the family Enterobacteriaceae (Gilson *et al.*, 1984; Stern *et al.*,

1984). In addition, Gilson and colleagues (1991) demonstrated that PU belong to a larger repeated DNA element, of up to 300 nucleotides, called BIME for bacterial interspersed mosaic element. Hulton and colleagues (1991) also identified the enterobacterial repetitive intergenic consensus (ERIC) sequences. These repetitive sequences display both features of similarity as well as differences, and will be discussed further below.

1.9.1 Repetitive extragenic palindromic (REP) sequences or palindromic unit (PU) sequences

The REP family is composed of hundreds of copies per genome distributed throughout the chromosome in both *E. coli* and *Salmonella typhimurium* (Gilson *et al.*, 1984; Stern *et al.*, 1984; Yang and Ferro-Luzzi Ames, 1990). REP consists of a highly conserved 33-bp inverted repeat sequence, and the copies are estimated to comprise about 0.5% of the genome. These sequences are arranged in clusters, from one to six copies, and are always found outside structural genes (Gilson *et al.*, 1990). No example of a REP sequence has yet been found within the coding sequence for a protein (Stern *et al.*, 1984). It has a highly conserved dyad symmetry that could permit the formation of a stem-loop structure in its corresponding RNA transcript and most are located at the 3′ terminus of a structural gene (Yang and Ferro-Luzzi Ames, 1990).

The origin and functions of REP sequences are still obscure (Gilson *et al.*, 1990). Several functions have been proposed, and one hypothesis is that they are involved in the regulation of intra-operonic gene expression. A second alternative is that REP sequences could play a role in chromosomal rearrangements and generation of genetic diversity (Stern *et al.*, 1984; Yang and Ferro-Luzzi Ames, 1988). A third hypothesis is that REP sequences are involved in chromosome structure and organization (Stern *et al.*, 1984; Yang and Ferro-Luzzi Ames, 1988). A final suggestion is that REP sequences could serve as specific

anchorage sites for the bacterial supercoiled nucleoid domains (Gilson *et al.*, 1987; Yang and Ferro-Luzzi Ames, 1988).

The observation that REP binds nucleoid-associated proteins could provide a plausible cause for their sequence homogeneity (Gilson *et al.*, 1990). Stern and colleagues (1991) demonstrated that deletion of the intergenic REP sequences from the chromosome has little effect on the transcription or translation of downstream sequences and does not cause a major alteration in the ratio of the upstream and downstream proteins. Although many roles for REP sequences have been suggested, it is unclear whether any of these could account for their extensive sequence conservation and distribution within and between species. As a result, it has been suggested that REP sequences may represent a 'selfish' sequence in nature, having no function other than self perpetuation. They are maintained by gene conversion (Higgins *et al.*, 1988), although they may sometimes confer an advantage on their carriers (Chao *et al.*, 1983).

1.9.2 Bacterial interspersed mosaic element (BIME)

BIME is a large repeat DNA element of up to 300 nucleotides (Gilson *et al.*, 1991a). The *E. coli* chromosome contains about 300 BIMEs (Espeli and Boccard, 1997), and these elements are located at the 3′ end of REP. BIME functions have been studied by two main approaches : effects of BIME on local gene expression, and biochemical characterization of BIME-protein interactions (Gilson *et al.*, 1991a). The abundance of REP sequences in BIME suggests that they may serve one (or several) physiologically important function(s). Their sequence homogeneity could indicate at least a common feature in the role of all REP, and BIME could belong to several functional classes depending on the nature of the motifs (Gilson *et al.*, 1991b). It has been shown that DNA gyrase and DNA polymerase I are able to specifically recognize BIME DNA *in vitro*. These

findings suggest that BIME could play a role in the functional organization of the bacterial nucleoid (Gilson *et al.*, 1991a). In addition, some BIMEs have been implicated in the protection of mRNA against 3′ exonucleolytic degradation, but the main role of elements belonging to this family remains to be elucidated (Espeli and Boccard, 1997).

1.9.3 Enterobacterial repetitive intergenic consensus (ERIC) sequences

The presence of ERIC sequences was demonstrated for the first time in the genome of *Mycobacterium tuberculosis* (Eisenach *et al.*, 1990). These sequences have been found in transcribed regions of the chromosomes of Gram-negative bacteria and identified in *E. coli* and *S. typhimurium* (Hulton *et al.*, 1991). Sequences are 126-bp long and appear to be restricted to transcribed regions of the genome, either in intergenic regions of polycistronic operons or in untranslated regions upstream or downstream of ORFs. The *V. cholerae* ERIC sequence is highly similar to those in *E. coli* and *S. typhimurium* (Hulton *et al.*, 1991).

ERIC sequences are imperfect palindromes (Sharples and Lloyd, 1990) which are highly conserved at the nucleotide sequence level, but their chromosomal locations differ between species (Hulton *et al.*, 1991). Several features of ERIC sequences resemble those of REP sequences (Stern *et al.*, 1984), although the nucleotide sequence is entirely different. However, REP sequences are much shorter than ERIC (33 bp compared with 126 bp), and while multiple REP sequences are often found at a single chromosomal location, all the ERICs identified so far occur singly. The estimated number of REP sequences per chromosome in *E. coli* is at least 10 times that of ERIC (Hulton *et al.*, 1991). ERICs are capable of forming two different secondary structures in single stranded RNA with free energies of less than -40 kcal mol^{-1} (Cromie *et al.*, 1997). Several functions for ERIC have been proposed, and they are similar to those proposed for REPs. It is considered that ERICs

and REPs are independent classes of 'selfish' DNA sequences in enterobacteria, and appear to be maintained and dispersed by gene conversion rather than providing any specific cellular function (Hulton *et al.*, 1991).

1.9.4 *Vibrio cholerae* repetitive (VCR) DNA sequences

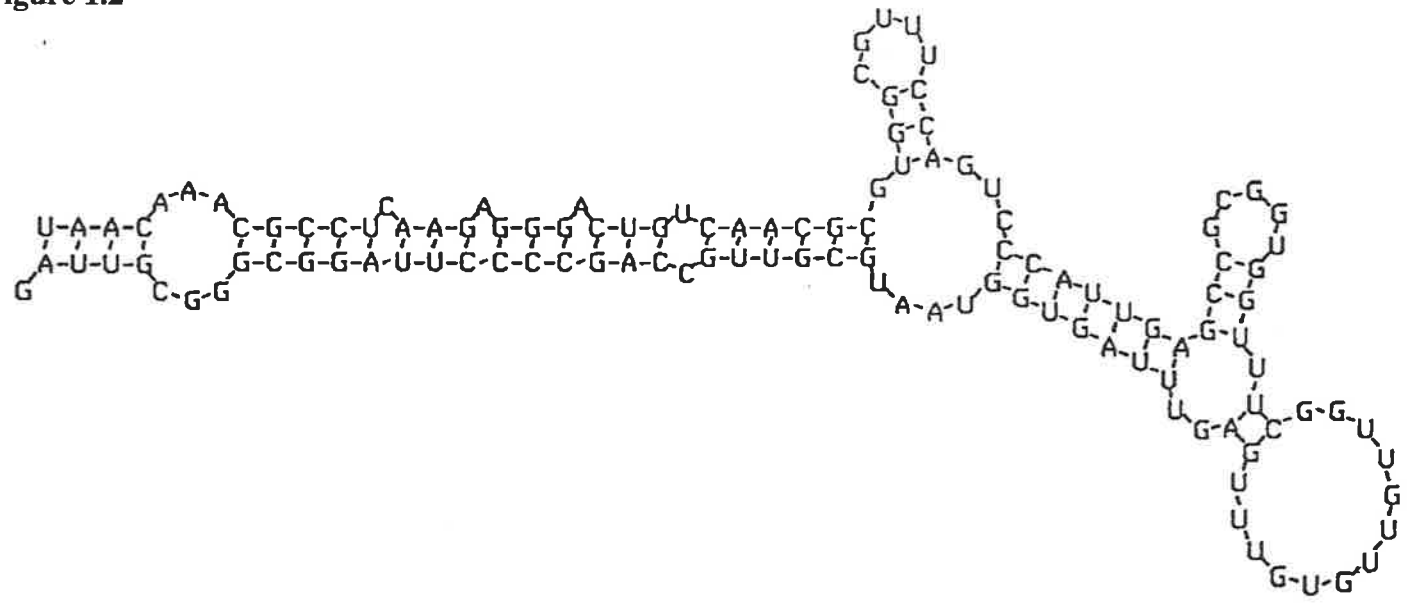
A large region representing over 10% of the smaller chromosome of *V. cholerae* has been found to contain arrays of single genes that are separated by short inverse repeat sequences of 123-126-bp, known as *V. cholerae* repetitive DNA sequences (VCRs) (Franzon and Manning, 1986; Franzon *et al.*, 1993; Barker *et al.*, 1994; Clark *et al.*, 1997; Rowe-Magnus *et al.*, 1999). More than 150 copies of VCR have been found in *V. cholerae* El Tor strain N16961 which have been shown to constitute a chromosomal integron (Rowe-Magnus *et al.*, 1999; Heidelberg *et al.*, 2000). VCRs are also found in a number of other *Vibrio* species including *V. metschnikovii, V. mimicus,* and *V. parahaemolyticus* (Mazel *et al.*, 1998).

Like many other bacterial repetitive sequence elements, the VCRs themselves are extremely conserved, such that pairwise comparisons reveal about 95-97% identity, with the major variation occurring at the ends of the repeats (Clark *et al.*, 1997). VCRs show several regions of dyad symmetry, and the whole repeat shows imperfect symmetry (Figure 1.2) (Barker *et al.*, 1994). VCR is able to form an extensive stem-loop structure like ERIC, and predicted free energies of the consensus sequence are -58.9 kcal mol^{-1} (Barker *et al.*, 1994). Although VCR and ERIC have a similar length and the potential to form a stem-loop structure, there is no sequence homology (Barker *et al.*, 1994), and their abundance and distribution are quite different (Clark *et al.*, 1997).

Barker *et al.* (1994) suggested that VCR might play a role in transcription termination in *V. cholerae*, but it seems unlikely that all copies of VCR act as

Figure 1.2 Proposed secondary structure for the VCR consensus sequence (a stem-loop free energy of -58.9 kCal / mol^{-1}) shows several regions of dyad symmetry, and the whole repeat shows imperfect symmetry (from Barker and Manning, 1994).

Figure 1.2



transcriptional terminators. This hypothesis is similar to REP sequences which have been shown to act as transcriptional terminators (Gilson *et al.*, 1986). However, Stern *et al.* (1991) showed deletion of REPs did not alter transcription of downstream sequence. It is now clear that VCR sequences share many features of the 59-base elements (59-be) found in class 1 integrons which will be discussed in the following section.

1.10 Integrons

1.10.1 Definition

Integrons are genetic elements that contain the determinants of a site-specific recombination system which enables them to recognize and capture mobile gene cassettes (Stokes and Hall, 1989; Hall and Collis, 1995). The essential components of an integron are an integrase gene (*intI*) encoding a site-specific recombinase, an adjacent site *attI*, which is the attachment site for the cassettes, and a promoter P_{cass} or P_c (formerly P_{ant}), suitably oriented for expression of the cassette-encoded gene (Hall and Collis, 1995; Nass *et al.*, 1998). Integrons consist of two conserved sequence regions, which are normally referred to as 5′- and 3′-conserved segments (5′- CS and 3′- CS), flanking a central variable region in which antibiotic resistance gene cassettes are found (Hall and Collis, 1995). The *intI* gene is located in the 5′- CS of the integron structure, and the *attI* site is closed to the *intI* gene (Figure 1.3A). Integron in *V. cholerae* is carried on the smaller chromosome which is 1.0-Mb in size, and contains all copies of the VCR sequence and 216 ORFs (Rowe-Magnus *et al.*, 1999; Heidelberg *et al.*, 2000).

1.10.2 Classes of integrons

To date, integrons have been classified into four classes according to the degree of divergence of their integrase gene (*intI*1, *intI*2, *intI*3) (Recchia and Hall, 1995; Hall and

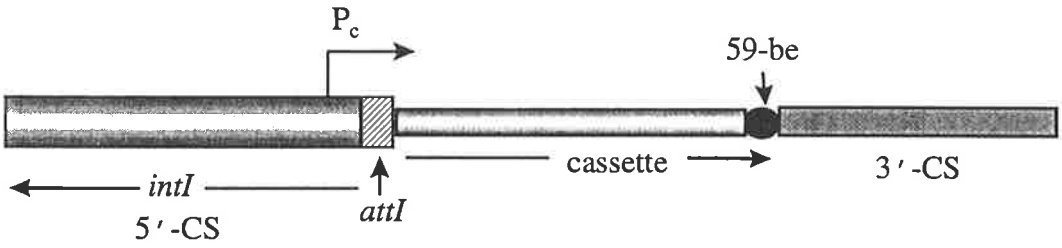
Figure 1.3 Generalized structure of an integron

(A) Generalized structure of an integron showing the integrase gene, *intI*, the recombination site *attI* (striped box), the promoter P_c, which transcribes the inserted cassettes, and one integrated gene cassette. The 59-base element is represented by a filled circle. The coding region of the *intI* and cassette gene is marked by horizontal arrows, and a vertical arrow indicates the position at which cassettes are inserted (adapted from Hall, 1995).

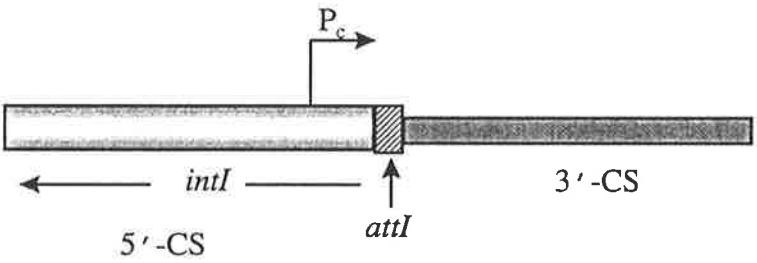
(B) The integron shown contains no cassettes. The *attI* site is represented as a striped box. In this cassette-free integron configuration, the 5′ and 3′ - CS (conserved segments) are connected (adapted from Hall and Collis, 1998).

Figure 1.3

(A)



(B)



Collis, 1998; Rowe-Magnus and Mazel, 1999) (Figure 1.4). The most recent, the class 4 integron containing the *intI*4 gene was identified. This integrase is chromosomally located in *V. cholerae* and unlike previously described (Barker *et al.*, 1994; Clark *et al.*, 1997; Recchia and Hall, 1997; Mazel *et al.*, 1998; Rowe-Magnus *et al.*, 1999). In addition, Rowe-Magnus *et al.* (1999) have recently described two new integrons in *V. mimicus* (IntI5) and *V. metschmikovii* (IntI6) forming a class 5 and class 6 integron, respectively, based on integrase gene divergence.

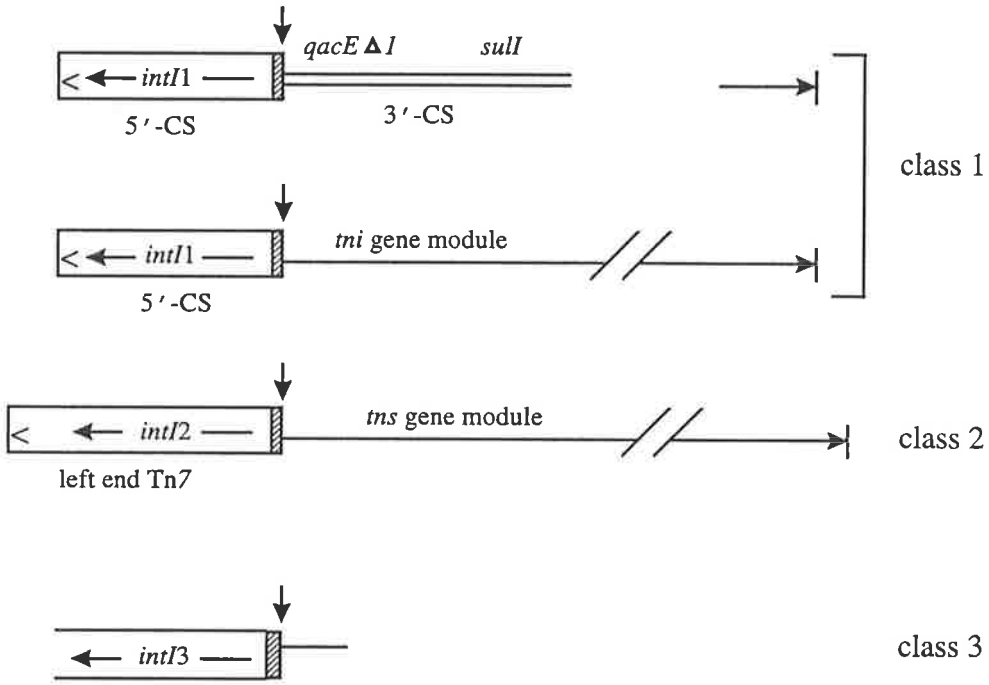
Class 1 integrons consist of three DNA regions, two conserved segments (5′-CS and 3′-CS) and one of variable length which is the sequence that includes the inserted antibiotic resistance gene cassettes (Recchia and Hall, 1997). However, the absence of an inserted gene in this class of integron has been found (Bissonnette and Roy, 1992; Collis and Hall, 1992a), with the 5′-CS and 3′-CS connected (Figure 1.3B). Many class 1 integrons are found in transposons of enteric bacteria and *Pseudomonas*, and commonly include a sulphonamide resistance determinant, *sulI*, located 3′ to the integrated cassettes known as the 3′-CS (Recchia and Hall, 1997). Class 1 integrons of the *sulI* type are the most prevalent in clinical isolates and have been found in many different organisms (Hall and Collis, 1998).

The transposon Tn*7* is the best known representative of class 2 integrons. Class 2 integrons have the defective *intI*2 gene, whose product is 40% identical to that of *intI*1 (Hall and Collis, 1998). The *intI*2 gene at the left end of Tn*7* is interrupted by an in-frame termination codon rendering the integrase inactive, and is located 5′ to the first cassette at the right end of Tn*7* (Hall and Vockler, 1987). Tn*7* and its close relatives are found in many different bacterial species, and the sequence of the region containing the *intI*2 gene and *attI*2 site from class2 integrons is quite distinct (Hall and Collis, 1998).

Figure 1.4 Structure of three known classes of integrons (with no cassettes shown).

The point of cassette insertion is marked by vertical arrows and the coding regions of the *intI* genes are indicated by horizontal arrows. Class 1 and class 2 integrons are transposons or defective derivatives and open arrowheads indicate the position of 25 bp inverted repeat sequences found at their ends. The integrase of known class 2 integrons (eg. Tn*7*) is interrupted by a stop codon. The majority of known class 1 integrons include the *sulI* gene in the conserved region known as the 3′-CS, and most of these do not include a complete version of the transposition gene (*tni*) module. A variety of structures lie in the gap in the class 1 integron containing the 3′-CS. The 3′-CS is not found in the class1 integron Tn*402* that contains the complete *tni* module (from Hall and Collis, 1998).

Figure 1.4


qacEΔ1
sulI
intI1
5′-CS
3′-CS
class 1
intI1
tni gene module
5′-CS
intI2
tns gene module
class 2
left end Tn7
intI3
class 3

There is only one example in *Serratia marcescens*, of a class 3 integron to date (Aragawa *et al.*, 1995; Senda *et al.*, 1996). This integron has the *intI*3 gene located 5′ to the metallo-β-lactamase IMP-1 gene (bla_{IMP}) cassette which was first found in a class 1 integron and is part of the transposon Tn*3* (Osano *et al.*, 1994; Hall and Collis, 1998). The putative integrase (IntI3) is 61% identical to the IntI1 integrase (Aragawa *et al.*, 1995) and is quite distinct from the *intI*2 gene in class 2 integrons. It is not known whether class 3 integrons are associated with transposons that are found in class 2 integrons.

Class 4 integrons are specific for *V. cholerae* and contain predominantly unknown genes that are not associated with antibiotic resistance (Clark *et al.*, 1997; Mazel *et al.*, 1998). A significant portion (greater than 10%) of one of the two *V. cholerae* chromosomes constitutes a large integron (Clark *et al.*, 1997; Mazel *et al.*, 1998; Trucksis *et al.*, 1998). It was reported that this region had all the key features of an integron, and because of its size it was proposed to constitute a "mega-integron" (Clark *et al.*, 1997) or "super-integron" (Mazel *et al.*, 1998, Rowe-Magnus and Mazel, 1999; Rowe-Magnus *et al.*, 1999). The *V. cholerae* integrase gene (*intI*4) has 45 to 50% identity with the three known integrases (Mazel *et al.*, 1998).

The class 1 integrons appear to be the most abundant (Nesvera *et al.*, 1998) and are more commonly found in clinical isolates. This could be simply a reflection of the fact that they contain antibiotic resistances and are thus easily detected. Most integrons have been found in Gram-negative bacteria (Rosser and Young, 1999), and only two cases of conserved integron sequences present in Gram-positive bacteria have been reported (Kazama, *et al.*, 1998; Nesvera *et al.*, 1998). One of these was reported in *Mycobacterium smegmatis* (Martin *et al.*, 1993) and the another in *Rhodococcus erythropolis* NI 86/21 (Nagy *et al.*, 1997). Recently, Nesvera and colleagues (1998) reported the presence of an integron, InCg, in *Corynebacterium glutamicum*. An important role of integrons in Gram-

negative bacteria is antibiotic-resistance gene acquisition. Jones and colleagues (1997) revealed that many antibiotic-resistant bacteria isolated from patients with nosocomial infections carried integrons with inserted regions of DNA varying in size from 800 to 3,900 bp.

1.10.3 Integrase

The integrase protein (IntI) is encoded by the *intI* gene, and is a member of the tyrosine recombinase (phage integrase) family (Argos *et al.*, 1986; Abremski and Hoess, 1992; Esposito and Scocca, 1997). One member of this family, the well-studied lambda Int protein (λ integrase), promotes integration and excision of the phage genome from that of the host (Landy, 1989). The λ integrase is involved in the rearrangement of DNA duplexes by means of conservative site-specific recombination reactions (Gravel *et al.*, 1998a; 1998b). The integron integrase or recombinase protein has also been demonstrated to be essential for the excision, integration, and rearrangement of gene cassettes (Martinez and de la Cruz, 1990; Collis and Hall, 1992a; Collis and Hall, 1992b; Collis *et al.*, 1993; Hall, 1995).

To date, the integrase family includes over 140 members which constitute a highly diverse range of proteins (Gravel *et al.*, 1998b). IntI1 is a member of the λ integrase family on the basis of similarity between the conserved domain shared by members of this family of site-specific recombinases (Recchia and Hall, 1995). The IntI1 integrase is the best characterized recombinase to date, and recognizes both the *attI* site in the integron and the 59-be which is a recombination site found in gene cassettes. The recombinase catalyzes both integration and deletion or resolution events (Hall and Collis, 1998). The integration or excision of gene cassettes by IntI1 has been demonstrated within different antibiotic resistance genes (Collis and Hall, 1992a; Collis and Hall, 1992b; Hall and Collis, 1995).

However, deletion occurs ten times more frequently than duplication (Collis and Hall, 1992a). Integron-encoded integrases range from 43 to 58% amino acid sequence identity, and are related to the temperate bacteriophage λ*int* genes. Although IntI2 and IntI3 share only 45 to 60% amino acid identity with IntI1, they are assumed to be able to catalyze the same events as IntI1, based on the observation that the same cassettes can be found inserted in different integrons (Hall and Collis, 1998).

1.10.4 *attI* site

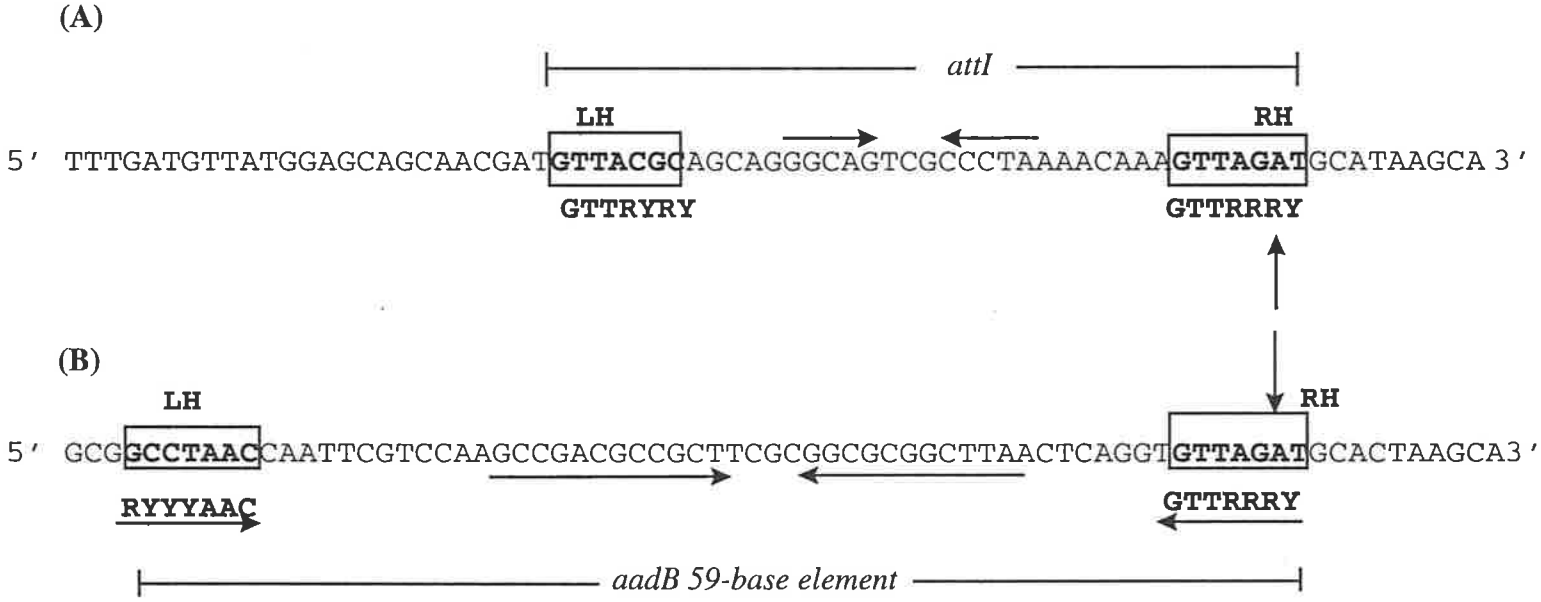
The *attI* site is the integron-associated recombination site that acts as a receptor for the insertion of cassettes into the integron (Recchia *et al.*, 1994). This site is located adjacent to the *intI* gene in the recipient structure and was previously assumed to be largely confined to the conserved sequences immediately 5′ to the inserted cassettes (Martinez and de la Cruz, 1990; Hall and Collis, 1995). The evidence that *attI* participates in recombination events was first reported by Martinez and de la Cruz (1990) and the sites were designated RHS for recombination hot-spots. The three *attI* recombination sites cannot be aligned to generate a consensus and are different from that of a short imperfect inverted repeat sequence, the 59-be or 59-base pair element (Stokes *et al.*, 1997; Hall, 1998).

A region of at least 40 bp but no more than 70 bp is required for full *attI* site activity (Recchia *et al.*, 1994), and this is made up of mainly conserved sequences located upstream of the first integrated cassette (Stoke *et al.*, 1997). The only obvious *attI* feature is the seven-base core site, GTTAGGC or GTTRRRY (R = purine, Y = pyrimidine), at the recombination crossover point (Figure 1.5). The GTTA residues are the most important in sequence recognition by the integrase protein (Gravel *et al.*, 1998a). However, the sequence GTTG instead of GTTA is found at the *attI* site in a few integrons, but

Figure 1.5 Comparison of the sequences of (A) *attI* and a representative (B) 59-base element (*aadB*).

The *attI* site is from class1 integrons. Horizontal arrows denote regions of inverted repeat sequences in each site. The core sites (GTTRRRY) are shown in bold and are boxed, and the vertical arrows show the position of the recombination cross-over. The inverse core sites (RYYYAAC) of the *aadB* 59-base element is also boxed. No inverse core site has as yet been identified in the *attI* sequence. Arrows indicate regions of imperfect dyad symmetry (from Recchia and Hall, 1995).

Figure 1.5


(A)
attI
LH
RH
5′ TTTGATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCTAAAACAAAGTTAGATGCATAAGCA 3′
GTTRYRY
GTTRRRY
(B)
LH
RH
5′ GCGGCCTAACCAATTCGTCCAAGCCGACGCCGCTTCGCGGCGCGGCTTAACTCAGGTGTTAGATGCACTAAGCA3′
RYYYAAC
GTTRRRY
aadB 59-base element

recombination at these sites is less common (Hansson *et al.*, 1997; Gravel *et al.*, 1998a). Therefore, a consensus sequence of GTTR would better define the specific recombination site (Gravel *et al.*, 1998a). The *attI* site was shown to participate in resolution events by Recchia and colleagues (1994), but no evidence for recombination between two *attI* sites was obtained.

1.10.5 59-base element (59-be)

1.10.5.1 Definition

A 59-be (also known as *attC*) is the cassette-associated recombination site and varies in length from 57 to 141 bp (Stokes *et al.*, 1997; Gravel *et al.*, 1998a). Although they are not highly conserved overall, comparison of the sequences of known 59-be has revealed that there are two regions of over 20 bp at each end that do show conservation and are related to a consensus sequence (Hall *et al.*, 1991; Collis and Hall, 1992b; Stokes *et al.*, 1997). The most highly conserved features within these regions are a 7-bp core site with the consensus GTTAGGC or GTTRRRY located at the right hand (RH) end and an inverse core site with the consensus GCCTAAC or RYYYAAC at the left hand (LH) end (Figure 1.5) (Hall *et al.*, 1991; Collis and Hall, 1992a; Stokes *et al.*, 1997). It seems likely that features of the 59-be, for example the differences between the LH and RH regions, have a role in ensuring that a cassette is inserted in the correct orientation (Stokes *et al.*, 1997).

Most 59-be contain inverted repeat sequences in their central region (imperfect dyad symmetry), and this region varies in length. When one gene is present in a cassette, the 59-be is located 3′ to the gene, i.e. between the coding sequence and the 3′-CS. When more than one gene is present, the 59-bes are found between gene pairs and at the end of the gene adjacent to the 3′-CS (Figure 1.3B) (Hall *et al.*, 1991). The term "59-be" was originally used because a consensus of 59 bases was derived by comparison of their

sequences (Cameron *et al.*, 1986), and although it is now known there is a range in both length and sequence, this name has been retained for all members of 59-be family (Hall *et al*, 1991).

1.10.5.2 Groups of 59-base elements

The sequences of the 59-be have been classified into three groups according to location and length by Hall and colleagues (1991). A 59-be in group 1 is found at the junction of the insert and 3′-CS, while in group 2 it is found between gene pairs. In group 3 it is found flanking the *aadA1* gene in Tn*7* (Cameron *et al.*, 1986; Hall *et al.*, 1991). Recchia and Hall (1997) identified several groups of closely related 59-be, and established that members of the largest group are 57 to 60 bp in length and differ pairwise by only a few bases (Table 1.1). A second group, including the *aacA4* and *aacA* (orfB) 59-be are 70 to 72-bp, while a third group includes the largest three 59-be discovered so far. The sequences of group 3 range from 127 to 141-bp in length (Recchia and Hall , 1997).

1.10.5.3 Functions

Although the 59-be are different in length and sequence, members of this element are active as sites for specific recombination or attachment recognized by the IntI1 integrases (Martinez and de la Cruz, 1990; Hall *et al.*, 1991; Collis and Hall, 1992a; Hall and Collis, 1995; Bunny *et al.*, 1995). Furthermore, Martinez and de la Cruz (1988, 1990) confirmed that two 59-be were the active sites for *int*-dependent site-specific recombination and these 59-be are located adjacent to the 3′-CS (Hall *et al*, 1991). Stokes and colleagues (1997) demonstrated that the recombination cross-over occurs close to one end of the 59-be, within a conserved core site with the consensus sequence GTTAGGC or GTTRRRY. In contrast, a particular 59-be is also associated with the 3′ end of a particular

Table 1.1 Groups of related 59-base elements.

Group 1 (57-60 bp)*	Group 2 (70-72 bp)*	Group 3 (127-141 bp)*
aadA1, aadA2	*aacA4*	bla_{IMP}
aadB	*aacA (orfB)*	*drfA7*
*aacA(*Iib)		*qacE*
sat		
catB3, catB5		
dfrB1, dfrB2, dfrB3		
orfd, orfE, orfF		

* Numbers in parentheses represent the range of lengths for members of each group (from Recchia and Hall, 1995).

gene to form a cassette or free circular cassette (Hall *et al.*, 1991; Collis and Hall, 1992b; Hall and Collis, 1995), and this cassette is mobilized within the integron areas by a gene deletion or insertion event.

Interestingly, 59-be have the same function as *attI*, although their structures are very different (Recchia *et al.*, 1994). However, the only obvious feature that *attI* shares with 59-be is a seven-base pair (7-bp) core site (GTTAGGC or GTTRRRY). Thus, IntI-mediated recombination can take place between either two 59-be sites or between *attI* and a 59-be, and the recombination cross-over occurs within the 7-bp core site at the 3′ end of the 59-be (or *attI*) (Hall and Collis, 1995).

1.10.5.4 59-base elements and VCR

In **Section 1.8** it was noted that the smaller chromosome of *V. cholerae* O1 carries a large integron island. This region has arrays of single genes that are separated by short inverted repeat sequences of 123-126 bp known as VCR (Barker *et al.*, 1994). Up to 150 copies of these VCRs are present in approximately 125-kb of this region of the *V. cholerae* O1 genome (Rowe-Magnus *et al.*, 1999). Recchia and Hall (1997) first reported that the outer ends of the VCR consensus sequence were found to be significantly related to the consensus sequence of 59-be. The 5′ and 3′ terminal sequences are conserved and correspond to the 59-be core sites (Manning *et al.*, 1999). VCRs have all the features of 59-be, but they are larger and associated with the $blaP_3$ β–lactamase gene which was found in group 3 of the 59-be. Although the $blaP_3$ β–lactamase is only partially sequenced, it is quite closely related to the VCR consensus sequence (63% over 89-bp) (Recchia and Hall, 1997). Both the 59-be and VCR can be folded to form a stem loop structure (Hall *et al.*, 1991; Barker *et al.*, 1994). These similarities suggest that the VCR region may have the same function for site-specific recombination as the 59-be.

1.10.5.5 Origins of 59-base elements

Little is known about the origins of the 59-be and why these elements are associated with gene cassettes. One suggestion is that the 59-be are related to the *E. coli* Rho-independent transcriptional terminators (Recchia and Hall, 1997). Three families of inverted repeat sequences, REP sequences (Stern *et al.*, 1984), ERIC sequences (Hulton *et al.*, 1991), and BIMEs sequences (Gilson *et al.*, 1991a) are found interspersed around the chromosomes of certain bacteria, usually in intergenic regions (Recchia and Hall, 1997). One of these families could also possibly represent the progenitor of the 59-be, although their consensus sequences do not conform to the outer ends of 59-be (Recchia and Hall, 1997). In these cases the 59-be would be expected to be part of the original transcript. In contrast, if the 59-be was added to the gene transcript before its conversion to DNA, a tRNA molecule could potentially serve as the source of the 59-be (Recchia and Hall, 1997). It is of interest to note that one of many integrase recognition sites, *attB* sites of lambdoid phages, lie within tRNA genes (Campbell, 1992).

1.10.6 Cassettes

1.10.6.1 Historical perspective

Cassettes were first identified in the early 1980s as the various integrated antibiotic-resistance genes found in integrons (Ward and Grinsted, 1982). Several antibiotic-resistance genes including *dfrB2* (*dbfrII*) from R388, *aadA2* from pSa, *aadA1* from Tn*21*, *oxa1* from Tn*2603* and *oxa2* from R46 were found to be flanked by identical sequences (Cameron *et al.*, 1986; Hall and Vockler, 1987; Ouellette *et al.*, 1987). Further antibiotic-resistance genes determining resistance to aminoglycosides (gentamicin, tobramycin, amikacin, kanamycin, streptomycin and spectinomycin), trimethoprim, chloramphenicol and β-lactam antibiotics were also found integrated at the unique site in the integrons (Hall

et al., 1991; Recchia and Hall, 1995). Subsequent studies demonstrated that the gene cassettes were mobile elements that could be excised from or integrated into an integron at the specific recombination sites, catalyzed by integron-encoded integrases (IntI) (Collis and Hall, 1992a; Collis *et al.*, 1993; Bunny *et al.*, 1995).

1.10.6.2 Definition

Cassettes are small mobile elements that most commonly consist of a single antibiotic-resistance gene and a short sequence ranging from 57 to 141-bp, i.e. 59-be (Hall *et al.*, 1991; Hall and Collis, 1995). The 59-be are located downstream or at the 3′ end of the gene (Figure 1.6A) and are recognized specifically by the integron integrase (Collis *et al.*, 1993). The sequences of over 40 cassettes are currently known and most of these contain an antibiotic resistance gene (Recchia and Hall, 1997). An open reading frame in cassettes is generally identified by the name of the gene they encode (Recchia and Hall, 1995), and many cassettes contain very little non-coding sequence (Recchia and Hall, 1997). In general, the initiation codons of the genes are located very close to one boundary of the cassette (Collis and Hall, 1995). Most cassettes do not contain a promoter and the cassette-encoded genes are expressed from a promoter P_c in the 5′ CS of the integron (Cameron *et al.*, 1986; Stoke and Hall, 1989; Hall and Collis, 1998). To date, all identified cassettes are in the same orientation, allowing expression of the associated genes from P_c (Stokes *et al.*, 1997). Gene cassettes differ from the well-characterized mobile elements (eg. insertion sequences, transposons, retroviruses) in that they do not encode the enzymes that catalyze their movement (Hall, 1995).

1.10.6.3 Functions

Cassettes vary considerably in total length ranging from 263 to 1549-bp. They can exist either free in circularized form or integrated at the *attI* site in the integrons (Figure 1.6B) (Recchia and Hall, 1995). The existence of circular gene cassettes is important in explaining the mechanism by which gene cassettes first became associated with integrons (Collis *et al.*, 1993). The integrase also catalyzes excisive recombination events that can lead to loss of cassettes from an integron and generate free circular cassettes (Hall and Collis, 1995). Although free circular cassettes are likely to be important intermediates in the dissemination of cassettes, they unable to replicate (Recchia and Hall, 1995; Hall and Collis, 1995). This spread of resistance genes occurs from one integron to another, as well as to non-specific sites (Hall and Collis, 1995). Furthermore, the integron cassettes may be transmissible from one bacterial cell to another, and even across species boundaries (Collis *et al.* 1993). The discovery of gene cassettes in a range of bacterial pathogens, particularly in the Enterobacteriaceae and Pseudomonadaceae indicates that inter-species transfer has occurred (Recchia and Hall, 1997).

1.10.6.4 Gene cassettes and transposons

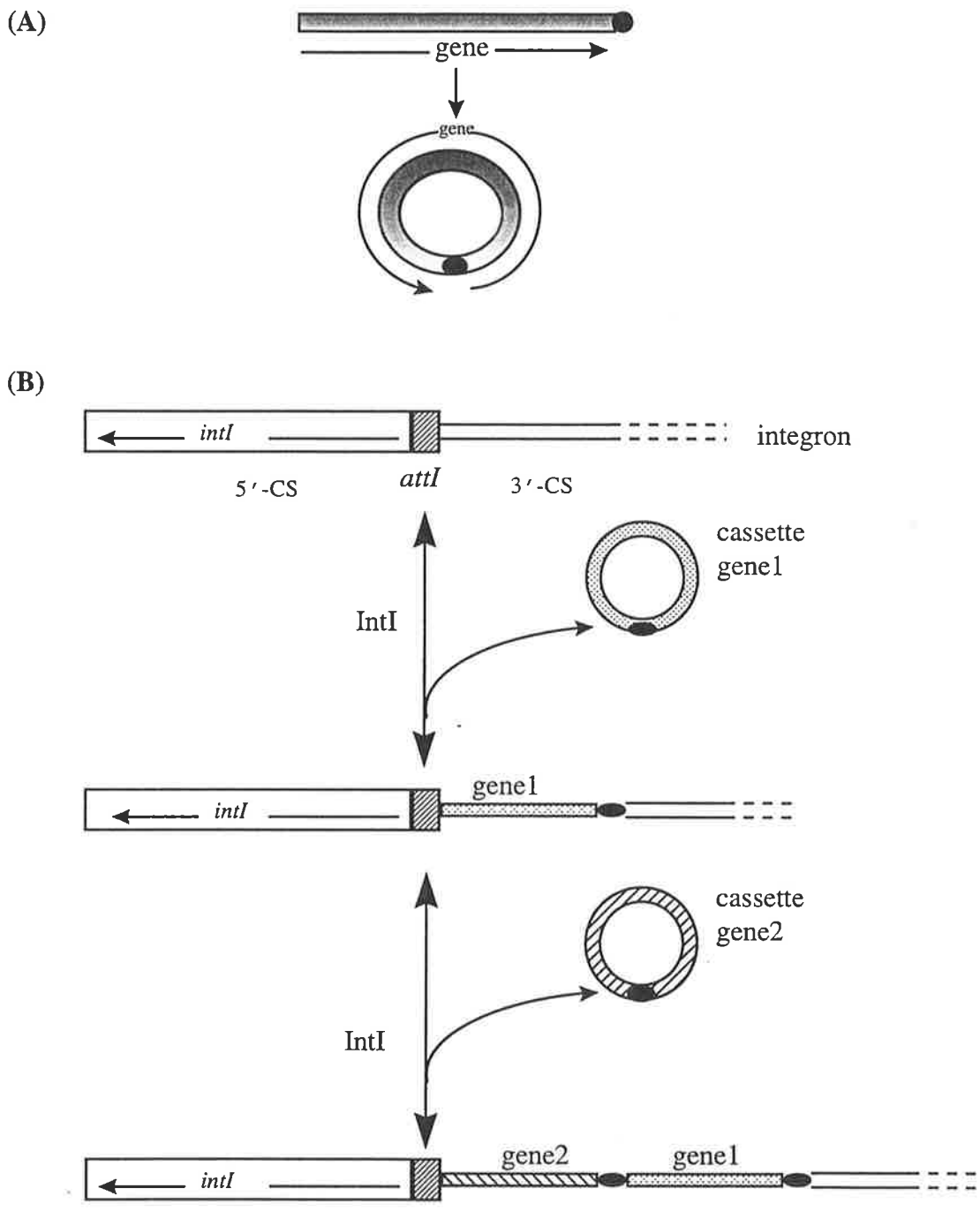
Although gene cassettes are mobile elements that can be integrated into or excised from integrons, they are quite distinct from transposons in many ways (Recchia and Hall, 1995). Firstly, gene cassettes do not encode any protein to catalyze their movement. Most other mobile elements encode the site-specific recombinase responsible for their movement, whereas the motility of cassette is catalyzed by an integrase encoded by the integron. Secondly, gene cassettes are not bounded by inverted repeat sequences, nor flanked by a duplication of target sequences. Thirdly, the mechanism used for mobilization of cassettes is distinct from transposition. Finally, all known gene cassettes are recognized

Figure 1.6 Structure, insertion and deletion of gene cassettes.

(A) Structure of a typical gene cassette in linearized form which can be formed as a circular gene cassette. The arrows indicate the extent of the gene coding regions, and the 59-base element recombination site is represented by a filled oval (adapted from Hall, 1995).

(B) Insertion and deletion of circular gene cassettes. Site-specific recombination between *attI* (striped box) and the 59-base element (filled oval) catalyzed by IntI results in integration of the gene cassette. The reverse reaction leads to loss of the cassette (from Hall and Collis, 1998).

Figure 1.6


(A)
gene
gene
(B)
intI
integron
5′-CS
attI
3′-CS
IntI
cassette
gene1
gene1
intI
IntI
cassette
gene2
gene2
gene1
intI

by the same integrase, despite differences in the precise sequences of their 59-be sites (Arakawa *et al.*, 1995; Recchia and Hall, 1995).

1.10.6.5 Fused cassettes

Although gene cassettes normally contain a single gene coding region and a 59-be, a cassette containing two potential coding regions has been reported (Bunny *et al.*, 1995). It seems likely that this cassette was created by the fusion of two cassettes each containing a single gene. Recchia and Hall (1995) suggested two potential routes to visualize how cassettes are fused (Figure 1.7). First, it involves a deletion event with end-points in two adjacent gene cassettes. This event will produce at least one truncated gene with the loss of a 59-be and a complete gene that still retains the 59-be, as has been demonstrated with the *catB4Δ1* - *aacA4* cassette fusion by Bunny and colleagues (1995). Similarly, fused cassettes that retain both complete genes may be created by any event involving loss of part or all of the 59-be from one cassette leading to its fusion with a second downstream cassette (Recchia and Hall, 1995). The complete 59-be normally associated with the *aadA1* gene is not present, but the element is present.

1.10.6.6 Origins of gene cassettes

Many models have been proposed to explain how different genes became part of a mobile gene cassette with a 59-be located downstream of the gene. Such models must explain why cassettes contain only one gene and no promoter. The most attractive hypothesis for the origin of gene cassettes is that they are formed by the reverse transcription of mRNA (Hall *et al.*, 1991; Recchia and Hall, 1995) as this would explain the fact that cassettes contain an open reading frame with little flanking DNA. Bacterial reverse transcriptases have been found to be encoded by retrons (Inouye and Inouye, 1991),

Figure 1.7 Two potential routes for fusion of adjacent cassettes. The 59-base elements are represented by filled boxes.

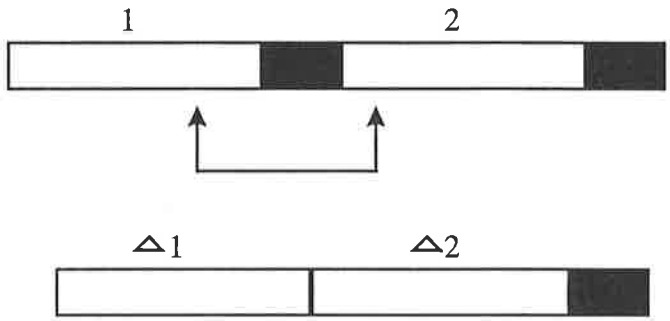
(A) Deletion event with end-points in each cassette resulting in the truncation of one or both genes.

(B) Loss of the 59-base element from one cassette resulting in the retention of both gene coding regions.

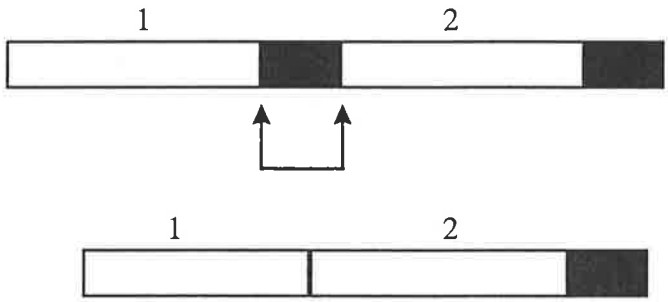
(from Recchia and Hall, 1995).

Figure 1.7

(A)



(B)



and it is possible that 59-be may have originated either from transcriptional terminators within the original transcript, or added later (Hall and Collis, 1995; Recchia and Hall, 1997) (Figure 1.8). However, the sources of single complete genes and the way in which the 59-be recombination sites become associated with them to create cassettes are not known (Hall, 1995). They may have evolved from independent common ancestors and cassettes are therefore likely to be very ancient structures, with the 59-be being functionally conserved over very long periods (Reccia and Hall, 1997).

1.10.7 Site-specific recombination

The recombination cross-over sites were previously identified in transposon Tn*21* and called the cross-over point. It has been clearly shown that the sequence GTTAG is the site of crossing-over in a conservative strand-exchange reaction (Martinez and de la Cruz, 1990). Hall and Collis (1995) reported that all genes found inserted at a specific site in integrons were shown to be associated with a GTTAG sequence at the 3′ ends of 59-be. The recombination cross-over point was localized to the GTT triplet. Interestingly, recombination at the highly conserved core site found in the inverse orientation at the 5′ end of the 59-be has not been detected (Hall *et al.*, 1991). Therefore, while the complete and correct orientation of a 59-be was found to be essential for cointegrate formation and Int-mediated site-specific recombination, the crossover point was restricted to the 3′ cassette (Hall *et al.*, 1991).

Stokes and colleagues (1997) demonstrated that the recombination cross-over was also localized to a unique position between the adjacent G and T residues. Substitution mutations in the 7-bp core site were generated, and it was demonstrated that the greatest effect was seen within the GTT triplet bases. The mutants containing either of the two substitution (T to G or T to A) at the second position of the triplet (G<u>T</u>T), showed an over

Figure 1.8 Model for the formation of cassettes.

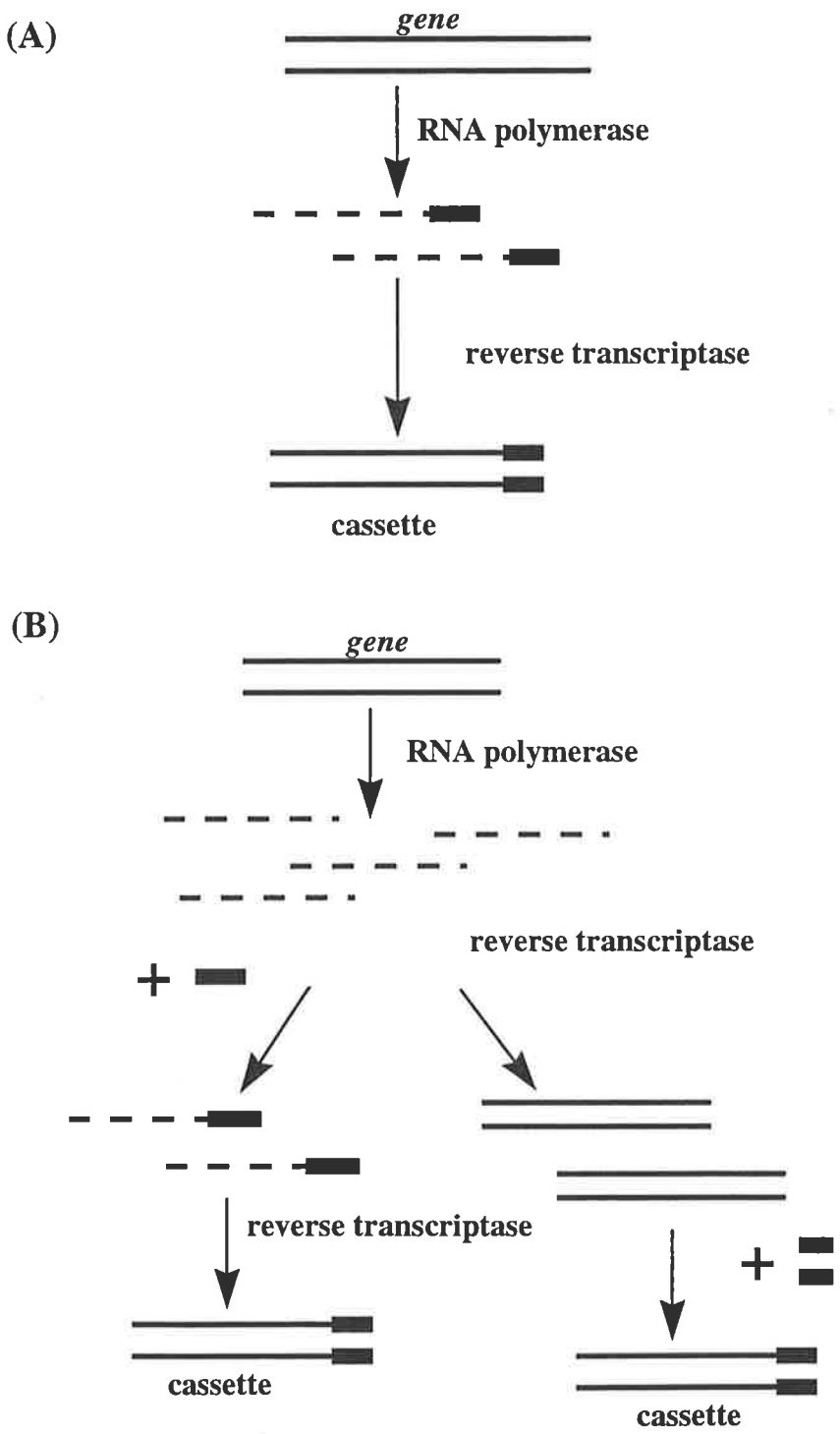
Alternative routes for the assembly of cassettes are presented. Cassette genes originate from mRNA molecules (dotted lines). The 59-base element recombination site represented by filled rectangles, may be either :

(A) part of the original transcript or

(B) added before or after the conversion of the mRNA to DNA (straight lines) by reverse transcription.

(from Recchia and Hall, 1997)

Figure 1.8


(A)
gene
RNA polymerase
reverse transcriptase
cassette
(B)
gene
RNA polymerase
reverse transcriptase
+
reverse transcriptase
cassette
+
cassette

1,000-fold reduction in the recombination frequency. Changes to the first and third positions (GTT) of the GTT triplet, and the transition mutation at the second position (GTT) reduced recombination activity 40 to 170-fold. Stokes *et al.* (1997) concluded that conservation of all three variant bases was very important for recombination activity, with the central T (GTT) being most critical. The last 4 bases (RRRY) of the identical 7-base core site are less critical than the GTT triplet. In contrast, base changes in the inverse core site (AAC triplet) found at the LH end of 59-be did not significantly affect the level of site activity, indicating that the identity of the invariant bases of the inverse core site is less critical than conservation of the core site GTT (Stoke *et al.*, 1997).

It was predicted by Collis and Hall (1992b) that if IntI1 used precisely the same mechanism as other integrases, recombination between two identical 59-be would occur at a higher frequency. However, it appears that the recombination efficiency is not enhanced by identity of sequences. Furthermore, Hansson *et al.* (1997) demonstrated that a region 27-36 bp at the 5′ end of *attI* influenced recombination with core sites in 59-be only, whereas a sequence 9-14 bp at the 5′ end of the cross-over point in *attI* was important for recombination with both *attI* and 59-be. In addition, it has been shown that the integron integrase is also able to catalyze recombination between one specific site (a 59-be) and secondary or non-specific sites outside an integron that conform to the consensus GNT (Gt/aT, Ga/tTNa/t) (Francia *et al.*, 1993; Recchia *et al.*, 1994), i.e. have some similarity to the core site consensus (Hall and Collis, 1995; Hall and Collis, 1998). Although these events occur at very low frequency, they are nonetheless likely to be important in the spread of cassette-associated genes to many different locations (Hall and Collis, 1995).

1.10.8 Evidence for insertion and deletion of gene cassettes in integrons

Hall *et al.* (1991) proposed that the recombination of a cassette into an integron was mediated by the integron-encoded integrase, and the event occurred between the *attI* site and the 59-be in a circular cassette (Figure 1.6B). Therefore, cassette insertion, deletion and rearrangement could occur only in the presence of the integrase (Hall *et al.*, 1991; Collis and Hall, 1992a; Collis and Hall, 1993). This was shown to be the case, in particular the insertion event, which was observed only if integrase was expressed in the recipient cell and the cassette DNA existed in a circular form prior to transformation (Collis *et al.*, 1993).

Several observations support the notion that site-specific insertion and deletion of genes can occur. It was demonstrated that the insertion and deletion of individual antibiotic-resistance genes were found in the variable regions of integrons (Collis and Hall, 1992a, 1992b; Collis *et al.*, 1993). The recombination sites of these cassette genes were found to occur at the 3′ end of the genes or the junction of the 5′ CS of 59-be. Collis *et al.* (1993) constructed circular gene cassettes *in vitro* which were identical in sequence to the circular molecules detected *in vivo*, and transformed recipient cells containing a plasmid based integron. It was found that integration of the transformed cassettes occurred at a specific insertion site (i.e. *attI*1) in an integron. In addition, excised and circularized cassettes have been found to integrate with preference for an *attI* at one end of the conserved sequence in integrons (Hansson *et al.*, 1997).

1.10.9 The expression of cassette-associated genes

Recchia and Hall (1995) revealed that recombination between *attI* or a cassette 59-be at a secondary site in the same genome can potentially lead to other genomic rearrangements. Different cassette genes may integrate at the same site, or integrate at the non-specific location in an integron. The multiple integration events are likely to be

important in the dissemination of several different cassettes which can be simultaneously resident in an integron (Hall and Collis, 1995). As mentioned in **Section 10.6.2**, cassette genes do not generally include a promoter to signal initiation of transcription. When more than one cassette is present, all transcripts detected commence at a common promoter P_{ant} or P_c (Hall and Collis, 1995) (Figure 1.9). The position of the cassette in the array is also important and influences the level of cassette-gene expression (Recchia and Hall, 1997). In all cases, the expression level is highest when the gene cassette is closest to P_c, and is reduced to different extents by the presence of individual upstream cassettes (Hall and Collis, 1995; Figure 1.10). This was explained by the fact that most of the integron mRNA is not full length, with only low levels of the full length transcript detected when more than one cassette is present. The shorter transcripts have different but discrete lengths, all originate from P_c, and they end at the positions that correlate well with the ends of the cassettes. In addition, the 59-be could either act as transcription terminators or be added to its conversion from mRNA to DNA mediated by reverse transcriptase (Recchia and Hall, 1995; Collis and Hall, 1995).

1.10.10 Integrons and transposons

Integrons are often found associated with insertion sequences (IS), transposons, and/or conjugative plasmids, which can serve as vehicles for intra- and inter-species transmission of genetic material (Hall and Collis, 1998, Rowe-Magnus *et al.*, 1999). It has been suggested that integrons are defective transposons located on mobile elements which facilitate their rapid spread to other strains and bacterial species (Jones *et al.*, 1997; Hall, 1998; Hall and Collis, 1998). Most class 1 integrons, whose structures have been examined in detail, are transposition-defective transposons or transposon remnants. Recently, the complete sequence of Tn*402*, which is a class 1 integron, has revealed that it resides on an

Figure 1.9 Locations of promoters in the 5′- conserved segment of integrons.

P_c is the common promoter present in all integrons, while P_2 is a secondary promoter present in only a few integrons, P_{intI} is the promoter for the divergently transcribed integrase gene (*intI*). Coding regions of inserted cassettes, designated gene 1 and 2, and of *int* are shown by horizontal arrows. 59-base elements (black ovals at the 3′ ends of cassettes), and *attI* (hatched box) are indicated (from Collis and Hall, 1995).

Figure 1.9

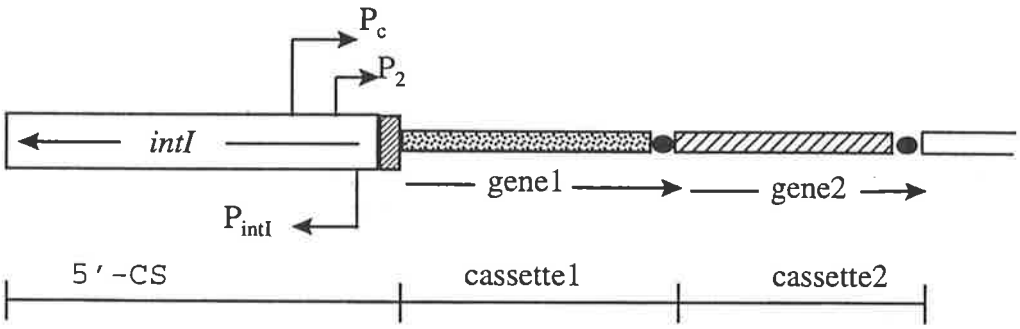
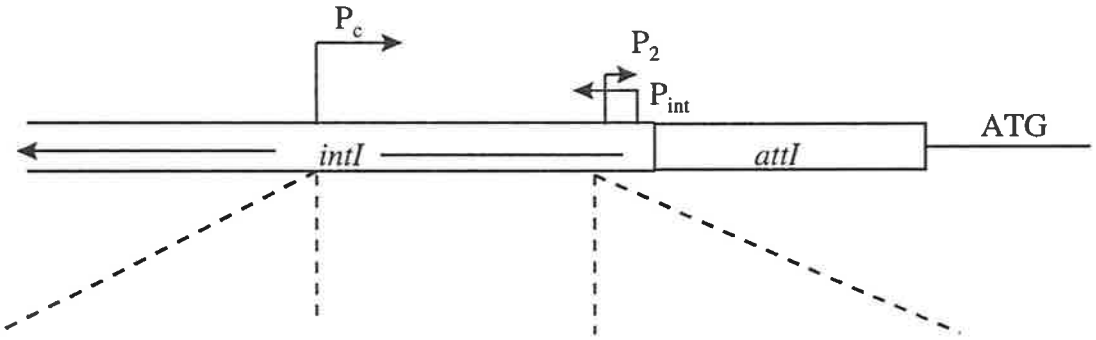


Figure 1.10 Promoter regions of class 1 integrons.

The position of the promoters P_c and P_2 used for transcription of antibiotic resistance genes, and of the *intI* gene (P_{int}), are represented by arrows that indicate the direction of transcription. Sequences of variants of the promoter -35 and -10 regions and the spacing between them are indicated. Strengths of the promoters relative to the strong version of P_c are represented as ++++, 100% activity; +++, 20-50% ; ++, approx. 10%; +, less than 10%; -, no activity. The initiation codon for translation of the cassette-associated (antibiotic resistance) genes is shown as ATG (from Hall and Collis, 1998).

Figure 1.10



TTGACA	TAAACT ++++	TTGTTA	TACAGT +++	strong
TGGACA	TAAACT ++	TTGTTA	TACAGT -	hybrid 1
TTGACA	TAAGCT ++			hybrid 2
TGGACA	TAAGCT +			weak
-35	-10	-35	-10	

active transposon (Hall *et al*, 1994; Rådström *et al.*, 1994; Recchia and Hall, 1995; Hall, 1998). Moreover, the general organization of the left end of Tn*7* (Figure 1.11) is similar to that of class 1 integrons. Integrons can be distinguished from transposons and insertion sequences by the mechanism of antibiotic-resistance gene acquisition which has been primarily associated with integrons (Hall and Collis, 1995). In addition, an integron does not of necessity contain any gene cassettes, indeed, a cassette-free class 1 integron has been found in nature (Recchia *et al.*, 1994), and also created experimentally (Recchia and Hall, 1992a).

1.10.11 An integron in the *V. cholerae* O1 genome

The *V. cholerae* chromosomally located integron was first described as a mega-integron by Clark *et al.* (1997), and later referred to as a super-integron by Mazel *et al.* (1998). The gene encoding the integrase in the *V. cholerae* integron, which was homologous to the IntI family of integrases, was identified as the *intI*4 gene (Mazel *et al.*, 1998). The product shows 45% to 50% identity with the three known integron integrases, but they still differ in several ways. The first recombination site, located at the beginning of the first cassette, has been named *attI*4 by homology to the attachment site in other classes of integrons. An array of four genes associated with VCRs is found upstream of *intI*4 in the characterized *V. cholerae* fragment (Mazel *et al.*, 1998). It was proposed that the *V. cholerae* integron exists for the purpose of gene capture and genomic evolution in *Vibrio* species (Mazel *et al.*, 1998; Manning *et al.*, 1999).

1.10.12 Integrons and evolutionary divergence

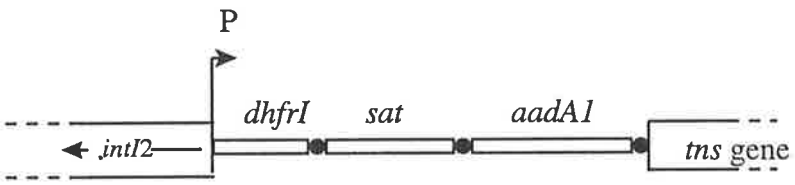
If integrons and cassettes make up a multicomponent system that is able to create arrays of antibiotic-resistance genes of enormous diversity, it seems reasonable to assume

Figure 1.11 Structure of the left end of (A) Tn*7* , compared with a (B) class 1 integron.

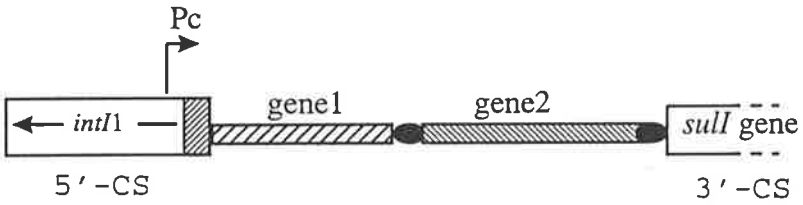
The *intI2* product is similar to IntI1 of *sulI*-associated integrons but it is interrupted by a stop codon. The position of the Tn*7* transposition genes (*tns*) is indicated. The promoter P in Tn*7* is presumed to be responsible for transcription of the genes in inserted cassettes. The position of P_c, the common promoter responsible for transcription of the inserted cassettes in the class 1 integron, is indicated. The *attI* region is present as a striped box and 59-base elements are filled circles (adapted from Hall and Collis, 1995).

Figure 1.11

(A)



(B)



that genes other than antibiotic-resistance genes may also be contained in cassettes. If this is so, the gene-cassette-integron system may be of much more widespread significance, and provide a general mechanism for the dissemination of modular gene units when gene cassettes act as modular gene packaging systems. This system is likely to be an important factor in facilitating the spread of antibiotic-resistance genes, and also in natural engineering of bacterial genomes. Another potential mechanism involves the ability of integrons to create cassette arrays or new operons or cassette arrays, that can then be fixed in place by deletion of 59-be. Stable new operons in plasmids and bacterial genomes may have occurred from cassette deletion or insertion and consequent fixation at secondary recombination sites (Recchia and Hall, 1995).

The fact that integrons are also located on mobile elements, transposons or defective transposon derivatives, which can relocate into other transposons or into plasmids (Hall and Collis, 1998), this suggests a role for integron in the gene cassette rearrangement. It is likely that cassette-associated genes have the ability to spread from one species to another either through the integron into a second integron, or directly via plasmids by transformation, transduction or conjugation. Not only are different cassettes found to spread within an integron, but integrons of different classes also appear to be able to acquire and carry the same gene cassettes (Sundstrom *et al.*, 1991; Aragawa *et al.*, 1995; Recchia and Hall, 1995). The degree of similarity between the three known integrases suggests that their evolutionary divergence extended beyond the half century of the antibiotic era, and probably much longer according to the substitution rates calculated by Ochman and Wilson (1987). Integrons act as natural cloning and expression vectors, and are responsible for gene acquisition (Recchia *et al.*, 1994; Stokes *et al.*,1997). Although integrons can mobilize various gene cassettes, integrons themselves are not mobile. Integrons are therefore thought to play a role in the horizontal transfer of bacterial genes.

As a result, integrons are an important factor in the development of multiply antibiotic resistant bacterial strains (Stokes *et al.*, 1997).

1.11 Lipocalins

1.11.1 General background

One of the genes encoded within the *V. cholerae* integron is the *V. cholerae* lipoprotein A, *vlpA* (Franzon *et al.*, 1993, Barker *et al.*, 1997). VlpA shows similarity to members of the lipocalin protein superfamily (Barker *et al.*, 1997). The lipocalin or α_2-microglobulin family is a large and diverse group of small extracellular proteins involved in the binding and transport of small hydrophobic ligands (Bishop and Weiner, 1996; Flower, 1996). Lipocalins are small proteins of around 200 residues, with molecular masses averaging 20 kDa. Most lipocalins have the features commonly found in extracellular soluble proteins, and exhibit an N-terminal signal peptide but lack other strong hydrophobic regions. Similarly, most lipocalins contain from one to three disulfide bridges that contribute to constraint of the overall structure by stabilizing the N- and C-terminal regions of the protein (Ganfornina *et al.*, 2000).

Lipocalins are generally found in eukaryotic organisms, mostly in vertebrates (Sansom *et al.*, 1994; Flower, 1996). The first lipocalin identified in prokaryotes was streptavidin in *Streptomyces avidinii*, a binding protein with a remarkable affinity for the vitamin biotin (Flower, 1993). However, Franzon *et al.* (1993) and Barker (1993) claimed to have first identified the first bacterial lipocalin in *V. cholerae*, a protein which has been designated VlpA. In addition, Bishop *et al.*, (1995) also claimed the bacterial lipocalin (Blc), the product of *f177* in *Escherichia coli* was the first lipocalin found in bacteria. A bacterial lipocalin has also been identified in *Citobacter freundii* OS60 (GenBank accession number U21727).

Lipocalins are also found in arthropods, for example butterfly insecticyanin (Holden *et al.*, 1987; Huber *et al.*, 1987), grasshopper lazarillo (Ganfornina *et al.*, 1995), cockroach Bla g 4 protein (Arruda *et al.*, 1995) and lobster crustacyanin (Keen *et al.*, 1991). There is evidence to suggest that carotenoprotein lipocalins may also be present in species from the phylum Coelenterata (Flower, 1996).

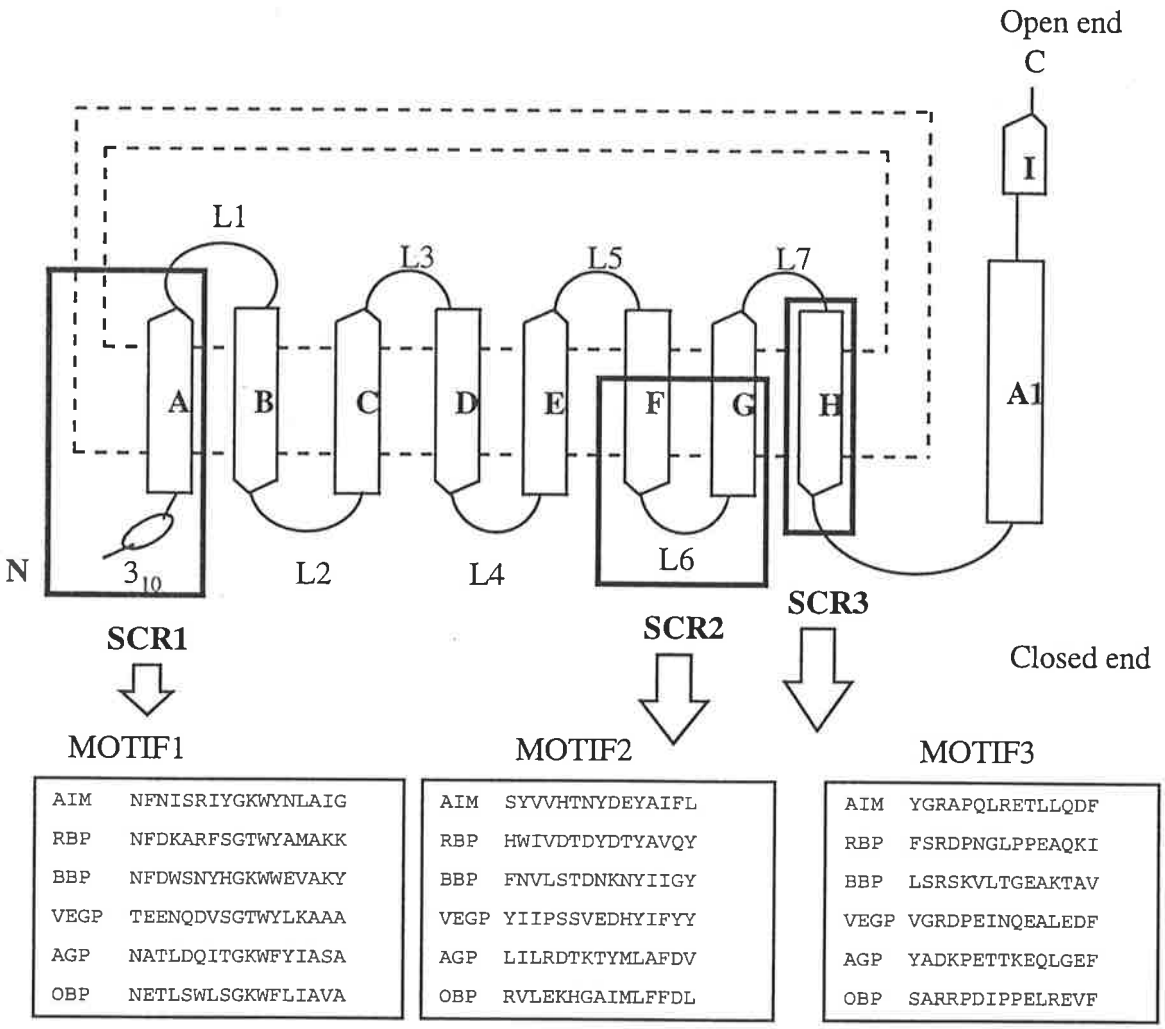
1.11.2 Structure of lipocalins

Members of the lipocalin family are characterized by their ability to bind a range of small, principally hydrophobic molecules (such as retinol), by their formation of soluble macromolecular complexes, and by their binding to specific cell-surface receptors (Flower, 1995; Flower, 1996). Lipocalins are a large and expanding group of proteins exhibiting great structural and functional diversity, both within and between species. The family demonstrates great diversity at the sequence level. However, they all share sufficient similarity, in the form of short characteristic conserved sequence motifs, providing a useful definition of family membership to be made (Flower *et al.*, 1991; Flower *et al.*, 1993). This analysis shows the lipocalin protein family to be composed of a core set of quite closely related proteins, the kernel lipocalins, and a smaller number of more divergent sequences, the outlier lipocalins (Figure 1.12) (Flower, 1996). Kernel lipocalins, which form by far the largest self-consistent subset within the whole set of related sequences, share three conserved sequence motifs, which correspond to the structurally conserved regions of the lipocalin fold (Flower, 1996). The first of these three characteristic motifs is shared by all lipocalins and is diagnostic for family membership. The outlier lipocalins match no more than two of these three motifs and are more diverse, forming distinct groups at the sequence level.

Figure 1.12 Characteristic features of the lipocalin fold.

An unwound view of the lipocalin fold orthogonal to the axis of the barrel. The nine β-strands of the antiparallel β-sheet are shown as arrows and labeled A-I. The N-terminal-like (3_{10}) helix and C-terminal α-helix (labeled A1) are also marked. The hydrogen-bonded connection of two strands is indicated by a pair of dotted lines between them. Connecting loops are shown as solid lines and labeled L1-L7. The two ends of the β-barrel are topologically distinct. One end has four β-hairpins (L1, L3, L5 and L7); the opening of the internal ligand-binding site is here and so it is called the Open end of the molecule. The other has three β-hairpin loops (L2, L4 and L6); the N-terminal polypeptide chain crosses this end of the barrel to enter strand A via a conserved 3_{10} helix affecting closure of this end of the barrel; the Closed end of the molecule. Those parts which form the three main structurally conserved regions (SCRs) of the fold, SCR1, SCR2 and SCR3, are marked as boxes. Three sequences motifs which correspond to these SCRs are shown (MOTIF 1, MOTIF 2 and MOTIF 3). The first three sequences are from kernel lipocalins and the second three from outlier lipocalins. Note that MOTIF 1 is well conserved in all sequences, whereas the other two, particularly MOTIF 2, are only well conserved in kernel lipocalin sequences (from Flower, 1996).

Figure 1.12


Open end
C
I
L1
L3
L5
L7
A
B
C
D
E
F
G
H
A1
N
3₁₀
L2
L4
L6
SCR1
SCR2
SCR3
Closed end
MOTIF1
MOTIF2
MOTIF3
AIM NFNISRIYGKWYNLAIG
RBP NFDKARFSGTWYAMAKK
BBP NFDWSNYHGKWWEVAKY
VEGP TEENQDVSGTWYLKAAA
AGP NATLDQITGKWFYIASA
OBP NETLSWLSGKWFLIAVA
AIM SYVVHTNYDEYAIFL
RBP HWIVDTDYDTYAVQY
BBP FNVLSTDNKNYIIGY
VEGP YIIPSSVEDHYIFYY
AGP LILRDTKTYMLAFDV
OBP RVLEKHGAIMLFFDL
AIM YGRAPQLRETLLQDF
RBP FSRDPNGLPPEAQKI
BBP LSRSKVLTGEAKTAV
VEGP VGRDPEINQEALEDF
AGP YADKPETTKEQLGEF
OBP SARRPDIPPELREVF

The molecular structures of four lipocalins, namely plasma retinol-binding protein, bilin-binding protein, insecticyanin and α-lactalbumin share a common structural motif that consists of a single eight stranded antiparallel β sheet folded back on itself to form a continuous hydrogen-bonded β-barrel, composed of two stacked orthogonal sheets. This together with a COOH-terminal α-helix encloses an internal ligand-binding site (Bishop *et al*, 1995; Flower, 1996). Despite the common lipocalin fold, only 25-30% amino acid sequence identity exists between lipocalins of known structure (Godovac-Zimmermann, 1988).

1.11.3 Function of lipocalins

Although the lipocalins have previously been classified as extracellular transport proteins, it is now clear that members of the lipocalin family fulfil a variety of different and potentially significant biological functions. This superfamily constitutes a phylogenetically conserved group of more than 40 proteins that function in the binding and transport of a variety of physiologically important ligands (Patel *et al.*, 1997). The lipocalins are perhaps best known for their binding of a remarkable array of small hydrophobic ligands. The structure of the lipocalin fold and a large cup-shaped cavity which forms a central hydrophobic binding pocket for the ligand, is also characteristic of the fatty-acid binding proteins (Bishop *et al.*,1995; Flower,1996). The diverse functions of lipocalins include roles in retinol transport, invertebrate cryptic coloration, olfaction, mediating pheromone activity, and prostaglandin synthesis (Flower, 1996). Lipocalins have also been implicated in enzymatic synthesis, the regulation of cell homeostasis and the modulation of the immune response, and as carrier proteins to act in the general clearance of endogenous and exogenous compounds. To date, there is no strong evidence for any real function of prokaryotic lipocalins.

1.11.4 Apolipoprotein D (ApoD)

The study by Bishop and Weiner (1996) has shown that two enterobacterial lipocalins or Blc proteins from *E. coli* and *C. freundii* share 34-39 % amino acid sequence identity with four mammalian ApoDs, whereas VlpA shares 26% with human ApoD (Barker and Manning, 1997). However, some important differences between VlpA and ApoD are apparent, such as the probable membrane association and the cysteine residues. VlpA seems to associate with the inner face of the outer membrane and there are no cysteine residues in the mature portion (Barker and Manning, 1997), while ApoD provides two disulphide bridges to maintain the β-barrel structure and form a dimer (Peitsch and Boguski, 1990).

ApoD is a glycoprotein originally characterized as a component of the HDL (high-density-lipoprotein) fraction of human plasma and known to be a member of the lipocalin superfamily (Holzfeind *et al.*, 1995). The primary structure and diverse expression of ApoD suggest that this protein is a multiligand, multifunctional glycoprotein (Yoshida *et al.*, 1996). Human ApoD is a serum glycoprotein that has no sequence similarity with other apolipoproteins but rather belongs to the α_2-microglobulin or lipocalin superfamily (Pietsch *et al.*, 1990). Most of the ApoD which is localized as a component of the human plasma HDL particle, represent a component of the lecithin : cholesterol acetyltransferase reaction (Bishop *et al.*, 1995; Flower, 1996). The mammalian plasma lipocalin, ApoD, has long been thought to function in cholesterol metabolism, although its precise physiological role remains uncertain (Flower,1996). About 80% of all plasma ApoD exists as disulphide-linked complexes predominantly with ApoA-II in HDL and plasma, and mainly with ApoB-100 in low-density lipoproteins and very-low-density lipoprotein (Flower, 1996).

Although the physiological function or the putative ligand of ApoD is unknown (Terrisse *et al.*, 1998), it has been implicated in several different functions. It has the ability

to bind phospholipids, cholesterol and other small hydrophobic molecules including sterols and steroid hormones (Suresh *et al.*, 1998). ApoD is also identified as a known progesterone / pregnenolone-binding protein. Balbin *et al.* (1990) reported that ApoD is the major protein component in cyst fluid from women with human breast gross cystic disease, and proposed a possible role of ApoD in progesterone transport in human mammary tissue.

Holzfeind *et al.* (1995) reported that ApoD might interact with meibomian lipids presents in human tear fluid and possibly contributes to the surface spreading of these lipids, or it may function as a clearance factor, protecting the cornea from harmful lipophilic molecules. Peitsch and Boguski (1990) hypothesized that heme-related compounds may be more favorable ligands for ApoD than either cholesterol or cholesteryl and they also showed that purified human ApoD binds bilirubin in an approximately one-to-one molar ratio.

ApoD may have an important function in the central nervous system (CNS), particularly in nerve regeneration. Boyles *et al.* (1990) and Spreyer *et al.* (1990) reported that ApoD was produced by astrocytes and oligodendrocytes in the central nervous system, and by neurolemmal cells and fibroblasts in the peripheral nervous system. In addition, the identification of a close homologue of ApoD has also been reported in Lazarillo, which is expressed by a subset of neurons in the developing nervous system of the American grasshopper (*Schistocera americana*) embryo (Ganfornina *et al.*, 1995; Sanchez *et al.*, 1995).

1.11.5 Lazarillo

Lazarillo is a close homologue of ApoD and is a novel cell surface glycoprotein expressed in the embryonic grasshopper nervous system (Ganfornina *et al.*, 1995; Sanchez

et al., 1995). It is anchored to the cell surface by a glycosyl-phosphotidylinositol group. Lazarillo is the first example of a lipocalin anchored to the plasma membrane, highly glycosylated, and restricted to a subset of developing neurons (Ganfornina *et al.*, 1995). The protein is necessary for the correct navigation of growing axons in the grasshopper embryo. Sanchez and colleagues (1995) suggested that Lazarillo may function as a signalling molecule mediating axonal path finding in the developing nervous system. Sequence analysis has defined Lazarillo as a member of the lipocalin family, extracellular carriers of small hydrophobic ligands, and most related to the porphyrin- and retinol-binding lipocalins. VlpA from *V. cholerae* shares 31 % amino acid identity with Lazarillo, 29% with rat ApoD and 26% with human ApoD (Barker and Manning, 1997).

1.11.6 Insect bilin-binding lipocalins

The roles for a number of invertebrate lipocalins have been identified, especially in coloration, such as the bilin-binding proteins in insects. A variety of different bile pigments, derived from heme breakdown products, are distributed widely in insects (Flower, 1996). The bile pigments are usually associated with protein as blue protein pigment in their complexes with biliverdin IXγ. Biliverdin IXγ is amongst the most common chromatophore, especially in the order Lepidoptera. The most closely related soluble lipocalins which have been studied in great detail are two insect bilin-binding lipocalins, insecticyanin (INS) from the tobacco hornworm *Manduca sexta* and bilin-binding protein (BBP) from the cabbage white butterfly *Pieris brassicae* (Flower, 1996; Holden *et al*, 1987; Huber *et al.*, 1987). The amino acid sequence of VlpA precursor shows 25% identity with bilin-binding protein from *P. brassicae* (Barker and Manning, 1997).

The first characterized insect bilin-binding lipocalin, insecticyanin, is a blue biliprotein found in the haemolymph and integument of the fifth instar larvae of butterflies

and moths. The blue colour is due to the presence of bile pigments, which are non-covalently bound to the apolipoprotein (Saito & Shimoda, 1993; Saito, 1998). Model studies of insecticyanin have shown that this protein binds a biliverdin chromophore. Interestingly, the overall three-dimensional fold of the insecticyanin subunit shows remarkable similarity to the structural motifs of bovine beta-lactoglobulin and the human serum retinol-binding protein (Holden *et al.*, 1987; Huber *et al.*, 1987). The second insect bilin-binding lipocalin, Bbp, is a blue pigment protein which is abundant in the butterfly *P. brassicae* (Schmidt & Skerra, 1994). The fold of Bbp protein is related to retinol binding protein, although the overall amino acid sequences of the proteins show only 10% homology. The ligands of bilin-binding protein and retinol-binding protein, biliverdin and retinol, respectively, are also similarly located (Huber *et al.*, 1987).

The precise functions of the insect bilin-binding lipocalins is unclear. It has been suggested that they may bind molecules and gain their colorant properties from interaction with their ligands (Flower, 1996). The protein biliverdin IXγ complexes have a role in pigmentation so they may function in photoreception and protection from photo-induced free radicals. In cyanobacteria, biliverdin-protein complexes are involved in light-harvesting (Huber *et al.*, 1987). In butterflies and moths, the pigments biliverdin IXγ are usually associated with proteins, and they contribute significantly to coloration in the epidermis or interlamellar space of the wing (Flower, 1996). Holden *et al.* (1987) reported that insecticyanin, isolated from the tobacco hornworm *M. sexta*, is involved in insect camouflage.

1.11.7 VlpA

1.11.7.1 General background

Barker and Manning (1997) compared the deduced amino acid sequence of VlpA precursor to entries in PIR and SWISS-PROT databases, using the FASTA algorithm (Pearson, 1990). Significant matches were obtained with a number of the α_2-microglobulin protein superfamily. Peitsch and Boguski (1990) have built a detailed structural model of human ApoD and then used this model to predict porphyrin-binding by ApoD. To see if VlpA might also fit into the porphyrin-group of lipocalins, a clustal V multiple alignment between the proteins was performed as shown in Figure 1.13 (Barker and Manning, 1997). Since Bbp binds a porphyrin and biliverdin IX (Huber *et al.*, 1987), the presence of residues in VlpA that are similar to residues which interact with the ligand in Bbp and IcyA raised the possibility that VlpA also binds a porphyrin.

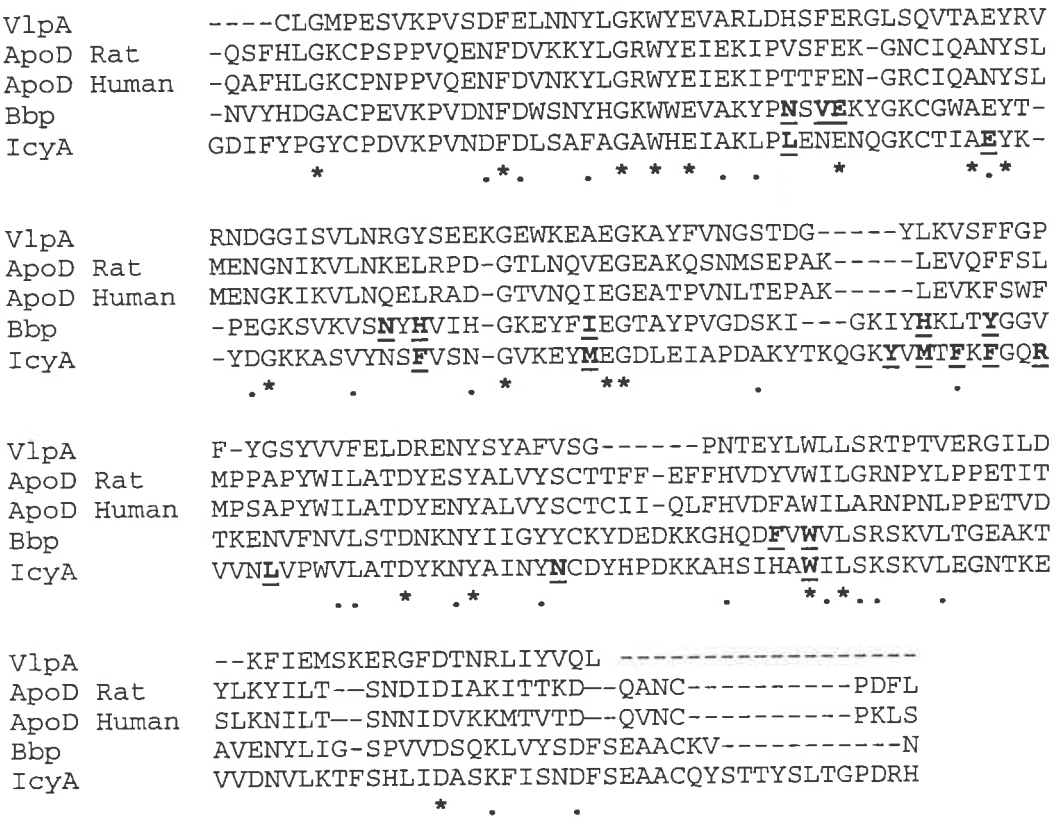
1.11.7.2 VlpA and bacterial lipocalin

Bacterial lipocalins have been reported in a restricted number of species. Lipocalins are expected to be absent in Gram-positive bacteria and those archaebacteria lacking of an outer membrane. However, no lipocalins were found in the completely sequenced genomes of the Gram-negative eubacteria *Haemophilus* and *Helicobacter* (Ganfornina *et al.*, 2000). While bacterial lipocalins have already been reported in *Citrobacter freundii*, *E. coli* and *V. cholerae* (Bishop *et al.*, 1995; Barker and Manning, 1997; GenBank accession no. U21727), four new putative lipocalins were discovered in three species of purple and green sulfur eubacteria: *Pseudomonas aeruginosa*, *Camphylobacter jejuni* and *Chlorobium tepidum*. The four new putative bacterial lipocalins showed significant similarity to the existing lipocalins (Ganfornina *et al.*, 2000).

Figure 1.13 ClustalW alignment of VlpA with porphyrin-binding-members of the α_2-microglobulin superfamily.

Identical residues are marked with an asterisk, similar residues with a dot. Residues in Bbp and IcyA that have been shown, by X-ray crystallography, to interact with the ligand are in bold type and underlined, while residues in IcyA in which the side chain lies close to the ligand are underlined. The alignment covers the mature portions of the respective proteins (from Barker and Manning, 1997).

Figure 1.13


VlpA ----CLGMPESVKPVSDFELNNYLGKWYEVARLDHSFERGLSQVTAEYRV
ApoD Rat -QSFHLGKCPSPPVQENFDVKKYLGRWYEIEKIPVSFEK-GNCIQANYSL
ApoD Human -QAFHLGKCPNPPVQENFDVNKYLGRWYEIEKIPTTFEN-GRCIQANYSL
Bbp -NVYHDGACPEVKPVDNFDWSNYHGKWWEVAKYPNSVEKYGKCGWAEYT-
IcyA GDIFYPGYCPDVKPVNDFDLSAFAGAWHEIAKLPLENENQGKCTIAEYK-
VlpA RNDGGISVLNRGYSEEKGEWKEAEGKAYFVNGSTDG-----YLKVSFFGP
ApoD Rat MENGNIKVLNKELRPD-GTLNQVEGEAKQSNMSEPAK-----LEVQFFSL
ApoD Human MENGKIKVLNQELRAD-GTVNQIEGEATPVNLTEPAK-----LEVKFSWF
Bbp -PEGKSVKVSNYHVIH-GKEYFIEGTAYPVGDSKI---GKIYHKLTYGGV
IcyA -YDGKKASVYNSFVSN-GVKEYMEGDLEIAPDAKYTKQGKYVMTFKFGQR
VlpA F-YGSYVVFELDRENYSYAFVSG------PNTEYLWLLSRTPTVERGILD
ApoD Rat MPPAPYWILATDYESYALVYSCTTFF-EFFHVDYVWILGRNPYLPPETIT
ApoD Human MPSAPYWILATDYENYALVYSCTCII-QLFHVDFAWILARNPNLPPETVD
Bbp TKENVFNVLSTDNKNYIIGYYCKYDEDKKGHQDFVWVLSRSKVLTGEAKT
IcyA VVNLVPWVLATDYKNYAINYNCDYHPDKKAHSIHAWILSKSKVLEGNTKE
VlpA --KFIEMSKERGFDTNRLIYVQL ----------
ApoD Rat YLKYILT--SNDIDIAKITTKD--QANC----------PDFL
ApoD Human SLKNILT--SNNIDVKKMTVTD--QVNC----------PKLS
Bbp AVENYLIG-SPVVDSQKLVYSDFSEAACKV----------N
IcyA VVDNVLKTFSHLIDASKFISNDFSEAACQYSTTYSLTGPDRH

In addition to VlpA, another bacterial lipocalin which has been examined is Blc from *E. coli*. Blc is an outer membrane protein like VlpA whose processing is globomycin-sensitive. Both VlpA and Blc can be labeled with [H^3] palmitate, suggesting that they are lipoproteins (Bishop *et al.*, 1995; Barker and Manning, 1997). Bishop *et al.* (1995) predicted that the mature protein of Blc is 159 residues (18,043-Da), while a mature protein of VlpA is 18 kDa (Barker and Manning, 1997). Analysis by Flower *et al.* (1991; 1993) suggested VlpA to be an outlier lipocalin, lacking the second characteristic lipocalin sequence motif. Blc contains a short deletion of 7 amino acid residues corresponding to a hydrophobic surface loop, which is thought to facilitate the physical interaction between ApoD and high density lipoprotein.

Based on the hydrophobic ligand binding capacity of ApoD, it was proposed that Blc also functions in a similar manner (Bishop *et al.*, 1995). The *blc* gene was demonstrated to be controlled by *rpoS*, which controls a program of gene expression induced under starvation conditions and at the onset of stationary phase (Kaasen *et al.*, 1992). Therefore it was suggested that Blc may function in an important starvation response function in *E. coli*. However, Condon and Weiner (1988) deleted the chromosomal copy of the *frd* operon in *E. coli* MI1443 encompassing the *blc* locus. This strain is capable of growth under both anaerobic and aerobic (Bishop, 1993) conditions, indicating that the *blc* gene is non-essential in *E. coli*.

Barker and Manning (1997) over-expressed *V. cholerae* VlpA in *E. coli* and demonstrated that this strain could bind haemin with low affinity, and possibly non-specifically suggesting that VlpA may be involved in binding of haemin. In addition, VlpA appears to mediate binding to the related compounds haematoporphyrin, protoporphyrin IX, and Congo red. Attempts to introduce mutations involving *vlpA* in *V. cholerae* strains were unsuccessful presumably due to both its chromosomal organization and the multiple

copies of *vlpA*. The activity of VlpA as a weak haemin binding protein may indicate a role for the protein in iron uptake (Barker and Manning, 1997).

1.11.7.3 Structure of bacterial and eukaryotic lipocalins

Flower and colleagues (1991, 1993a) have analyzed the prokaryotic lipocalin structures which display characteristic lipocalin sequence motifs, although the greatest similarity is to a subset of the lipocalin family that includes insecticyanin, Lazarillo, crustacyanin, and ApoD (Flower *et al.*, 1995b). Bishop and Weiner (1996) reported that the biochemically distinguishing feature between bacterial and eukaryotic lipocalins, i.e. ApoD and Lazarillo, is their peripheral membrane location. Blc is anchored to the outer membrane of the Gram-negative cell envelope by a lipid-modified amino-terminal cysteine residue (Bishop and Weiner, 1996), while ApoD associates with plasma lipoproteins by a solvent-exposed hydrophobic loop on the protein surface (Peitsch and Boguski, 1990; Yang *et al.*, 1994), and Lazarillo is anchored to membranes by a glycosyl-phosphatidylinositol (GPI) anchor at the carboxyl terminus (Ganfornina *et al.*, 1995).

When the two bacterial lipocalins are compared, VlpA lacks cysteine residues in the mature portion of the protein (Flower *et al.*, 1995b) while Blc in *E. coli* exhibits a consensus prokaryotic lipoprotein cleavage site (Braun, 1993). In addition, Barker (1993) predicted the putative role of VlpA may be providing a membrane anchor for the associated adhesins which has been proposed for TcpC and the TCP fimbrial adhesin. However, VlpA is not required for anchoring of the MFRHA (Barker, 1993).

1.11.7.4 Study of bacterial lipocalin

To date, very little is known about bacterial lipocalins, and the particular functional specialization and the possible roles are still undefined. Since no lipocalins have yet been

identified from an archaeon, the possibility remains that bacterial lipocalins were acquired by horizontal transmission from a eukaryote (Bishop and Weiner, 1998). Conversely, a biased %G + C in the first and third codon positions provides no support for this hypothesis (Ganfornina *et al.*, 2000; Lawrence and Ochman, 1997).

Barker and Manning (1998) have suggested that the bacterial lipocalin gene itself has probably been pilfered from another organism, possibly a member of the *Enterobacteriaceae*. Bishop *et al.* (1995) supported two hypotheses: the first being that the primordial lipocalin originated in bacteria, and served a function in the Gram-negative cell envelope before it was acquired by eukaryotes and adapted to a number of functions in multicellular organisms. Alternatively, the lipocalins originated in eukaryotes from where they adapted to function in the Gram-negative cell envelope. However, the suggestion that the lipocalin protein family is far more ancient than has been generally assumed and should prompt a renewed search for lipocalins in organisms from all branches of the tree of life.

1.12 Aims of this study

VlpA is an outer-membrane protein that belongs to the lipocalin superfamily. There is no evidence for the function of this prokaryotic lipocalin. A preliminary characterization of the *vlpA* gene was undertaken by Barker (1993) in which one copy of the structural gene from *V. cholerae* 569B was cloned and sequenced. However, the construction of specific *vlpA* mutants in *V. cholerae* was unsuccessful. Therefore, further studies attempts using the mutants to determine the role of VlpA in the pathogenesis of cholera were not carried out. Further analysis by Barker (1993) revealed that *vlpA* is present as multiple copies in the *V. cholerae* genome and always associated with a VCR which themselves were found to be significantly related to the consensus sequences of 59 - base element (Recchia and Hall, 1997; Mazel *et al.*, 1998; Manning *et al.*, 1999). It remains to be establish experimentally

whether *vlpA* is a mobile cassette gene and VCRs can function as integrase-mediated recombination sites.

The *V. cholerae intI*4 gene encoding the integrase IntI4 has also recently been cloned and expressed in *E coli* (C. Clark, personal communication). However, the chromosome insertion and/or deletion of gene cassettes in *V. cholerae* catalyzed by IntI4 has not till now been demonstrated. The role of IntI4 in integrase mediated movement of the *vlpA* gene cassette containing VCR elements will be investigated in this study. Insertionally inactivated *vlpA* genes using the antibiotic resistance cartridges as markers will be used to detect the deletion of this gene cassette both in *E. coli* as a cloned plasmid encoded copy and in *V. cholerae* where it naturally occurs on the chromosome.

The objective of this study is to address the above issues and also to insertionally inactivate the copies of the *V. cholerae vlpA* genes and examine their role in this organism virulence, using the infant mouse model. Finally, an analysis of the expression of the genes encoding VlpA including the promoter region of the *vlpA* genes will be undertaken.

Chapter 2

Materials and Methods

2.1 Bacterial strains

Bacterial strains used in this study are *Vibrio spp.*, *Aeromonas spp.* and *E. coli* strains. All strains are listed in Table 2.1 to Table 2.6

Table 2.1 *V. cholerae* O1 strains used in this study.

Lab. strain numbers [a]	strains	biotype	serotype	source[b] (year of isolation)
V30	569B	classical	Inaba	a 1946
V35	CA 401	classical	Inaba	b 1953
V584	Z17561	classical	Inaba	c 1985
V32	CA411	classical	Ogawa	b 1953
V585	AA14041	classical	Ogawa	c 1985
V586	AA13993	El Tor	Inaba	c 1985
V587	AA14073	El Tor	Ogawa	c 1985
V33	O17 SR	El Tor	Ogawa	a pre-1965
V777	C5	El Tor	Ogawa	d 1957
V793	C31	El Tor	Ogawa	d 1957

[a] numbers used in this laboratory

[b] sources are :

a = Dr. K. Bhaskaran (Central Drug Institute, Lucknow, India); b = Dr. J. Berry (University of Texas, Austin, Texas, USA); c = Dr. B. Kay [International Center for Diarrhoeal Disease Research, Dhaka, Bangladesh (I.C.D.D.R.B.)]; d = Dr. V. Franzon (University of Adelaide, Adelaide, South Australia).

Table 2.2 *V. cholerae* O139 strains used in this study.

Lab. strain numbers [a]	strains	source [b] (year of isolation)
V911	AI-1837	a
V912	AI-1838	a
V913	AI-4450	a
V914	AI-1841	a
V915	AI-1852	a
V916	AI-1854	a
V917	AI-1855	a
V918	AI-4260	a
V921	Arg 3	a
V961	AI-4260B no1	c
V962	AI-4260B no2	c
V996	AJ-18360	a
V997	232/95	a
V998	X-004	d
V999	MDO-25	e

[a] numbers use in this laboratory

[b] sources are :

a = Dr. M. John Albert (I.C.D.D.R.B.); b = Prof. Paul A. Manning (Strain was from Argentina); c = Dr. Gunhild Jonson; d = Dr. M. John Albert (Strain was from China); e = John Albert (Strain was from India)

Table 2.3 *V. cholerae* non-O1 strains used in this study.

strains	source
V. cholerae O2 to *V. cholerae* O155	Dr. M. John Albert (I.C.D.D.R.B.)

Table 2.4 *Vibrio spp.* strains used in this study.

Lab.strain numbers [a]	species	strains	source [b]
V943	*V. anguillarum*	ATCC43305	a
V944	*V. anguillarum*	ATCC43306	a
V947	*V. anguillarum*	ATCC43307	a
V948	*V. anguillarum*	ATCC43308	a
V949	*V. anguillarum*	ATCC43309	a
V950	*V. anguillarum*	ATCC43310	a
V951	*V. anguillarum*	ATCC43311	a
V952	*V. anguillarum*	ATCC43312	a
V953	*V. anguillarum*	ATCC43313	a
V954	*V. anguillarum*	ATCC43314	a
V957	*V. anguillarum*	85-3954-1	a
V958	*V. anguillarum*	85-3954-2	a
V959	*V. anguillarum*	86-3674	a
V960	*V. anguillarum*	89-3748-1	a
V563	*V. mimicus*	non-O1	b
V800	*V. mimicus*	non-O1	c
V128	*V. mimicus*	M-33	b
V1209	*V. mimicus*	M-35	b
V1210	*V. mimicus*	N-4459	b
V691	*V. parahaemolyticus*	102	d
V1211	*V. parahaemolyticus*	AA-3853	b

Table 2.4 *Vibrio spp.* strains used in this study (continued).

Lab.strain numbers [a]	**species**	**strains**	**source** [b]
V1212	*V. parahaemolyticus*	X-4844	b
V1213	*V. parahaemolyticus*	Y-17233	b
V1213	*V. parahaemolyticus*	Y-17233	b
V564	*V. fluvialis*	non-O1	c
V1214	*V. fluvialis*	AA-15385	b
V1215	*V. fluvialis*	AL-1577	b
V1216	*V. fluvialis*	AL-14413	b

[a] numbers use in this laboratory

[b] sources :

a = Dr. Peter Hanna (Deakin University); b = Dr. M. John Albert (I.C.D.D.R.B.); c = Dr. Steve Attridge (The University of Adelaide, Adelaide, South Australia); d = Dr. Dharam Sharma (1988).

Table 2.5 *Aeromonas spp.* strains used in this study.

Lab. strain numbers [a]	species	strains	source [b]
V994	*A. trota*	1203	a
V995	*A. trota*	1354	a
E1129	*A. sobria*	A187	b
E1130	*A. sobria*	A191	b
E1131	*A. caviae*	A321	b
E1133	*A. caviae*	V14	b
E1128	*A. hydrophila*	A006	b
E1132	*A. hydrophila*	A335	b
E1324	*A. hydrophila*	AE-18	b
E1325	*A. hydrophila*	Th4	b
E1326	*A. hydrophila*	AH-1	b
E1327	*A. hydrophila*	AH-2	b

[a] numbers use in this laboratory

[b] sources are :

a = Dr. John Albert (I.C.D.D.R.B.); b = Dr. S. Kirov (University of Tasmania, Australia).

Table 2.6 *E. coli* strains used in this study.

strains	genotype	source
DH5α	*F⁻,ϕ80d lacZ ΔM15, Δ(lacZYA-argFU169, deoR, recA1, endA1, hsdR17 (r_k^-, m_k^+) supE44, thi1, gyrA96, relA1,λ⁻*	a
S17-1λ pir	A lysogen of S17-1[*pro hsdR RP4-2-Tc:: Mu Km::Tn7*] used for conjugal transfer of plasmids with R6K replicon.	b

sources are :

a = Bethesda Research Laboratories; b = de Lorenzo *et al.* (1990).

2.2. Growth media

V. cholerae, *V. mimicus*, and *V. fluvialis* strains were cultivated in Nutrient Broth (NB) (Oxoid) consisting of Bacto-peptone (100g/l), Lab Lemco (10 g/l) (Oxoid) and NaCl (5 g/l). *V. parahaemolyticus* strains were grown in NB containing NaCl at 20 g/l, while *V. anguillarum* and *Aeromonas spp.* strains required Trypticase Soy Broth (TSB) (BBL Microbiology Systems) at 30°C. Luria Broth (LB) was prepared as described by Miller (1972) and consisted of Bacto-tryptone (10 g/l), yeast extract (5 g/l) (Difco), and NaCl (5g/l). The pH was adjusted to 6.5 prior to autoclaving. M9 Minimal medium consisted of M9 salts [$Na_2HPO_42H_2O$ (75.2g/l), KH_2PO_4 (10g/l), and NaCl (5 g/l)], supplemented with filter sterilized $MgSO_4$ (0.1 mg/ml), glucose (5 mg/ml), and thiamine (Vitamin B) (100μg/ml) prior to use. AKI medium was used to assess TcpA/TCP and cholera toxin producing strains (Jonson *et al.*, 1991a). AKI medium was prepared fresh for each experiment as described (Iwanaga *et al.*, 1986). This was consisted of NaCl (5 g/l), Bacto-peptone (Difco) (15 g/l), yeast extract (4 g/l) and 20 ml of 7.5% $NaHCO_3$ (filter sterilized and prepared fresh prior to use). The cultures were gassed with 5% CO_2 prior to incubation

(Voss and Attridge, 1993), with incubations standing for 4 hours at 30°C, followed by 16 hours at 30°C with vigorous shaking (250 oscillations/min; Orbital shaking water bath, Paton industries). Solid media, i.e. Nutrient Agar (NA), Trypticase Soy Agar (TSA), and Minimal medium Agar were made from NB, TSB, and M9 Minimal medium, respectively, by solidification with Bactoagar (15 g/l) (Difco). Thiosulphate-citrate-bile salt-sucrose agar (TCBS) (Oxoid) was used to select for *V. cholerae* strains when necessary. Isopropyl-β-D-thiogalacto-pyranoside (IPTG) (24 μg/ml) and 5-bromo-4-chloro-3-indolyl-β-D-galcto-pyronaside (X-gal) (20 μg/ml) (purchased from Boehringer Mannheim) were used to supplement the solid media when required.

Antibiotics were added to broth and solid media at the following final concentrations : ampicillin (Ap) 50 μg/ml, chloramphenicol (Cm) 25 μg/ml, kanamycin (Km) 50 μg/ml, tetracyclin (Tet) 4 μg/ml for *V. cholerae* and 10 μg/ml for *E. coli*, and rifampicin (Rif) 200 μg/ml.

Incubations were at 37°C unless otherwise specified. Normally, liquid cultures were grown in 20 ml McCartney bottles or 100ml conical flasks. Culture densities were measured at 600nm (A_{600}) using a Pharmacia LKB-Ultrospec Plus spectrophotometer.

2.3. Maintenance of bacterial strains

Strains were maintained as lyophilized cultures stored *in vacuo* in sealed glass ampoules for long term storage. The working stocks were maintained at -70°C in 1% Bacto-peptone (Difco) containing 30% (v/v) glycerol. When required, a loop-full of the glycerol suspension was streaked onto NA and incubated overnight at 37°C . Unless indicated, bacteria were cultured in NB with aeration at 37°C. Strains from ampoules were opened and their contents suspended in several drops of the appropriate sterile broth. Half

of the contents were then transferred to a 10 ml bottle of NB and grown with shaking overnight. The other half was streaked onto two NA plates and incubated overnight at the appropriate temperature.

Bacterial strains were prepared for long term storage by suspension of several colonies in 300 µl of 10% sterile skimmed milk solution in Milli-Q water. Approximately 100 µl aliquots of this thick bacterial suspension were dispensed into sterile 0.25 x 4 inch freeze drying ampoules and the end of each ampoule was plugged with cotton wool. The cotton wool plugs were pushed well down the ampoules and a constriction was made just above the level of the plug. The ampoules were then lyophilized in a freezer dryer at -50°C (Dynavac engineering). After the vacuum was released, the ampoules were evacuated to a partial pressures of 30 microns and then sealed at the constriction without releasing the vacuum. Finally, the labelled ampoules were stored at 4°C.

2.4. Chemical and reagents

All chemicals used were analytical grade. Chemicals purchased from BDH were : ethanol, methanol, propane-2-ol (iso-propyl-alcohol), iso-amyl-alcohol, hydrochloric acid, acetic acid, sodium chloride, ammonium acetate, sodium hydroxide, caesium chloride, phenol, polyethylene glycol (PEG), sodium dodecyl sulphate (SDS), formaldehyde, formamide, ethylene-diamine-tetra acetic acid disodium salt (EDTA), iso-octylphenoxypolyethoxyethanol (Triton X-100), chloroform, sucrose and glucose.

The four deoxyribonucleotide triphosphates (dATP, dCTP, dGTP, and dTTP), IPTG, X-gal, and 5-bromo-4-chloro-3-indolyl phosphate (X-pho) were from Boehringer Mannheim. Tris tris [hydroxy-methyl] amino-methane, Tween 20, 3-[N-Morpholino] propanesulfonic acid (MOPs), adenosine-5′-triphosphate (ATP), guanidine thiocyanate,

and diethyl pyrocarbonate (DEPC) were purchased from Sigma. Magnesium chloride, magnesium sulphate, calcium chloride, potassium chloride, and glycerol were from Ajax. Glycine was from Amresco. Herring sperm DNA and bovine serum albumin (fraction V) (BSA) were from Ciba-Geigy. Ficoll (Type 400) was obtained from Pharmacia.

Digoxigenin (DIG) DNA labelling and detection kits were purchased from Boehringer-Mannheim. Other chemicals were from Ajax, Sigma, or BDH. Milli-Q water was purified using a Milli-Q water purification system (Millipore Corp.) with a measured resistance to conductivity of 18MΩ/cm. Sterile, deionized and filtered water was used to prepare all buffers and reagents for DNA and RNA manipulation. All other reagents and buffers were prepared with deionized water.

Antibiotics, i.e. Ap, Km, Rif were purchased from Sigma, and Cm, Tet from Calbiochem. All other antimicrobial agents (dyes, detergents and antibiotics) were purchased from Sigma, BDH, Glaxo, or Calbiochem. Electrophoresis grade reagents were purchased from Boehringer-Mannheim, Bio-Rad, Bethesda Research Laboratories, and FMC Corp. Acrylamide and ultra pure N-N′-methylene bis-acrylamide were from Boehringer-Mannheim. Ammonium persulphate and pulsed field certified agarose were from Bio-Rad. Ultrapure™ low gelling temperature (LGT) agarose was from Bethesda Research Laboratories. High gelling temperature (HGT) agarose (Seakem), and N, N, N′, N′-tetramethyl ethylenediamine (TEMED) were from FMC Corp.

2.5. Enzymes and Immunoconjugates

Deoxyribonuclease I was obtained from Geneworks. Ribonuclease A (RnaseA) and lysozyme were from Sigma. Glycogen, pronase, and proteinase were from Boehringer-Mannheim. All restriction endonucleases were purchased from either Boehringer-

Mannheim, New England Biolabs, Pharmacia or Amersham and used according to the suppliers instructions.

Other DNA modifying enzymes were purchased from the following suppliers. T4 DNA ligase was from New England Biolabs, while Klenow fragment of DNA polymerase I, molecular biology grade alkaline phosphatase and terminal transferase were from Boehringer-Mannheim. Taq polymerase (Ampli-Taq) was purchased from Perkin Elmer Cetus Corp.

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

2.6. Animals

Infant Swiss mice weighing between 2.4-2.7 g from the Central Animal House of the University of Adelaide were used for *in vivo* studies of *V. cholerae* strains. Antisera were raised in outbred rabbits from the same source.

2.7. Bacterial plasmids and cloning vectors

Bacterial plasmids used in this study are listed in Table 2.7.

Table 2.7 Plasmids used in this study.

Plasmids	Description / Relevant phenotypes	References or sources
pACYC184	Tet^R Cm^R	Chang and Cohen (1972)
pBAD24	Ap^R cloning vector, arabinose promoter	Guzman *et al.*(1995)
pBBR1MCS-4	Ap^R *lacZ*	Kovach *et al.* (1995)
pBluescript II KS+/-	Ap^R *lacZ*	Stratagene
pBluescript II SK+/-	Ap^R *lacZ*	Stratagene
pBSL121	cloning vector, Cm^R Ap^R	Takeshita *et al.* (1987) and Bartolomé *et al.* (1990).
PCVD442	Suicide vector with Pir-dependent RK6 replicon. Contain *sacB* gene for positive selection of recombinants (Ap^R).	Donnenberg and Kaper (1991)
pFLAG-1™	Ap^R cloning vector, *tac* promoter	International Biotechnologies
pGEM5Zf(+)	Ap^R *lacZ*	Promega
pGEM-T	pGEM5Zf(+) linearized with T-overhang for direct cloning of PCR products (Ap^R *lacZ*).	Promega
pGEM-T Easy	pGEM-T Easy linearized with T-overhang for direct cloning of PCR products (Ap^R *lacZ*).	Promega
pGP1-2	Km^R T7 *pol* gene under heat inducible promoter	Tabor and Richardson (1985)

Table 2.7 Plasmids used in this study (continued).

Plasmids	Description / Relevant phenotypes	References or sources
pPM471	6.2-kb *Bam*HI fragment contains *mrhA,B* genes and *vlpA* gene inserted into pBluescript II KS+ (Ap^R).	Franzon (1988)
pRMH232	*Sph*I-*Hin*dIII fragment of In6 (pSa) inserted into pACYC184 (Cm^R Int^-)	Collis *et al.* (1993)
pRMH313	263-bp *Taq*I fragment included *attI*1 from Tn*21* cloned into pACYC184 at *Eco*RV and *Taq*I sites	Dr Christina Collis, CSIRO Division of Biomolecular Engineering, Sydney Australia
pSU23	cloning vector, Cm^R *lacZ*	Bartolomé *et al.* (1990)
pSU2056	1.176-kb *Bam*HI-*Rsa*I fragment of In2 from Tn*21* inserted into pUC9 (Ap^R Int^+).	Martinez and de la Cruz (1990)
pUC18/19	Ap^R *lacZ*	New England Biolabs

2.8 Oligodeoxynucleotides

Oligodeoxynucleotides primers used in this study were purchased from GeneWorks, except those refer to the references were synthesized from the Sequencing Center of the Institute of Medical and Veterinary Science (I.M.V.S.) (Adelaide, South Australia), and are described in Table 2.8.

Table 2.8 Oligodeoxynucleotides primers used in this study.

Numbers	Primer sequences (5′→3′)	References
#922	CCCCTTAGGCGGGCGTTA	Franzon (1988)
#923	CCCTCTTGAGGCGTTTGTTA	Franzon (1988)
#1009	CACACGCAAACTCAAGGA	Ogierman (1998)
#1019	CCTTTGAAGTAGTATGGC	Ogierman (1998)
#1071	TTTGGACATATGATGAGAGCTATC	Barker (1993)
#1072	GTTAGGTATTTATTGCAGCTG	Barker (1993)
#2265	GGCATTTTGGTTCTTATTCG	This study
#2266	GGGAAACTCAGCATTCAG	This study
#2267	AAACTCCAGCCAAAAGTC	This study
#2268	GAAAAATGGCGGTCATATC	This study
#2272	GACTTTCAAGCAAATGTG	This study
#2273	CGACACAATAGATTCTCC	This study
#2277	GGCTCATTTACGTTCACG	This study
#2385	AGAAAGGCAGGCAAAATTAC	This study
#2386	GTTCATTTTTTCTACCTAGTAC	This study
#2421	GGACAAGTTCATAGAAATGTCGAAAGAGCGTGG	This study
#2422	GGTTTTACTGATTCGGGCATGCCCAAGCAGCCA	This study

Table 2.8 Oligodeoxynucleotides primers used in this study (continued).

Numbers	Primer sequences (5′→3′)	References
#2476	AGATATGTGTTCACAGAACTCGCTTATATCGTG	This study
#2477	CCCGAGCAATAAAGCTGTGACATCTGAATAACC	This study
#2478	CGCTTGGTCAGCGATGGGAGTTATTAGACACTA	This study
#2479	GCCCTACCCCAACACCAGTGCTAACATTCGTAC	This study
#2480	GGTTTTGGTGGGTGCGTCACCGCAAGTTTATGC	This study
#2481	CCTGGGGTTCTTAGGAATACACTTTCTAAGTGC	This study
#2483	GCGCTTGAGAGAACTCGG	This study
#2484	AAGTTGTCATGTAAGCGG	This study
#2495	GGCACTGAACATTGAACCTAC	This study
#2679	AATTTAATCGCGGCCTCGAGC	This study
#2680	GCAGTTTCATTTGATGCTCGA	This study
#2681	CCAGTGTCGGATTTTGAACTG	This study
#2682	GCCGATTTGTATCAAAACCAC	This study
#2816	CTTCCTTCCACTCACCTT	This study
#2817	GAATGGCTCAACAGATGG	This study
#2904	GGCAATAGCCAGTAGCTCAAG	This study
#2905	GGTGGATCAGTGATGAACAGA	This study
#2926	TAATACGACTCACTATAGGG	This study
#2993	TTAGGTGACACTATAGAATAC	This study
#3082	CAACTGGTCCACCTACAACAA	This study

Table 2.8 Oligodeoxynucleotides primers used in this study (continued).

Numbers	Primer sequences (5′→3′)	References
#3092	GATAGCTCTCATCATATGTCCAAA	This study
#3093	CAGCTGCAATAAATACCTAAC	This study
#3141	GATTCCGAGAGCTTGGCTTAC	This study
#3142	TTAGGCTAGTTATCCACTACC	This study
#3148	CAATGCGCTCATCGTCATCCT	This study
#3162	ACGCCGAAACAAGCGCTCATG	This study
#3164	GAGTTAGCTCACTCATTAGGC	This study
#3165	TTCGCTATTACGCCAGCTGGC	This study
#3198	GATCCATTCCCACTGTAGCAA	This study
#3199	AGGAGCATCTTTCACACGTGC	This study
#3200	CGCTCTGTAATAGCGTGTTAT	This study
#3201	GTGAGAACTATATGAAATCCC	This study

2.9. Transformation and electroporation

2.9.1 Transformation of *E. coli*

2.9.1.1 Preparation of competent cells

Competent cells used in transferring plasmids between *E. coli* K-12 strains were performed essentially according to the method described by Brown *et al.* (1979). *E. coli* strains were grown shaking in NB overnight, then diluted 1:20 into NB and incubated with shaking until an A_{600} = 0.6. The cells were chilled on ice for 20 min, pelleted at 4°C in a bench centrifuge, resuspended in 0.5 volume of ice-cold 100 mM $MgCl_2$, centrifuged again and resuspended in 0.5 volume of ice-cold 100 mM $CaCl_2$. The competent cells were

allowed to stand on ice for 60 min before adding of 1ml ice-cold 100 mM $CaCl_2$, 60% (v/v) glycerol. Cells were then snap frozen in 300 µl aliquots and stored at -80°C.

2.9.1.2 Transformation procedure

Competent cells were thawed on ice and then 0.1 ml of cells were mixed with 1.2 µg of DNA [made to a volume of 10 µl with 10 mM Tris-HCl pH 8.0, 1 mM EDTA] and incubated on ice for 30 min. The mixture was heat-shocked at 42°C for 90 sec, returned to ice for 1-2 min and 1 ml of NB was added to the transformation mixture, followed by incubation at 37°C for 1-2 hours. The mixture was plated directly onto selection plates.

2.9.2 Electroporation of *E.coli* and *V. cholerae*

2.9.2.1 Preparation of electrocompetent cells

Electrocompetent *E. coli* DH5α cells were freshly prepared according to the Bio-Rad protocol. Briefly, 100 ml of LB was inoculated with 1/100 volume of a fresh overnight culture and cells were grown at 37°C with vigorous shaking to an A_{600} of 0.5-0.8. Before harvesting, the cells were transferred to sterile SS-34 tubes (Nalgene), allowed to chill on ice for 15-30 min, and centrifuged at 10,000 rpm for 7 min at 2°C (JA-20 rotor, Beckman). The supernatant was discarded and the cells gently resuspended in 100 ml of ice-cold, sterile, distilled water (dH_2O). The cells were subjected to two more washes; the first in 50 ml ice-cold dH_2O and a final wash in 2 ml of ice-cold 10% glycerol. Finally, the cell pellet was resuspended to a final volume of 200-300 µl in ice-cold 10% glycerol. The suspension was frozen in aliquots on dry ice, and stored at -70°C. The cells are good for at least 6 months under these conditions, and were retained for no longer than this period of time.

Preparation of electrocompetent *V. cholerae* cells was essentially similar to the procedure described for *E. coli*. Strains were grown overnight on NA and a single colony was inoculated into a 20 ml LB, then grown to an A_{600} = 0.5-0.8. The exception for *V. cholerae* is that the cells were first washed with an equal volume of ice-cold sterile 1 mM $MgCl_2$, then in ice-cold sterile 1 mM $CaCl_2$ solution, followed by a wash with a sterile, ice-cold 0.5 mM $MgCl_2$, then in 0.5 mM $CaCl_2$ solution. *V. cholerae* cells were finally resuspended in 1 ml of the latter solution and kept on ice.

2.9.2.2 Electroporation procedure

The plasmid DNA or PCR product in TE or sterile Milli-Q water were mixed with electrocompetent *E. coli* or *V. cholerae* at a ratio 1:10 (v/v) in a sterile microfuge tube on ice. The mixture was transferred to an ice-cold sterile Gene Pulser™ cuvette (electrode gap of 0.2 cm, Bio-Rad). The Gene Pulser (Bio-Rad) was set at 25 μF capacitance, 200Ω resistance and electroporation was performed at 2000V or 2500V for *E. coli* and *V. cholerae* respectively, with the time constant in the range of 4.6-4.7 msec. The suspension was immediately diluted in 1 ml of LB, in order to maximize cell survival (Marcus *et al.*, 1990), transferred to a sterile microfuge tube, and incubated at 37°C for 60-90 min. After centrifugation at 15,000 rpm for 1-2 min, the supernatant was discarded, the pellet gently resuspended in 400 μl of sterile saline, and plated onto selection plates containing appropriate antibiotics.

2.10. Bacterial conjugation

Overnight broth cultures grown in NB or LB were diluted 1:20 and grown at 37°C shaking to early exponential phase (A_{600} = 0.5). Excess antibiotics were washed from cells

before conjugation with 10 ml LB twice by centrifugation at 4500 rpm for 10 min at RT, and resuspended in 10 ml LB. Donor and recipient bacteria were mixed at a ratio of 1:10 by volume and the cells pelleted by centrifugation for 10 min, 5000 rpm, in a bench centrifuge. The pellet was gently resuspended in 0.1 ml of LB and spread onto a cellulose acetate membrane filter (0.45 mm, type HA, Millipore Corp.) on a NA plate. This plate was incubated for 4-6 hours at 37°C, and the bacteria on the filter were collected by vortexing the filter in 10 ml NB. After removing the filter, serial dilutions of cell suspension were plated onto selective agar and incubated at the appropriate temperature.

2.11. DNA extraction procedures

2.11.1 Small scale plasmid isolation

Small scale plasmid isolation procedures were employed using a modified alkaline/SDS method to prepare the clear lysate. Overnight bacterial cultures (10 ml) were pelleted by centrifugation for 10 min at 5000 rpm at room temperature. The pellets were resuspended in 300 µl solution P1 (50 mM Tris, 10 mM EDTA, adjusted to pH 8.0 with HCl, and RNase added to 100 µg/ml). Cells were lysed by the addition of 300 µl solution P2 [200mM NaOH, 1.0% SDS (w/w)] and mixed gently, but thoroughly, until the lysate appeared to be homogeneous. After 5 min incubation at RT, the mixture was neutralized by the addition of 300 µl solution P3 (3.2M potassium acetate, adjusted with acetic acid to pH 5.0) and mixed immediately by inverting the tube 5 times, and left on ice for 10-15 min. The cellular debris was removed by centrifugation 15000 rpm for 15 min, at 4°C. The supernatant was extracted twice with 400 µl of chloroform. Plasmid DNA was then precipitated by the addition of 700 µl of iso-propanol and incubated on ice for 30 min. DNA was collected by centrifugation (15 min, 15000 rpm, 4°C), washed with 70% ethanol,

and dried at 65°C. The DNA pellet was resuspended in Milli-Q water or 10 mM Tris-HCl pH 8.0, 1 mM EDTA and stored at 4°C.

2.11.2 Large scale plasmid isolation

Large scale plasmid isolation was performed by the three step alkaline lysis method (Garger *et al.*, 1983). Cells from a one litre culture were harvested (7000 rpm, 10 min, 4°C, GS-3, Sorvall) and resuspended in 24 ml of solution 1 (50 mM glucose, 25 mM Tris HCl pH 8.0, 10 mM EDTA). Freshly prepared lysozyme (4 ml of 20 mg/ml in solution 1) was mixed with the cell suspension and incubated at RT for 10 min. Total cell lysis was achieved by the addition of 55 ml of solution 2 [0.2M NaOH, 1% (w/w) SDS], swirled gently, followed by a 5 min incubation on ice. After the addition of 28 ml solution 3 (50 mM potassium acetate, 11.5% glacial acetic acid) the mixture was incubated on ice for 15 min, and centrifuged at 8000 rpm, 4°C, for 20 min (GSA rotor, Sorvall) to remove the cellular debris. The supernatant was extracted with an equal volume of Tris-saturated phenol (pH 7.5), chloroform, isoamyl alcohol (25 : 24 : 1) mixture. Plasmid DNA from the aqueous phase was precipitated by the addition of 0.6 volume of iso-propanol, incubated at room temperature for 10 min and collected by centrifugation (11000 rpm, 20min, 4°C, GSA rotor, Sorvall). The DNA pellet was washed with 70% (v/v) ethanol, dried briefly *in vacuo* and resuspended in 4.8 ml 10 mM Tris-HCl pH 8.0, 1 mM EDTA (TE). CsCl was added to a density of 1.8 g/ml, followed by 0.8 ml of 10 mg/ml ethidium bromide and the mixture overlaid with 8 ml CsCl in 10 ml TE (density of 1.47 g/ml) in two 5/8 x 3 inch Beckman Quickseal polyallomer tubes. The tubes were filled with the 1.47 g/ml CsCl solution, sealed, and centrifuged (65000 rpm, 5 hours, 20°C, Ti-80, Beckman L8-80). After centrifugation the DNA band was removed by side puncture of the tubes with a 19 gauge

needle attached to a 1 ml syringe. The ethidium bromide was removed by several extractions using isoamyl alcohol, while CsCl was removed by dialysis against three changes of 2 litres 10 mM TE at 4°C. DNA was precipitated by the addition of 0.1 volume 3M sodium acetate pH 5.2 and an equal volume of iso-propanol, followed by incubation on ice for 10 min. DNA was collected by centrifugation (15 min, microfuge), washed with 70% (v/v) ethanol, dried *in vacuo* , resuspended at a concentration of 1 µg/µl in 10 mM Tris-HCl pH 8.0, 1 mM EDTA and stored at 4°C.

2.11.3 Preparation of bacterial genomic DNA

Genomic DNA from *Vibrio spp.* and *Aeromonas spp.* were prepared according to the method of Manning *et al.* (1986). Cells from a 20 ml shaken overnight culture were pelleted in a bench centrifuge for 10 min and resuspended in 2 ml 25% sucrose, 50 mM Tris-HCl pH 8.0. 1 ml of 10 mg/ml lysozyme in 0.25M EDTA pH 8.0 was added and the mixture incubated on ice for 20 min. 0.75ml of TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and 0.25 ml of lysis solution [5% (w/v) sodium n-lauryl sarcosine, 50 mM Tris-HCl pH 8.0, 62.5 mM EDTA pH 8.0] were then added, along with 2 mg solid pronase (Boehringer-Mannheim) and mixed gently. The mixture was incubated at 65°C for a minimum of 60 min. This was followed by three extractions with TE-saturated phenol (pH 7.5) and two extractions with diethyl-ether. The genomic DNA was precipitated with four volumes of 100% ice cold ethanol, washed with 70% ethanol and resuspended in 1 ml of TE or Milli-Q water, then stored at 4°C.

2.12 Analysis and manipulation of DNA

2.12.1 DNA quantitation

DNA concentration was determined using a Pharmacia LKB Ultrospec Plus spectrophotometer, assuming that A_{260} of 1 is equal to 50 µg dsDNA/1 ml (Maniatis *et al.*, 1982).

2.12.2 Restriction endonuclease digestion of DNA in solution

Most cleavage reactions were done using the restriction enzyme buffer specified and recommended by the manufacturers of the enzymes. 0.1-0.5 µg of DNA was incubated with 2 units of each restriction enzyme in a final volume of 20 µl, at 37°C, for 2 hours, with the exception of digests using *Sma*I, which were incubated at 25°C. Reactions for gel electrophoresis were terminated by heating at 65°C for 10 min. Prior to loading onto a gel, the reactions were mixed with a one-tenth volume of tracking dye [15% (w/v) Ficoll, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol and 0.1 mg/ml RNaseA.

2.12.3 Analysis of restriction fragments

Electrophoresis of digested DNA was carried out at RT on horizontal 0.8% or 1% (w/v) agarose gels (Seakem HGT) in Easy-Cast™ gel tanks (OWL Scientific Inc.). Gels were run at 100V in either TBE buffer (67 mM Tris base, 22 mM boric acid and 2 mM EDTA, final pH 8.8), or TAE buffer (40 mM Tris acetate and 2 mM EDTA). After electrophoresis the gels were stained in distilled water containing 2 mg/ml ethidium bromide for 10-15 min. DNA bands were visualized by trans-illumination with UV light and photographed using either Polaroid 667 positive film or 665 negative film.

2.12.4 Calculation of DNA restriction fragment size

The size of restriction enzyme fragments were calculated by comparing their relative mobility with that of *Eco*RI digest *Bacillus subtilis* bacteriophage SPP1 DNA. The calculated sizes of the SPP1 *Eco*RI fragments used were (in kb) : 8.55; 7.42; 6.10; 4.90; 3.64; 2.80; 1.95; 1.88; 1.51; 1.41; 1.16; 0.99; 0.71; 0.49; 0.36; and 0.081. The size of small restriction enzyme fragments were calculated by comparing their relative mobility with that of DNA ladder markers (DMW-100L) purchased from GeneWorks. The sizes of DMW-100L were (in bp) : 1000; 900; 800; 700; 600; 500; 400; 300; 200; and 100.

2.12.5 Extraction of DNA restriction fragments from an agarose gel

The required DNA restriction fragments from an agarose gel were excised by using QIAquick Gel Extraction Kit purchased from QIAGEN, according to the manufacturer's instruction.

2.12.6 Dephosphorylation of DNA using alkaline phosphatase

Restriction enzyme-digested DNA was dephosphorylated with alkaline phosphatase according to Maniatis *et al.*(1982). After 0.1-0.5 µg DNA was completely digested with restriction enzymes, 1 unit of alkaline phosphatase [calf intestinal phosphatase (C.I.P.) was added and the mixture was incubated at 37°C for 30 min. The reaction was terminated by the addition of EDTA pH 8.0 to a final concentration of 3 mM, followed by heating at 65°C for 10 min. The reaction mix was extracted with TE-saturated phenol and the DNA precipitated with 0.1 volume of 3 mM sodium acetate pH 4.8 and 2 volume of ethanol, followed by incubation at -20°C for 1 hour. DNA was collected by centrifugation (15000

rpm, 15 min, 4°C), washed with 70% ethanol (v/v), and dried in *vacuo*. DNA was resuspended in an appropriate volume of TE or Milli-Q water and stored at -20°C.

2.12.7 End-filling of linear DNA with Klenow-fragment

Protruding 5′ ends created by cleavage with restriction endonuclease were end-filled using the Klenow fragment of *E. coli* DNA polymerase I (Klenow) by the following method. Completely digested DNA of approximately 1 µg in a final volume of 20 µl was heated at 65°C for 20 min to terminate the digest, with the exception of digests using *Bam*HI, which were terminated by heating at 80°C for 20 min. A mixture of 2 mM dNTPs to a final concentration of 40 µM was added with 2 units of Klenow, and incubated at 37°C for 30 min. The reaction was terminated by heating at 65°C for 30 min, and the volume made up to 0.1 ml with Milli-Q water. The mixture was extracted with Tris-saturated phenol (pH 7.5), chloroform, isoamyl alcohol (25 : 24 : 1). DNA was precipitated with 0.1 volume 3M sodium acetate pH 5.2, 3 volumes of 100% ethanol and incubation at -20°C for 30 min. 20 µg glycogen was added as a carrier. DNA was recovered by centrifugation (15 min, microfuge), pellets were washed with 70% (v/v) ethanol and dried *in vacuo*. DNA was resuspended in an appropriate volume of Milli-Q and stored at -20°C.

2.12.8 Ligation of DNA fragments

Ligations were performed using approximately 100 ng vector DNA, and insert DNA at an approximate 3-fold Molar excess. The final volume of 20 µl, in 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5% (w/v) PEG, 40 mM ATP, 0.5 mM DTT (dithiothreitol) using 2 units of T4 DNA ligase was incubated overnight at 16°C for "sticky end" ligations

or 4°C for "blunt end" ligations. The ligated DNA was either electroporated or transformed into *E. coli* strains.

2.12.9 Exonuclease III treatment of linear DNA

Exonuclease III treatment of the ligation products was performed as described by Collis *et al.* (1993). The portion of ligation mix was diluted with an equal volume of water, followed by addition of 150 units exonuclease III (New England Biolabs) and incubated at 37°C for 60 min. The ligation products were separated by electrophoresis, transferred to a membrane and visualized by hybridization with a *vlpA*-specific DNA probe.

2.12.10 Labeling of DNA fragments

Purified DNA fragments either from restriction endonuclease fragments or DNA from PCR products were labeled to use as DNA probes. Digoxigenin-11-dUTP (DIG-11-dUTP; Boeringer-Mannheim) was used to label these DNA fragments according to the manufacturer's protocol. DNA fragments were labeled using a random-priming reaction (Feinberg and Vogelstein, 1983). DNA to be labeled (10 ng-3 ng) was denatured at 95°C for 10 min, and chilled quickly on dry ice or ice / alcohol for 3 min before the addition of 2 µl hexanucleotide mixture, 2 µl dNTP labeling mixture, and 2 units of Klenow. The reaction volume was made up to 20 µl with sterile Milli-Q water, and incubated at 37°C for at least 2 hours or up to 20 hours (longer incubation can increase the amount of labelled DNA). The reaction was stopped by the addition of 2 µl of 0.2M EDTA pH 8.0. DNA was precipitated by the addition of 2.5 µl 4M LiCl, and 60 µl of pre-chilled 100% ethanol. After the mixture was mixed well and incubated for a minimum of 2 hours at -20°C or 30 min at -70°C, DNA was collected by centrifugation at 12000 g for 10 min. The DNA pellets were

washed with 50 µl pre-chilled 70% ethanol (v/v), dried *in vacuo* and resuspended in 20 µl 10 mM Tris-HCl, 1 mM EDTA pH 8.0 (TE) or in Milli-Q water.

2.13 Southern transfer and hybridization (Southern blot)

Digested chromosome DNA samples and markers were electrophoresed overnight on a 1.0% (w/v) agarose gel, stained with EtBr and photographed. The gel was denatured by soaking in several volumes of 1.5M NaCl and 0.5M NaOH solution for 1 hour with slow agitation. The chromosome DNA from the agarose gel was transferred overnight onto Hybond^{TM} - N+; positively charged nylon membrane (Amersham), at room temperature using 10X SSC as the transfer buffer. The method followed was as described by Southern (1975) as modified by Maniatis *et al.* (1982). After transfer the filter was either fixed for 20 min with 0.4M NaOH or baked for 1 hour at 80°C. Filters were incubated in pre-hybridization buffer at 42°C for 2 hours prior to hybridization with the probes. Pre-hybridization buffer for DNA fragment probes consisted of 5% SSC (0.15M NaCl, 0.015M sodium citrate), 50 mM sodium phosphate buffer pH 6.4, 50% formamide, 5X Denhardt's reagent (0.1% Ficoll, 0.1% polyvinylpyrollidone, 0.1% Fraction V BSA) and 0.1 mg/ml herring sperm DNA. Pre-hybridization buffer for oligonucleotide probes consisted of 1M NaCl, 0.1M Tris-HCl pH 7.6, 5X Denhardt's reagent, 0.05% SDS, 0.025M EDTA and 0.1 mg/ml herring sperm DNA. After prehybridization, the DIG-labeled probes were heated to denature them for 10 min (100°C for fragment probes, 75°C for oligonucleotide probes) and added to the filter in pre-hybridization buffer, followed by incubation for 18 hours at 42°C (37°C for oligonucleotide probes). Following hybridization with oligonucleotide probes, stringency washes (3 x 10min) were performed at 5°C below the melting temperature (Tm) of the oligonucleotide with 5X SSC, 0.1% SDS. Filters hybridized with

fragment probes were washed two times with shaking for 5 min in 2X SSC, 0.1% SDS at room temperature, followed by two times with shaking for 15 min in 0.2X SSC, 0.1% SDS at 65°C. Filters were incubated in blocking reagent [5% skimmed milk in buffer 1 (0.1M Tris-HCl, 0.15M NaCl pH 7.5)] for 1 hour before being incubated with anti-digoxigenin-POD Fab fragments (1/10 K dilution in buffer 1, Boehringer-Mannheim) for at least 30 min. Unbound antibody-conjugate was removed by washing the filters four times for 5 min in buffer 1, and once with PBS (10 mM potassium phosphate pH 7.4, 0.82% (w/v) NaCl). DNA bands were visualized using ECL (Amersham or Boehringer Mannheim), and the filter was exposed to X-ray film at room temperature.

2.14 Colony hybridization (colony blot) using labeled DNA probes.

The colony hybridization procedure was followed the method described by Grunstein and Hogness (1975). Individual colonies were grown in 100 µl of appropriate broth in a 96 well microtritre tray (Falcon) at the appropriate temperature without agitation for 16 hours. After incubation, 5 µl of individual cultures were spotted onto Hybond^{TM} - N+; nylon membrane (Amersham) and allowed to air dry. The filter was treated with 0.5M NaOH for 5 min followed by soaking of the filter in 0.1M NaOH, and placed on Whatmann 3M paper soaked with lysis solution (10% SDS) with spotted colonies on the upper side. After 5 min at room temperature, the filter was transferred to Whatmann 3M paper soaked with denaturation solution (0.5M NaOH, 1.5M NaCl) for a further 5 min. The filter was transferred to Whatmann 3M paper soaked with neutralization solution (1.5M NaCl, 0.5M Tris-HCl pH 8.0) for 5 min, and air dried for 35-40 min before the treated colonies were fixed to the membrane by treatment with 0.4M NaOH for 20 min, followed by a brief wash in 5X SSC. Cell debris was removed from the filter by placing the filter in

0.1% SDS and gently scrubbing the surface. The filter was incubated at 42 ° C for 1-2 hours with agitation in prewashing solution (50 mM Tris-HCl pH 8.0, 1M NaCl, 1mM EDTA pH 8.0, 0.1% (w/v) SDS). Prehybridization, hybridization and detection were performed as described above for Southern hybridization.

2.15 Polymerase chain reaction (PCR) and long range PCR

2.15.1 Standard PCR

PCR products were generated by the protocol as described by Delidow (1993). The PCR reaction was performed in GENE-Amp™ reaction tubes (0.5 ml, Perkin Elmer) in a 50 µl volume containing Taq buffer (50 mM KCl, 10mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin), 2 µM deoxynucleoside triphosphates (dNTP) (at final concentration), 20 pmol of primer, 100 ng of plasmid template or genomic DNA and 2 units Taq polymerase (Perkin-Elmer). The reaction was overlaid with a drop of light mineral oil (Nujol, Perkin-Elmer). The initial denaturation period was 30 sec at 95°C, and the reaction cycle was subjected to 25 cycles of amplification (95°C x 30 sec, 55°C x 1 min, 72°C x 1 min) using a DNA thermal Cycler (Perkin-Elmer). PCR products were extracted and cleaned by using QIAquick PCR Purification Kit (QIAGEN), according to the manufacturer's instruction.

2.15.2 Long range PCR

Amplification of long sequence DNA (up to 35-kb) was performed by the PCR method as described by Barnes (1994), using the Expand™ Long Template PCR System Kit (Boehringer-Mannheim) according to the manufacturer's protocol. 0.2 ml thin-walled PCR tubes (Perkin Elmer) were used to perform the reactions. The primers used in the

reactions consisted of 33 mers with Tm > 68°C. The reaction consisted of buffer (25 mM Tris pH 9.1, 10 mM $(NH_4)_2SO_4$, 250 µM dNTPs, 3.5 mM $MgCl_2$), 100 ng of each primer, 2 units Taq/Pfu (80 : 1), 100 ng genomic template, and was made up to a final volume of 50 µl with sterile Milli-Q water. After overlay with a drop (40 µl) of light mineral oil, the reaction cycle was subjected to 10 cycles of amplification (94°C x 10 sec, 54°C x 30 sec, 68°C x 15 min). The reaction cycle was linked to another 10 cycles of amplification (94°C x 10 sec, 65°C x 30 sec, 68°C x 15 min). PCR products were extracted and cleaned as described above for standard PCR.

2.15.3 PCR DIG - labeled DNA probes

The protocol used for DIG-labeled PCR is essentially that described in **Section 2.12.10**. The PCR reaction was performed in 0.5 ml reaction tubes in a 50 µl volume containing supplied Taq buffer, 200 µM each of deoxynucleoside triphosphate dATP, dCTP, and dGTP, 190 µM of dTTP and 10 µM of DIG-11-dUTP, 200 ng of plasmid template or genomic DNA and 2.5 units of Taq polymerase (Perkin-Elmer). The reaction was overlaid with a drop of light mineral oil and subjected to the cycle detailed in **Section 2.15.1**.

2.16 Analysis of DNA sequences

DNA nucleotide sequences were determined by the dideoxy chain terminator procedures as described by Sanger *et al.* (1977, 1980) using a DNA thermal cycler (Perkin-Elmer). DNA sequencing was performed using dye terminator sequencing kits purchased from Perkin-Elmer. The sequencing reactions were performed in GENE-Amp™ reaction tubes (0.5 ml, Perkin Elmer) consisting of 8 µl Terminator Ready Reaction Mix (Perkin-

Elmer), 1-2 µg double-stranded DNA template, and 3.2 pmol primer. This reaction mixture was made up to a final volume of 20 µl with sterile Milli-Q water, and overlaid with *ca.* 40 µl of light mineral oil (Nujol, Perkin-Elmer). The reaction was subjected to 25 cycles of 96°C x 30 sec, 50°C x 15 sec, 60°C x 4 min. The products were removed from GENE-Amp™ reaction tubes and transferred into a sterile 1.5 ml microcentrifuge. DNA was precipitated by the addition of 2.0 µl of 3M sodium acetate pH 5.2 and 50 µl of 100% prechilled ethanol and incubated on ice for 20 min. DNA was collected by centrifugation at 15000 rpm for 30 min at 4°C. The pellets were washed with 70% ethanol, and dried *in vacuo*. DNA sequencing was performed by the Sequencing Center of the Institute of Medical and Veterinary Science (I.M.V.S.) (Adelaide, South Australia) using an Applied Biosystems model 377 DNA sequencer. Raw sequence data sets were analyzed using protein analysis programs DNASIS and PROSIS. DNA sequences were compared with entries in the GenBank and EMBL databases using the BlastX algorithm (Atschul *et al.*, 1997). The predicted protein sequences were compared to PIR and SWISS-PROT databases using FASTA (Pearson, 1990) or BlastX. Multiple sequence alignments were performed using the programme CLUSTAL W (Higgins and Sharp, 1989).

2.17 RNA analysis

2.17.1 RNA isolation from bacterial cultures

The method for RNA isolation is based on lysis of cells in the presence of guanidine followed by centrifugation through a caesium chloride cushion to pellet the RNA. A single colony was picked up from a fresh plate and grown overnight, then subcultured in the ratio 1 : 50 (v/v), and grown to mid-late log phase (A_{600} = 0.8). Cultures were chilled on ice for 10 min, and cells harvested by centrifugation at 7000 rpm for 10

min at 4°C. The pellets were resuspended in 2 ml guanidine isothiocyanate (GTC) solution [4M guanidine isothiocyanate, 0.1M Tris-HCl pH 7.4, 0.5% sarkosyl, dissolved in diethyl pyrocarbonate (DEPC) treated water (1 ml DEPC in 1 litre distilled water, left overnight, and then autoclaved)]. The lysate was transferred into a 3 ml syringe with a 19.5 gauge needle and syringed 6 times to shear chromosomal DNA, and overlaid on top of the CsCl cushion (2.5 ml of 5.7M CsCl, 10 mM EDTA) in the ½ x 2 inch Beckman Quick Seal polyallomer tubes (capacity 5.1 ml). The tubes were filled with GTC solution to the upper shoulder and balanced in pairs, sealed, and centrifuged (50000 rpm, 20°C, TLA 100.4) for 16-18 hours in a Beckman Optima™ TLX Ultracentrifuge. RNA pellets were visible as a translucent bulge on the side of the tubes near the bottom. RNA was dissolved in 150 µl of DEPC treated water, transferred into sterile screw-capped micrcentrifuge tubes, and precipitated by the addition of 0.1 volumes of 3M sodium acetate pH 5.2, and 3 volumes of prechilled 100% ethanol. This was incubated at -70°C for 2-3 hours, collected by centrifugation at 15000 rpm for 15 min at 4°C, washed with 70% (v/v) ethanol, dried in *vacuo*, and finally resuspended in sterile Milli-Q water at a concentration of 1 µg/µl and stored at -70°C.

2.17.2 RNA quantitation

The concentration of RNA in solutions was determined by measurement of absorption at 260 nm using an LKB Ultroscope Plus spectrophotometer, assuming an A_{260} of 1.0 is equal to 40 µg RNA/ml (Miller, 1972; Maniatis *et al.*, 1982).

2.17.3 DNA probes preparation for use in Northern hybridization

The GIGA Prime DNA Labeling Kit (GeneWorks) was used to label purified DNA fragments used as probes in the Northern blots. DNA fragments were labeled using a random-priming reaction with ^{32}P-dNTP according to the manufacturer's instruction. DNA was denatured at 95°C for 10 min, then chilled on ice before the addition of 12 µl decanucleotide, 12 µl buffer, 5 µl sterile Milli-Q water, 2 units Klenow, and 5 µl ^{32}P-dNTP. 5 µl of glycogen was added to reaction volume along with, 5 µl 0.5M EDTA pH 8.0, 5 µl 10% SDS, and made up to a volume of 100 µl with Milli-Q water. DNA was precipitated with the addition of 100 µl 4M ammonium acetate, 600 µl prechilled 100% ethanol, and incubated at -70°C for 1 hour. DNA was collected by centrifugation at 15000 rpm for 20 min at 4°C, dried *in vacuo*, and dissolved in 194 µl 10 mM Tris-HCl pH 8.0, 1mM EDTA, and 6 µl 10N NaOH. DNA was neutralized by boiling at 95°C for 12 min, snap-cooling to 0°C, followed by the addition of 260 µl ice cold 4M ammonium acetate. This mixture was ready to use as a probe for Northern hybridization.

2.17.4 Northern hybridization

Separation of RNA was on denaturing agarose gels (Seakem HGT), followed by transfer to nylon membrane (Hybond^{TM} - N+ , Amersham) and hybridization with the specific DNA probes.

RNA was heated to remove secondary structure prior to electrophoresis. 25µg of each RNA sample was mixed with 50% (v/v) formamide, 1.1M formaldehyde prior to running in the agarose gel. RNA sample was heated at 65°C for 7 min, snapped cool on ice, and 5X loading buffer (25% (v/v) glycerol, 0.5% (w/v) xylene cyanol) was added immediately prior to loading onto a 1.25% (w/v) agarose gel (Seakem HGT) with 1.1M

formaldehyde in 1X MOPs pH 7.0. The gel was run in a running buffer of 1X MOPs (1X MOPs electrophoresis buffer consisted of 0.4M 3-(N-Morpholino) propanesulfonic acid, and 0.01M EDTA in DEPC treated water) at 100V constant voltage in a 16 x 22 cm gel apparatus (Easy-Cast^{TM} - OWL Scientific Inc.).

RNA from the gel was transferred overnight to a nylon membrane using 10X SSPE buffer pH 7.4 consisting of 175.3 g NaCl, 27.6g NaH_2PO_4, 7.4 g Na_2 EDTA, and made up to 1 litre with sterile Milli-Q water. The filter was baked at 80°C *in vacuo* for 2 hours to fix RNA to the membrane, then added to prehybridization solution consisted of 50% (w/v) formamide, 5X SSPE pH 7.4, 5X Denhardt's reagent, 0.5% (w/v) SDS, 0.05% (w/v) NaPPi, 200 µg/ml sheared/denatured DNA (herring sperm DNA in DEPC treated water and sheared by multiple passage through an 18G needle and denatured by boiling 12 min). Prehybridization proceeded for at least 3 hours at 42°C, a DNA probe was then added, and hybridization was allowed to proceed overnight at 42°C. The filters were washed twice with slow shaking for 30 min at 60°C in 500 ml of 1X SSC, 0.1% (w/v) SDS, 0.05% (w/v) NaPPi, followed by one wash in 500 ml of 0.12X SSC, 0.1% (w/v) SDS, 0.5% (w/v) NaPPi. After washes, filters were exposed to a phospho-image screen or exposed to X-ray film.

2.17.5 Primer extension

The method was followed from Ausubel *et al.* (1994). The 5′ terminus of a RNA was mapped by extending a primer using reverse transcriptase. A labeled oligonucleotide primer specific for the mRNA of interest was hybridized to the RNA, and extended by reverse transcriptase towards the start of the mRNA using unlabeled deoxynucleotides to form a single-stranded DNA complementary to the template RNA. The resultant DNA was

analyzed on a sequencing gel. The length of the product generated by this primer extension defines the start site of transcription.

An oligonucleotide primer specific for the mRNA of interest was labeled at the 5′ end using polynucleotide kinase and γ^{32} P-ATP. RNA (10 ng or 20 ng) was added to the annealing reactions. The extension products were precipitated using 3M NaAc pH 5.2 and ethanol and dried 10-20 min *in vacuo* (Speedivac). The dried extension products were dissolved by vigorous vortexing in 50% formamide loading solution and heated at 95-100°C for 5 min before being separated on a conventional DNA sequencing gel (6% acrylamide, 7M urea, 1X TBE). The gel size was 20 x 40 cm with a thickness of 0.4-0.5 mm. The extension products were run at a constant voltage setting of 1200V (~25 mA). The gel was placed overnight initially in contact with a phosphoimage screen, but various exposure times were used.

2.17.6 Reverse transcription PCR (RT-PCR)

The procedures for RT-PCR were combined from Aasinki (1997), and SuperScript™ II (Life Technologies). Reverse transcriptase enzyme used was Moloney Murine Leukemia Virus (MMLV) (SuperScript™ II). Total RNA (50 ng) was used as a template in which RNA is reverse transcribed to produce complementary DNA (cDNA) templates (the first strand DNA synthesis). Ten percent of the first strand cDNA was used in the subsequent reaction with initial denaturation at 95°C for 5 min. This was followed by 25 cycles of 30 sec denaturation at 95°C, 1min annealing at 55°C, and 1 min extension at 72°C using a DNA Thermal Cycler (Perkin Elmer). PCR products were extracted and cleaned as described above for PCR.

2.18. Protein analysis

2.18.1 Over-expression of *vlpA* using the T7 RNA polymerase expression system

The over-expression of *vlpA* gene was performed under the control of the T7 RNA polymerase promoter system by a modification of the method described by Tabor and Richardson (1985). *E. coli* strains (E2096) carrying the plasmid pGP1-2 were used as a source of T7 RNA polymerase. E5218 containing pPM5007 and pGP1-2 was streaked onto nutrient agar plate containing Km (50 μg/ml) and Ap (50 μg/ml), and incubated at 30°C overnight. A single colony was used to inoculate 10 ml of LB containing Km and Ap, then incubated overnight at 30°C with shaking. The culture was diluted 1 in 20 (v/v) into fresh LB with kanamycin and ampicillin and incubated at 30°C with shaking to an A_{600} of *ca.* 0.6. The flasks were shifted to 42°C with shaking for 30-45 min to induce expression of the T7 RNA polymerase. Rifampicin (4 mg/ml freshly prepared stock in methanol) was then added to a final concentration of 200 μg/ml and the culture maintained at 42°C for a further 30 min before being transferred to 37°C and incubated with shaking overnight.

2.18.2 Preparation of whole cell samples

Samples of whole cells were prepared for analysis by centrifugation of 1 ml of a mid-exponential phase culture for 5 min, at 15000 rpm at room temperature. The pellets were resuspended in 100 μl normal saline, and mixed with an equal volume of 2X SDS-sample buffer (62.5 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol and 0.05% (w/v) bromophenol blue) (Lugtenberg *et al.*, 1975). Samples were heated to 100°C for 3 min before storage at -20°C.

2.18.3 Preparation of whole membrane

Ten ml of mid-exponential phase culture ($A_{600} \approx 1$) were centrifuged at 10000 rpm for 10 min at 4°C. Pellets were resuspended in 10 ml of 30 mM Tris-HCl pH 8.1 and centrifuged again as above. The addition of 0.2 ml 20% (w/v) sucrose in 30 mM Tris-HCl pH 8.1 to the pellets was followed by 20 µl of 1mg/ml lysozyme (in 0.1M EDTA pH 7.3). The mixture was incubated on ice for 30 min, with 3 ml of 3 mM EDTA pH 7.3 added, and brief sonication with a Benson Model B15 Sonifer™ for 60 sec or until lysis occurred. Unlysed cells and large cell debris were removed by low speed centrifugation (5000 rpm, 5 min, 4°C). Whole membranes was collected by centrifugation at 15000 rpm for 60 min at 4°C and the drained pellets were stored in 1X SDS-sample buffer at -20°C.

2.18.4 Cell fractionation

The cell fractionation procedure was a modification of that described by Osborn *et al.* (1972). Cells were grown in LB to mid-exponential phase at 37 ° C (50 ml, A_{600}=0.6). Cells were pelleted in a Beckman SS-34 rotor, (7000 rpm. 10 min, 4 ° C) and resuspended in 1 ml of 20% (w/v) sucrose, 30 mM Tris-HCl pH 8.1, transferred to SM-24 tubes and chilled on ice. Cells were converted to sphaeroplasts with 0.1 ml of 1mg/ml lysozyme in 0.1M EDTA pH 7.3 for 30 min on ice, centrifuged and the supernatant was collected (periplasmic fraction). The cell pellet was frozen in an ethanol dry ice bath for 30 min, thawed and dispersed vigorously in 3 ml 3 mM EDTA, pH 7.3. Cells were lysed with a Branson Ultrasonifer (50% cycle, intermittent), by alternative freezing in dry ice with ethanol and thawing at 37 ° C. Unlysed cells and inclusion bodies were removed by slow speed centrifugation (7000 rpm, 10 min, 4 ° C). The supernatant containing the membranes and the cytoplasm was centrifuged at 35,000 rpm using a 50Ti or 80Ti rotor for 90 min at

20 ° C in a Beckman L8-80 ultracentrifuge. The supernatant (cytoplasmic fraction) was collected and the whole membrane pellet was resuspended in 1 ml H_2O. 500µl of Triton solution (4% Triton X-100, 2 mM $MgCl_2$, 50 mM Tris pH 7.5) was added to an equal volume of the whole membrane sample which was vortexed intermittently for 30 min at RT. The inner (soluble) membrane fraction was separated from the outer (insoluble) membrane fraction by centrifugation at 35000 rpm for 90 min in a 50Ti or 80Ti rotor (Beckman L8-80) at 20 ° C. The outer membrane fraction was resuspended in Milli-Q H_2O.

2.18.5 Preparation of outer membrane proteins (OMPs)

The bacterial cultures were grown to $A_{600} \approx 1$, then centrifuged at 10000 rpm for 10 min and the supernatants were discarded. Pelleted cells were resuspended to *ca.* 10^{11} cells/ml in 20% sucrose, 30 mM Tris HCl pH 7.8 and left on ice. 1 mg/ml lysozyme in 0.1M EDTA pH 7.3 was then added to the suspension to have a final lysozyme concentration of 0.1 mg/ml, and left on ice for 20-30 min. Cells were pelleted at 10,000 rpm for 10 min at 4 ° C. Pelleted cells were frozen in ethanol dry ice bath, thawed at 37 ° C 4 times, and resuspended in 3 mM EDTA pH 7.3. The suspension was sonicated twice for 1 min at 50% power. Cracked cells suspensions were centrifuged at 7000 rpm for 10 min to pellet any intact cells; and supernatants were centrifuged at 35000 rpm for 90 min (Beckman L8-80 Ultra-centrifuge) to pellet cell membranes. Membrane pellets were extracted with 2% Triton X-100 in 2 mM $MgCl_2$ 50 mM Tris-HCl pH 7.5 by vortexing intermittently for 30 min and kept on ice. Outer membranes (Triton insoluble material) were collected by spinning at 35000 rpm for 90 min. The pellets were resuspended in 0.25 mM Tris-HCl pH 7.5 and stored at -20°C as a 4 µg/ml stock.

2.18.6 Sodium dodecyl sulphate - polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed on 15% polyacrylamide gels using a modification of the procedure by Lugtenberg *et al.* (1975). Bacterial suspensions (10 µl of 10^{11} cells/ml in saline) were diluted with an equal volume of 2X SDS-sample buffer (Aiba *et al.*, 1981) and heated at 100°C for 5 min prior to loading onto the gel. Gels were generally electrophoresed at 100V while samples ran through the stacking gel, and then at 150V for 3-4 hours. Proteins were slowly stained with gentle agitation overnight in 0.04% (w/v) Coomassie brilliant blue and 3.5% (v/v) perchloric acid in distilled water, and destained in 5% acetic acid. Size markers (Pharmacia) were phosphorylase B (94kDa), bovine serum albumin (67kDa), ovalbumin (43kDa), carbonic anhydrase (30kDa), soybean trypsin inhibitor (20.1kDa) and alpha-lactalbumin (14.4kDa). Size markers for Western blotting gels were New England Biolabs prestained markers. The molecular masses were MBP-galactosidase (175kDa) [MBP=maltose-binding protein], MBP-paramyosin (83kDa), glutamic dehydrogenase from bovine liver (62kDa), aldolase from rabbit muscle (47.5kDa), triosephosphate isomerase from rabbit muscle (32.5kDa), β-Lactoglobulin A from bovine milk (25kDa), lysozyme from chicken egg white (16.5kDa) and aprotinin from bovine lung (6.5kDa).

2.18.7 Western transfer and detection

Separated proteins on an unstained SDS-PAGE gel were transferred to nitrocellulose membrane (Schliecher and Schuell) at 200mA for 2 hours in a trans-blot cell (Bio-Rad). The transfer buffer used was 25 mM Tris-HCl pH 8.3, 192 mM glycine and 20% (v/v) methanol (Towbin *et al.*, 1979). After transfer, the blot was incubated for 1 hour in Blotto (5% skim milk powder in TTBS [0.05% (v/v) Tween 20, 20 mM Tris-HCl pH

7.4, 0.9% (w/v) NaCl] to block non-specific protein binding sites before immersion in anti-VlpA serum (diluted 1/5000 in Blotto). After incubation with gentle agitation at RT for 2-16 hours, unbound antibody was removed by washing the immunoblot three times (10 min each) in TTBS with shaking. The filter was incubated with secondary antibody (HRP-conjugated goat anti-rabbit IgG at a dilution 1/40,000 in TTBS) for 2 hours at RT (gentle agitation). The filter then received four (5 min) washes with TTBS, followed by two (5 min) washes in TBS [20 mM Tris-HCl pH 7.4, 0.9% (w/v) NaCl] to remove unbound secondary antibody.

The antigen-antibody complexes were visualized using Enhanced Chemiluminescence (ECL) detection. ECL detection reagent was prepared according to the manufacturer's directions (Amersham); the filter was incubated in the reagent for 1 min, drained, covered with clear plastic and exposed to X-ray film (Kodak X-Omat) for 5 min.

2.18.8 Purification of proteins by electroelution.

Whole cell samples were electrophoresed on polyacrylamide gels (see **Section 2.18.6**). The gel was stained and destained as described above. Proteins can be electroeluted from gel slices by the method described by Leppard *et al.* (1993). After destaining the gel, the corresponding protein band was cut out of the gel ensuring that no other proteins were also excised. The gel slice was placed into small length dialysis tubing containing 1 ml of 0.2M Tris-acetate (pH 7.4), 1.0% SDS and 100 mM dithiothreitol (DTT) per 0.1 g of wet polyacrylamide gel. The tubing was placed in a horizontal electrophoresis tank and submerged in running buffer (50 mM Tris-acetate (pH 7.4), 0.1% SDS and 0.5 mM sodium thioglycolate), and allowed to electrophorese for 3 hours at 100V. The gel slices were removed from the tubing and stained with Coomassie blue to

check that the protein had been electroluted. The solution containing the protein of interest was dialyzed against distilled water. The purified protein was stored at -20° C.

2.18.9 Determination of protein concentration

Protein concentration was determined using the BCA™ protein assay agent (Pierce) according to the manufacturers instructions. The samples were incubated in a microwell plate at 37°C for 30 min. After incubation and cooling the plate to room temperature, the absorbance was measured at A_{570} in a Dynatech model MR 5000 ELISA tray reader. Concentrations were obtained by interpolation from a standard curve generated from known concentrations of BSA.

2.19 *In vivo* assay for virulence using the infant mouse cholera model

2.19.1 Infant mouse cholera model

Challenge *V. cholerae* strains were picked from fresh plates and grown in NB (Oxoid) at 37°C with shaking for 3 hours to an $A_{600} \approx 1.0$. The suspensions of the test strain were diluted in serial ten fold dilution and each was used to feed one group of 5 mice. The infant mice used were 3-4 days old, with a body weight at 2.4-2.7 g, and were removed from their mothers about 5-6 hours before use, to permit emptying of stomach contents. Each mouse received 0.1 mls of bacterial suspension, administered by means of a smooth-tipped hypodermic needle (26 gauge). Mice were kept on tissue-lined plastic containers in the laboratory for 48 hours at 25°C. After challenge, the number of survivors within each group was noted and these data used to construct a plot of cumulative percentage mortality versus $\log_{10}$ challenge dose (Attridge,1983). The estimate of the LD_{50} (the dose of vibrios capable of killing 50% of the mice within 48 hours) was obtained.

Strains with an $LD_{50} > 1 \times 10^8$ are considered avirulent in this model. Viable counts of the administered culture were also determined by plating dilution on NA with appropriate selection.

2.19.2 LD_{50} studies

The virulence of *V. cholerae* strain was assessed using the infant mouse model of cholera infection by determining the 48 hour LD_{50} value as described in **Section 2.19.1** (Attridge and Rowley, 1983). Cells from NB were pelleted (10 min, bench centrifuge) and resuspended in peptone-saline [0.1% (w/v) proteose-peptone (Difco), 0.85% (w/v) NaCl (PS)]. Serial dilutions of the strain of interest were made in peptone-saline, and each dilution orally administered to a group of 5-7 mice. The challenge dose (100µl dose per mouse) was delivered and viable counts of the administered cultures were determined by plating dilutions on NA with appropriate selection. Using these data, a plot of cumulative percentage mortality versus $\log_{10}$ challenge dose (Reed and Muench, 1938) is plotted and the 48 hours LD_{50} dose determined by interpolation.

2.19.3 Competition experiments

Competition experiments of parent and mutant strains *in vivo* were performed by feeding the infant mice a mixed inoculum comprising approximately equal numbers of two strains; the dose of each strain was equal to *ca.* 10 LD_{50} doses of the parent strain (Attridge, 1979). The mice were maintained at 25°C for 22-24 hours before being sacrificed and their entire small intestine removed and homogenized in 2 ml saline [0.9% (w/v) NaCl) using a homogenizer (Ultraturrax). To determine the competition ratio of parent and mutant organisms, which are present at the time of sacrifice, suitable dilutions of each gut

homogenate were spread onto NA with appropriate selection, and incubated at 37°C for viable counting. From the number of colonies present on these plates the ratio of parent to mutant was calculated.

2.20 Production of antisera

2.20.1 Production of antisera against unprocessed *V. cholerae* O139 VlpA

Antiserum to VlpA protein from *V. cholerae* O139 was raised by subcutaneous immunization of a rabbit with the outer membrane protein preparation of the over-expressed VlpA protein from *E. coli* (E5218) (see **Chapter 6**). A rabbit from the Central Animal House of the University of Adelaide was pre-bled and this serum shown not to contain antibodies that cross-reacted with the *E. coli* DH5α strain. The rabbit was immunized with 100 μg of the partially purified fraction emulsified in Freund's incomplete adjuvant (Commonwealth Serum Laboratories) in at least four sites and boosted four times at two week intervals with 100 μg of the partially purified VlpA protein. Ten days after the final immunization the rabbit was exsanguinated and the final serum stored at -20°C. A working stock of the antiserum was kept at 4°C with 0.05% (w/v) azide. The antiserum was absorbed four times with outer membrane protein from an *E. coli* with no VlpA clone present. Alternating absorptions were incubated at 37°C for 4 hours or overnight at 4°C. Following each absorption the serum was clarified by centrifugation and after the final absorption passed through a 0.2 μm Millipore filter. This serum is hereafter referred to as anti-VlpA serum and was confirmed by Western blot.

2.20.2 Antisera against cholera toxin and OMPs

Rabbit polyclonal anti-TCP and mouse anti-*V. cholerae* OMPs were kindly provided by Dr. S. Attridge (The University of Adelaide).

2.21 *In vitro* characterization of mutant strains from growth rate

The growth rate of *V. cholerae* O139 single and double *vlpA* mutants were compared with those of the corresponding wild-type strains by performing *in vitro* competition experiments in both nutrient and M9 minimal media. Mixed inocula of *ca.* 5 x 10^3 cells of each strain were prepared by dilution of overnight or early exponential phase cultures. The mixed cultures were incubated at 37°C with shaking for 7-8 hours (*ca.* 10^8 bacteria/ ml). To prevent the NB cultures from reaching stationary phase during this period, they were subcultured after ca. 4 hours. The initial ratio of the two strains in the mixed culture was determined by plating dilutions of the culture onto NA and selective media; this ratio was compared with that present in the culture after 7-8 hours incubation at 37°C.

2.22 Integron and gene cassettes

2.22.1 Construction of circular gene cassettes

Circularized cassettes were constructed by digesting 10µg of plasmid DNA containing the *vlpA* gene insertionally inactivated with the antibiotic resistance cartridge (Km^R) and flanked by the repeat sequences (VCRs) on both sides. After digestion with enzyme *Bsu*36I in both of VCRs sequences, this fragment was separated in a 1% agarose gel and isolated from a gel slice using QIAquick Gel Extraction Kit (QIAGEN), according to the manufacturer's instructions. Fragments were ligated in 200 µl of ligation buffer (25 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 5% (w/v) PEG, 1 mM ATP, 10 mM dithiothreitol

and 8 units T4 DNA ligase (New England Biolabs) at 4°C for 16 hours as described by Collis *et al.* (1993). After phenol extraction, DNA was isolated by precipitation with prechilled 100% ethanol, and dissolved in TE. A fraction of this DNA was run on a 1.5% agarose gel to confirm that circularization had occurred. The ligation products were also treated with exonuclease III to confirm that exonuclease III digestion had removed linear monomers and dimers from the ligation mix as described in **Section 2.12.9**.

2.22.2 Insertion of circular cassettes into integrons

Circular cassettes from **Section 2.22.1** were electroporated into *E. coli* containing the relevant plasmid using the Gene Pulser (Bio-Rad). The electroporation parameters were set at capacitance 25 μF, resistance 400 Ω and field strength 12.5kV/cm as described by Collis *et al.* (1993). The electroporated cells were immediately diluted in 1 ml of SOC media (2% Bacto-tryptone, 0.5% bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20% glucose). The suspension was grown with shaking at 37°C for 60 min. Aliquots were then plated in duplicate on LB plates containing appropriate antibiotics, resistance to which was specified by the incoming cassettes. A further 200 μl of the culture was diluted 25-fold into LB containing the appropriate antibiotics to final volume 5 ml, and grown shaking at 37°C for 16 hours. After centrifugation, the pellets were plated on three selective plates and incubated at 37°C for 2-3 days. Colonies which appeared after 2-3 days were patched on selective media and only those colonies that grew at this stage were scored.

2.22.3 Analysis of cassette deletion

The method used for cassette deletion was followed as described by Collis *et al.* (1992a). *E. coli* DH5α with plasmid vectors containing part of an integron (integrase *intI*1 or *intI*4, and *attI*1 or *attI*4), including gene cassettes inserted with antibiotic resistance cartridges, were initially grown on LB supplemented with all antibiotics to which the strain was resistant. After overnight incubation at 37°C, the culture was diluted 2000-fold into 2 ml of fresh LB containing only antibiotic for vector-encoded determinants. Cells were grown at 37°C to stationary phase, plasmid DNA was isolated by the alkaline lysis method. DNA was dissolved in Milli-Q water in a final volume of 50 µl. DNA was diluted 1/10 before introduced into *E. coli* under conditions in which fewer than 1 in 100 transformants received more than one DNA molecules. Transformants were selected on media containing only antibiotic for vector-encoded determinants, and colonies were then patched onto media containing appropriate antibiotics to screen for cassette deletion.

Chapter 3

The distribution and variation of the *vlpA* genes and linked VCRs in *Vibrio* and related genera

3.1 Introduction

The presence of multiple copies of *vlpA* genes in *V. cholerae* was first demonstrated by Barker and Manning (1997). It was also shown that *vlpA* is located in a region of the *V. cholerae* chromosome known to contain of 124-bp direct repeat sequences (VCRs). Recent reports have revealed that VCR represent a significant portion (about 12% or approximately 125-kb) of the 1.0-Mb smaller *V. cholerae* O1 chromosome (Clark *et al.*, 1997; Mazel *et al.*, 1998; Heidelberg *et al.*, 2000; Waldor and RayChaudhuri, 2000), and more than 150 copies of VCR were found in the *V. cholerae* El Tor strain N16961 chromosomal integron (Rowe-Magnus *et al.*, 1999). It is now clear that these repeat sequences share many features of the 59-base elements (59-be) or *attC* sites which are associated with the integron gene capture system (Recchia and Hall, 1997). Thus, if this chromosomally located integron-like element (or integron island) is functional in *V. cholerae*, and *vlpA* shares characteristic of cassette genes recognized by this gene capture system, it will be of interest to characterize the distribution and genetic variation of this gene which is always flanked by VCRs. Although integrons have site-specific recombination systems capable of capturing genes that are contained in gene cassettes (Stoke and Hall, 1989; Hall and Collis, 1995; Recchia and Hall, 1995a), the reasons for the appearance of multiple gene copies in the chromosome still needs to be explained. This information, derived from the *vlpA* genes in *Vibrionaceae* and related genera, will

contribute to the understanding of mobile cassette genes, and the gene capture system. It may also allow predictions to be made with regard to the genomic evolution in these bacteria.

3.2 Results

3.2.1 Distribution of the *vlpA* genes in *Vibrio* and related genera.

Prior to examining the copy number of *vlpA* genes, colony hybridization was performed in strains of *V. cholerae* O1 classical and El Tor biotypes, O139, and 155 strains of non-O1 serogroups, using PCR DIG-labeled *vlpA* (PCR product from oligonucleotides #1071 and #1072) as a probe. Colony hybridization was observed in all strains of *V. cholerae* O1 classical and El Tor biotypes and O139 tested, but not in all strains of the non-O1 serogroup (Figures 3.1A, B, C). In addition, some strains of *V. mimicus*, *V. fluvialis*, *V. parahaemolyticus*, and *Aeromonas spp.* were tested with no positive hybridization signal being detected (Figure 3.1D).

PCR analysis was performed to confirm the presence of *vlpA* in *V. cholerae* O1, O139, non-O1, other *Vibrio spp.* and *Aeromonas spp.* using the internal *vlpA* oligonucleotide primers (#1071 and #1072) which bind at the start and the end of *vlpA* to give a 516-bp PCR product. The results from PCR analysis correlated well (Figures 3.2A, B, C, and Figures 3.3A-J) with the results shown by colony hybridization. These data indicated that the *vlpA* gene is present in all strains of *V. cholerae* O1, O139, *V. anguillarum* and in some strains of *V. cholerae* non-O1. However, PCR products were not observed in *V. cholerae* non-O1 serogroup strains O113 and O124 which showed a positive result in the colony hybridization. In contrast, PCR products were observed in eight strains of *V. cholerae* non-O1 (O116, O117, O120, O122, O136, O137, O140 and

Figure 3.1 Colony blots of *V. cholerae* O1, O139 , non-O1, *Vibrio spp.* and *Aeromonas spp.* using a DIG-labeled *vlpA* (PCR product of oligonucleotides #1071 and #1072) as a probe.

Cultures were grown in microtitre trays to an A_{600} = 0.6. 5 µl of the samples were spotted onto Hybond™ - N^+; positively charged nylon membrane (Amersham), and treated as described in **Section 2.14**, then probed for *vlpA*. The serotypes of the strains are indicated on the Figures and are listed in Table 3.1. The positive controls are *V. cholerae* O139 (V911) and *E. coli* DH5α harboring pPM5007. The negative control is *E. coli* DH5α.

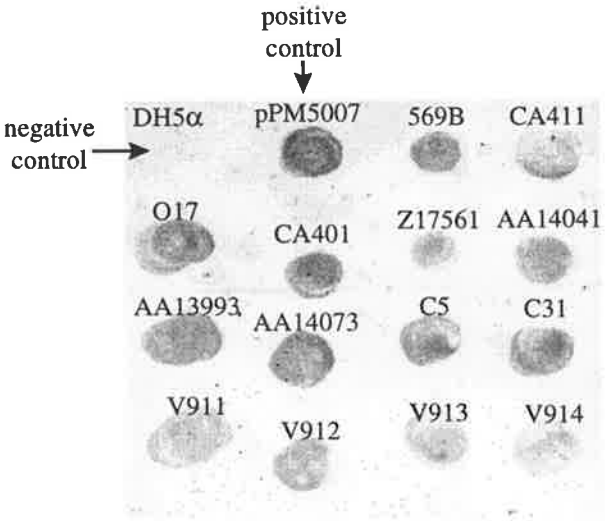
(A) *V. cholerae* O1 and O139 strains

(B) and (C) *V. cholerae* non-O1 strains

(D) *Vibrio spp.* and *Aeromonas spp.* (V563 : *V. mimicus*; V564 : *V. fluvialis*; V691: *V. parahaemolyticus*; V800 : *V. mimicus*; V994 : *A. trota* #1203; E1128 : *A. hydrophila* #A006; E1129 : *A. sobria* #A187; E1130 : *A. sobria* #A191; E1132: *A. hydrophila* #A335; E1326 : *A. hydrophila* #AH-1; E1327 : *A. hydrophila* #AH-2;

Figure 3.1

(A)


positive control
negative control
DH5α
pPM5007
569B
CA411
O17
CA401
Z17561
AA14041
AA13993
AA14073
C5
C31
V911
V912
V913
V914

(B)

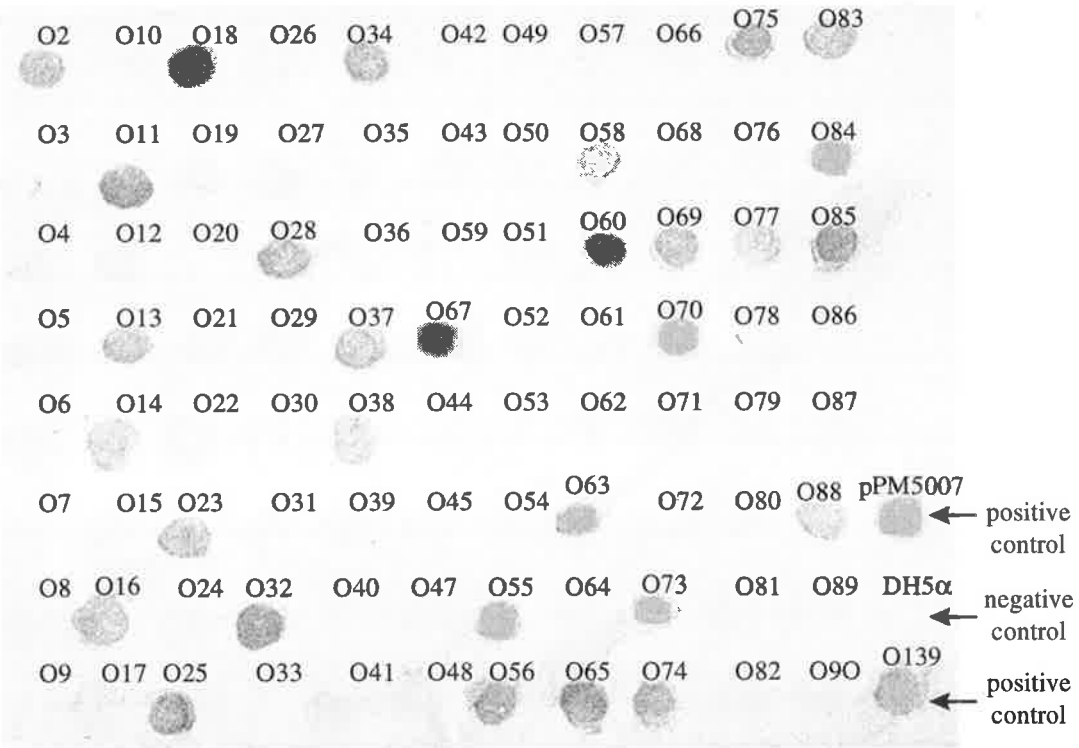

O2 O10 O18 O26 O34 O42 O49 O57 O66 O75 O83
O3 O11 O19 O27 O35 O43 O50 O58 O68 O76 O84
O4 O12 O20 O28 O36 O59 O51 O60 O69 O77 O85
O5 O13 O21 O29 O37 O67 O52 O61 O70 O78 O86
O6 O14 O22 O30 O38 O44 O53 O62 O71 O79 O87
O7 O15 O23 O31 O39 O45 O54 O63 O72 O80 O88 pPM5007
positive control
O8 O16 O24 O32 O40 O47 O55 O64 O73 O81 O89 DH5α
negative control
O9 O17 O25 O33 O41 O48 O56 O65 O74 O82 O90 O139
positive control

Figure 3.1

(C)


O91 O99 O108 O117 O125 O132 O140 O148
O92 O100 O109 O118 O46 O133 O141 O149
O93 O101 O110 O119 O126 O134 O142 O150
O94 O102 O111 O120 O127 O135 O143 O151
O95 O104 O113 O121 O128 O136 O144 O152
O96 O105 O114 O122 O129 O137 O145 O153
O97 O106 O115 O123 O130 O138 O146 O154
O98 O107 O116 O124 O131 O139 O147 O155
pPM5007
positive control
DH5α
negative control
O139
positive control

(D)

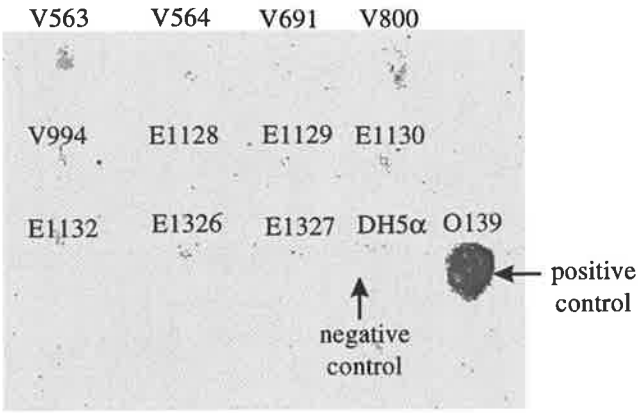

V563 V564 V691 V800
V994 E1128 E1129 E1130
E1132 E1326 E1327 DH5α O139
positive control
negative control

Figure 3.2 PCR amplification using *vlpA*-specific oligonucleotide primers (#1071 and #1072) in *V. cholerae* O1, O139, *V. anguillarum*, *Vibrio spp.* and *Aeromonas spp.*

(A) PCR amplification of *V. cholerae* O1 (Lanes 1-10), and O139 (Lane 11). Lanes contain: **1**, 569B; **2**, CA411; **3**, O17; **4**, CA401; **5**, Z17561; **6**, AA14041; **7**, AA13993; **8**, AA14073; **9**, C5; **10**, C31; **11**,V911; **12**, *E. coli* DH5α.

(B) *V. anguillarum.* Lanes contain : **1**, ATCC43305; **2**,ATCC43306; **3**,ATCC43307; **4**,ATCC43308; **5**,ATCC43309; **6**,ATCC43310; **7**,ATCC43311; **8**,ATCC43312; **9**,ATCC43313; **10**,ATCC43314; **11**,85-3954-1; **12**,85-3954-2; **13**, 86-3674; **14**,89-3748-1; **15**, V911; **16**, *E. coli* DH5α.

(C) *Vibrio spp.* and *Aeromonas spp.* Lanes contain PCR product from strains: **1**, *V. mimicus* (V563); **2**, *V. parahaemolyticus* (V691); **3**, *V. fluvialis* (V564); **4**, *A. trota* #1203; **5**, *A. trota* #1354; **6**, *A. hydrophila* #A006; **7**, *A. sobria* #A191; **8**, *A. hydrophila* #A335; **9**, *A. hydrophila* #AH-1; **10**, *A. hydrophila* #AH-2; **11**, *E. coli* DH5α; **12**, V911.

(*E. coli* DH5α and V911 were used as negative and positive controls, respectively. The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.)

Figure 3.2

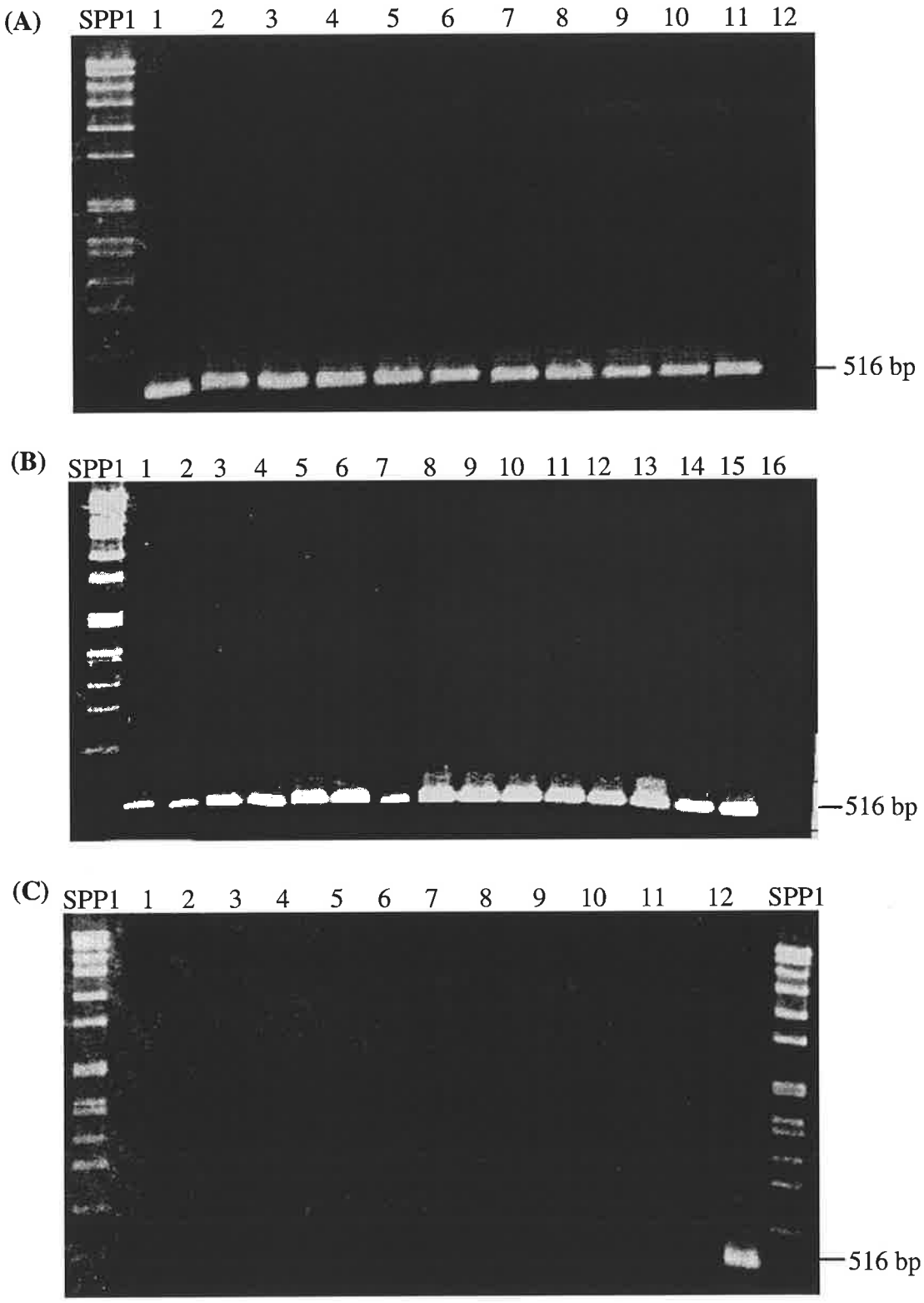

(A) SPP1 1 2 3 4 5 6 7 8 9 10 11 12
516 bp
(B) SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
516 bp
(C) SPP1 1 2 3 4 5 6 7 8 9 10 11 12 SPP1
516 bp

Figure 3.3 PCR amplification of *V. cholerae* non-O1 strains using *vlpA*-specific oligonucleotide primers (#1071 and #1072).

(A) Lanes contain: **1**, O2; **2**, O4; **3**, O5; **4**, O6; **5**, O7; **6**, O8; **7**, O9; **8**; O11; **9**, O12; **10**, O13; **11**, O14; **12**, O15; **13**, O16; **14**, O17; **15**, O18; **16**, O19; **17**, *E. coli* DH5α; **18**, V911.

(B) Lanes contain: **1**, O21; **2**, O22; **3**, O23; **4**, O24; **5**, O25; **6**, O26; **7**, O27; **8**, O28; **9**, O29; **10**, O30; **11**, O31; **12**, O32; **13**, O33; **14**, O34; **15**, O35; **16**, O36; **17**, *E. coli* DH5α; **18**, V911.

(C) Lanes contain: **1**, O37; **2**, O38; **3**, O40; **4**, O41; **5**, O43; **6**, O59; **7**, O67; **8**, O44; **9**, O45; **10**, O47; **11**, O48; **12**, O49; **13**, O50; **14**, O51; **15**, O52; **16**, O53, **17**, *E. coli* DH5α; **18**, V911.

(D) Lanes contain: **1**, O54; **2**, O55; **3**, O56; **4**, O57; **5**, O58; **6**, O60; **7**, O61; **8**, O62; **9**, O63; **10**, O64; **11**, O65; **12**, O66; **13**, O68; **14**, O69; **15**, O70; **16**, O71, **17**, *E. coli* DH5α; **18**, V911.

(*E. coli* DH5α and V911 were used as negative and positive controls, respectively. The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.)

Figure 3.3 (continued)

(E) Lanes contain: **1**, O72; **2**, O73; **3**, O74; **4**, O75; **5**, O76; **6**, O77; **7**, O78; **8**, O79; **9**, O80; **10**, O81; **11**, O82; **12**, O83; **13**, O84; **14**, O85; **15**; O86; **16**, O87; **17**, O88; **18**, V911.

(F) Lanes contain: **1**, O89; **2**, O90; **3**, O91; **4**, O92; **5**, O93; **6**, O94; **7**, O95; **8**, O96; **9**, O97; **10**, O98; **11**, O99; **12**, O100; **13**, O101; **14**, O102; **15**, O104; **16**, O105; **17**,*E. coli* DH5α; **18**, V911.

(G) Lanes contain: **1**, O106; **2**, O107; **3**, O108; **4**, O109; **5**, O110; **6**, O111; **7**, O113; **8**, O114; **9**, O115; **10**, O116; **11**, O117; **12**, O118; **13**, O119; **14**, O120; **15**, O121; **16**, O122; **17**, *E. coli* DH5α; **18**, V911.

(H) Lanes contain: **1**, O123; **2**, O124; **3**, O125; **4**, O46; **5**, O126; **6**, O127; **7**, O128; **8**, O129, **9**, *E. coli* DH5α; **10**, V911.

(I) Lanes contain: **1**, O130; **2**, O131; **3**, O132; **4**, O133; **5**, O134; **6**, O135; **7**, O136; **8**, O137, **9**, O138; **10**, O139; **11**, O140; **12**, O141; **13**, O142; **14**, O143; **15**, O144; **16**, O145; **17**, O146; **18**, *E. coli* DH5α; **19**, V911.

(J) Lanes contain: **1**, O147; **2**, O148; **3**, O149; **4**, O150; **5**, O151; **6**, O152; **7**, O153; **8**, O154, **9**, O155; **10**, V911; **11**, *E. coli* DH5α.

(*E. coli* DH5α and V911 were used as negative and positive controls, respectively. The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.)

Figure 3.3

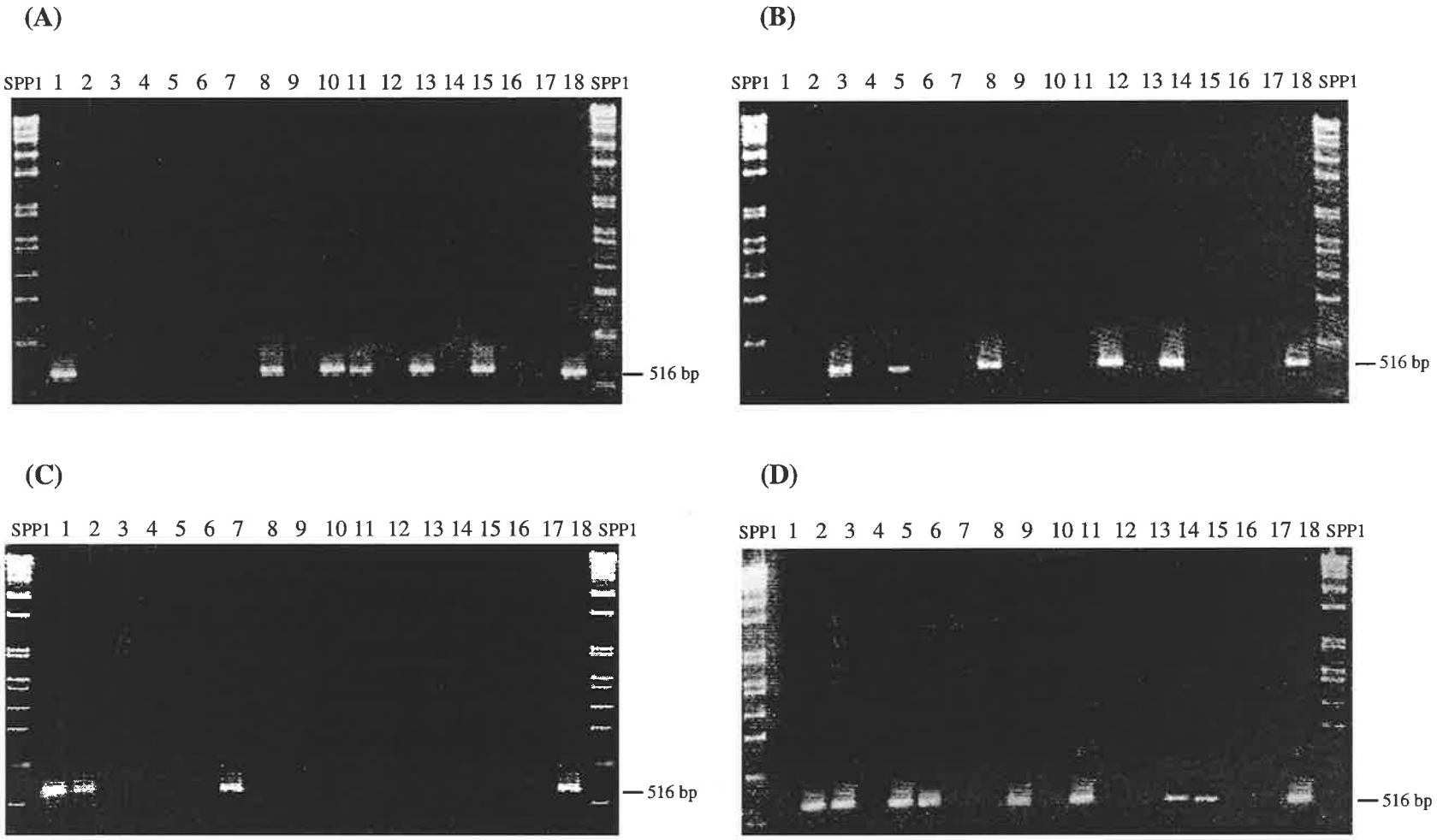

(A)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(B)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(C)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(D)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp

Figure 3.3

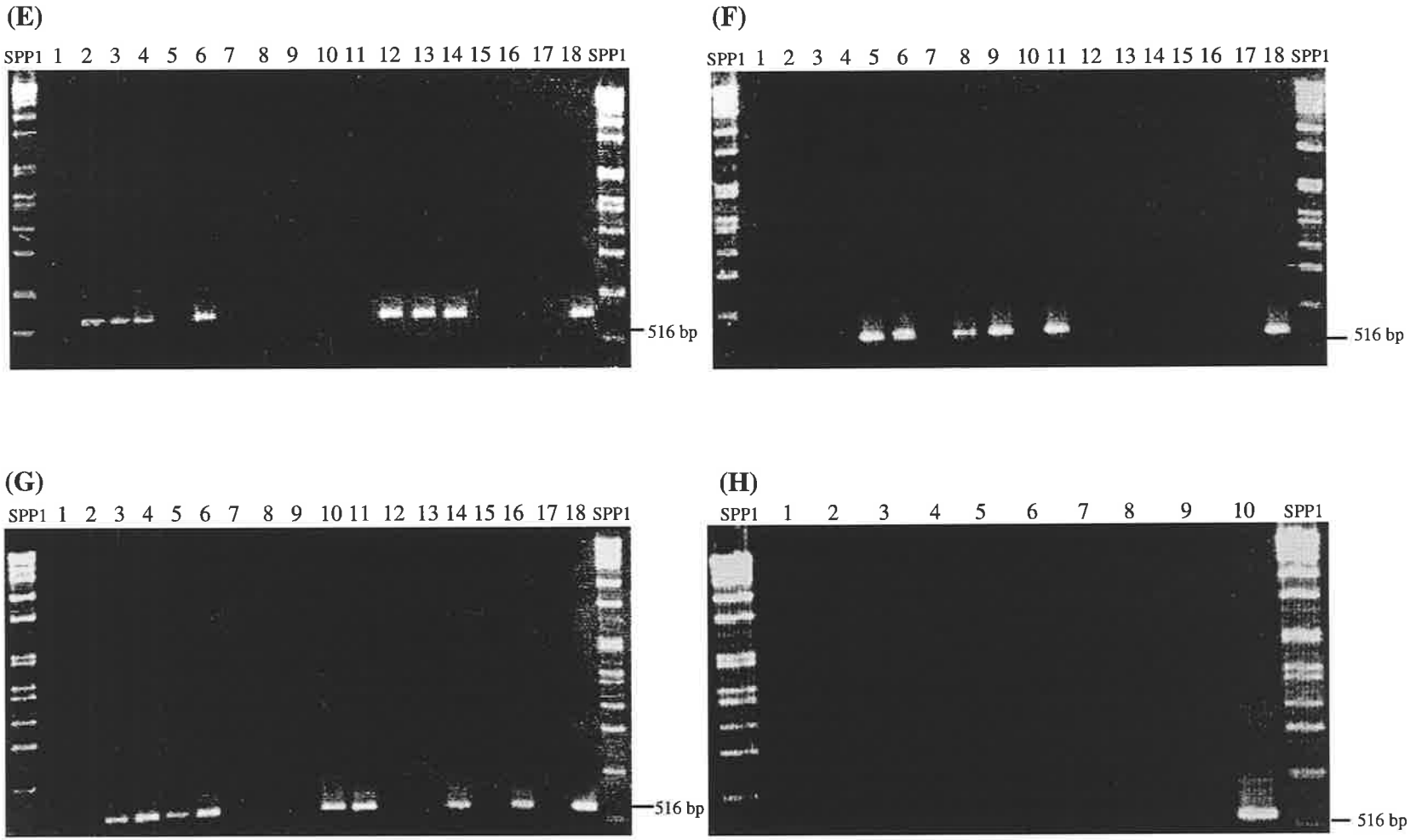
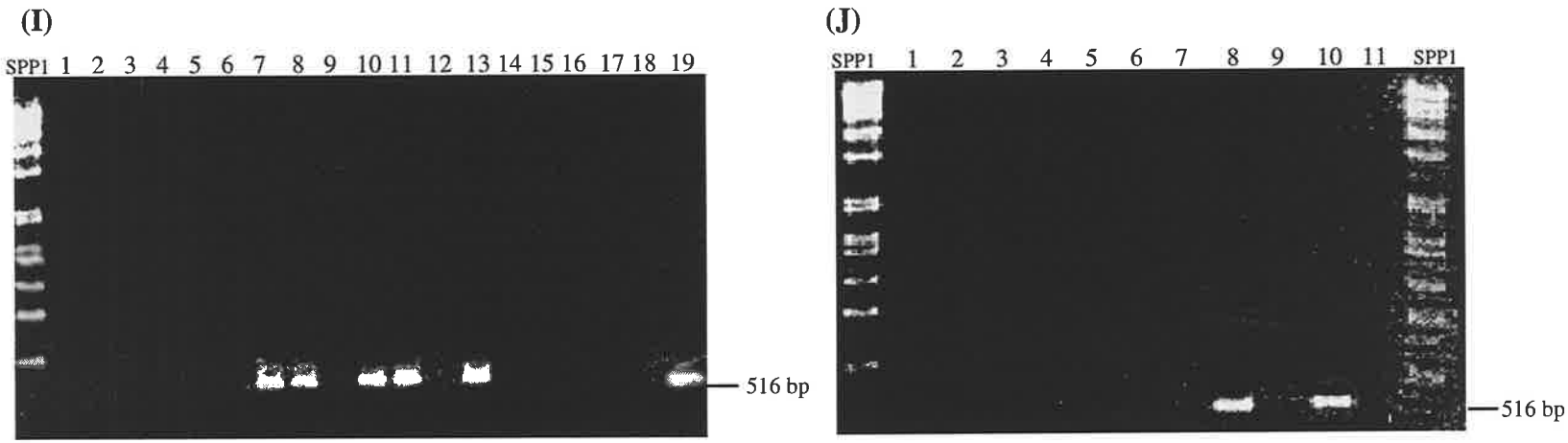

(E)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(F)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(G)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 SPP1
516 bp
(H)
SPP1 1 2 3 4 5 6 7 8 9 10 SPP1
516 bp

Figure 3.3


(I)
SPP1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
516 bp
(J)
SPP1 1 2 3 4 5 6 7 8 9 10 11 SPP1
516 bp

O154) which were negative for *vlpA* in the colony blots. The data are summarized in Table 3.1. PCR product was not observed in *V. mimicus*, *V. fluvialis*, and *V. parahaemolyticus*, supporting the data from the colony hybridization, i.e. that the *vlpA* gene is not present in these strains. In addition, PCR product was not detected in *Aeromonas spp.*, although members of the genus *Aeromonas* are grouped in the family *Vibrionaceae*.

The occurrence of multiple copies of *vlpA* were investigated by Southern hybridization using the PCR DIG-labeled *vlpA* probe as described for the colony blots. Chromosomal DNA from different strains of *V. cholerae* O1, O139 and *Vibrio spp.* were digested with *Eco*RI, while for *V. cholerae* non-O1 chromosomal DNA was digested with *Eco*RV to give the suitable fragment sizes to be hybridized on the nylon membrane. Both enzymes *Eco*RI and *Eco*RV do not digest within *vlpA* (sequence from EMBL GenBank database accession number X64097), thus it is possible to assume that these restriction enzyme sites are also not present in other copies of *vlpA*. The number of positive bands to appear for each strain were used to determine the copy number of *vlpA*. All strains of *V. cholerae* O1 and O139 examined showed multiple hybridization bands (Figures 3.4A, B) except for strain C31 (Figure 3.4A, lane 14). It is interesting to note that a large region of the genome of the strain C31 was reported to be deleted (van Dongen and De Graaf, 1986). The number of *vlpA* genes in *V. cholerae* and *Vibrio spp.* are summarized in Tables 3.2A, B, C, D, and E. *V. cholerae* O1 strains of the classical and El Tor biotypes tested in this study varied in the *vlpA* copy number present from two to four. All the *V. cholerae* O139 strains tested in this study contained only two hybridization bands (Figure 3.4B), whereas several strains of the *V. cholerae* non-O1 serogroup also showed variation in the number of bands ranging from one to three (Figure 3.4C).

Table 3.1 Presence of *vlpA* in *V. cholerae* non-O1 strains detected by colony hybridization (Data from Fig 3.1 B, C), and PCR amplification (Data from **Figure 3.3**).

Strain characteristics	No of strains containing *vlpA*	%	Serogroup
vlpA positive by both PCR and colony hybridization	41/154	26.62	O2, O11, O13, O14, O16, O18, O23, O25, O28, O32, O34, O37, O38, O55, O56, O58, O60, O63, O65, O67, O69, O70, O73, O74, O75, O77, O83, O84, O85, O88, O93, O94, O96, O97, O99,O108, O109, O110, O111, O139, O142
vlpA negative by both PCR and colony hybridization	103/154	66.88	O3, O10, O12, O15, O17, O19 - O22, O24, O26 O27, O29 - O31, O33, O35, O36, O39 - O54, O57, O59, O61,O62, O64, O66, O68, O71, O72, O76, O78 - O82, O86, O87, O89-O92 ,O95, O98, O100-O107, O112, O114, O115, O118, O119, O121, O123, O125, O135, O138, O141, O143 - O153, O155
vlpA positive by PCR only	8/154	5.2	O116, O117, O120, O122, O136, O137, O140, O154
vlpA positive by colony hybridization only	2/154	1.3	O113, O124

Table 3.2 (A) Genomic copy number of *vlpA* and DNA fragment size after digestion with *Eco*RI in different strains of *Vibrio* species. (Data from Southern hybridization **Figure 3.4A**).

No	Laboratory strain designation	Strain, serogroup and biotype	*vlpA* copy number	Estimated fragment size (kb)			
				1	2	3	4
1	V30	569B (O1 classical)	2	~	8.0	7.5	-
2	V32	CA411 (O1 classical)	3	~	~	8.0	-
3	V33	O17 SR (O1 El Tor)	4	~	8.0	6.1	3.0
4	V35	CA401 (O1 classical)	3	~	~	8.0	-
5	V563	*V. mimicus*	-	-	-	-	-
6	V564	*V. fluvialis*	-	-	-	-	-
7	V565	*V. cholerae* O53	-	-	-	-	-
8	V584	Z17561 (O1 classical)	2	~	8.5	8.0	-
9	V585	AA14041 (O1 classical)	3	~	8.0	7.5	-
10	V586	AA13993 (O1 El Tor)	4	~	8.0	7.0	3.0
11	V587	AA14073 (O1 El Tor)	4	~	~	8.5	7.5
12	V691	*V. parahaemolyticus*	-	-	-	-	-
13	V777	C5 (O1 El Tor)	4	~	7.5	7.0	2.5
14	V793	C31 (O1 El Tor)	1	7.5	-	-	-
15	V911	O139	2	~	7.2	-	-

~ too large to estimate size.

Table 3.2 (B) Genomic copy number of *vlpA* and the DNA fragment size digested with *Eco*RI in different strains of *V. cholerae* O139. (Data from Southern hybridization **Figure 3.4B**).

No	Laboratory strain designation	Strain	*vlpA* copy number	Estimated fragment size (kb)	
				1	2
1	V911	AI-1837	2	6.5	4.8
2	V912	AI-1838	2	6.5	4.8
3	V913	AI-4450	2	6.5	4.8
4	V914	AI-1841	2	6.5	4.8
5	V915	AI-1852	2	6.5	4.8
6	V916	AI-1854	2	6.5	4.8
7	V917	AI-1855	2	6.5	4.8
8	V918	AI-4260	2	6.5	4.8
9	V996	AI-18360	2	6.5	4.8
10	V997	232/95	2	6.5	4.8
11	V998	X-004	2	6.5	4.8
12	V999	MDO-25	2	6.5	4.8

Table 3.2 (C) Genomic copy number of *vlpA* and the DNA fragment size digested with *Eco*RI in different strains of *V. cholerae* non-O1. (Data from Southern hybridization **Figure 3.4C**).

No	Laboratory strain designation	Strain	*vlpA* copy number	Estimated fragment size (kb)		
				1	2	3
1	V1051	O2	2	7.35	6.6	-
2	V1053	O3	0	-	-	-
3	V1054	O4	0	-	-	-
4	V1060	O11	2	4.0	3.6	-
5	V1062	O13	1	7.35	-	-
6	V1063	O14	1	7.35	-	-
7	V1065	O16	2	7.0	6.3	-
8	V1067	O18	2	6.7	6.6	-
9	V1072	O23	1	6.6	-	-
10	V1074	O25	3	6.6	3.7	2.5
11	V1077	O28	2	3.4	2.0	-
12	V1081	O32	2	6.3	2.8	-

Table 3.2 (D) Genomic copy number of *vlpA* and the DNA fragment size digested with *Eco*RI in different strains of *V. anguillarum*. (Data from Southern hybridization **Figure 3.4D**).

No	Laboratory strain designation	Strain	*vlpA* copy number	Estimated fragment size (kb)		
				1	2	3
1	V943	ATTC 43305	1	3.3	-	-
2	V944	ATTC 43306	1	3.4	-	-
3	V947	ATTC 43307	2	6.0	3.7	-
4	V948	ATTC 43308	1	5.5	-	-
5	V949	ATTC 43309	1	3.8	-	-
6	V950	ATTC 43310	3	6.2	5.0	3.0
7	V951	ATTC 43311	1	6.0	-	-
8	V952	ATTC 43312	2	7.0	6.0	-
9	V953	ATTC 43313	1	6.0	-	-
10	V954	ATTC 43314	1	7.0	-	-
11	V957	85-3954-1	1	3.2	-	-
12	V958	85-3954-2	1	3.2	-	-
13	V959	86-3674	1	3.2	-	-
14	V960	89-3748-1	1	6.2	-	-

Table 3.2 (E) The presence of *vlpA* in *V. mimicus*, *V. fluvialis*, and *V. parahaemolyticus*. (Data from Southern hybridization, not shown).

No	Laboratory strain designation	Species	Strain	*vlpA* copy number
1	V563	*V. mimicus*	non-O1	0
2	V800	*V. mimicus*	non-O1	0
3	V1208	*V. mimicus*	M-33	0
4	V1209	*V. mimicus*	M-35	0
5	V1210	*V. mimicus*	N-4459	0
6	V691	*V. parahaemolyticus*	102	0
7	V1211	*V. parahaemolyticus*	AA-3853	0
8	V1212	*V. parahaemolyticus*	X-4844	0
9	V1213	*V. parahaemolyticus*	Y-17233	0
10	V564	*V. fluvialis*	non-O1	0
11	V1214	*V. fluvialis*	AA-15385	0
12	V1215	*V. fluvialis*	AL-1577	0
13	V1216	*V. fluvialis*	AL-14413	0

Figure 3.4 Southern hybridization for *vlpA* in *Vibrio spp.*

Chromosomal DNA from *Vibrio spp.* were cleaved with *Eco*RI (**A, B, D**) or *Eco*RV (**C**) and probed with a PCR DIG-labeled *vlpA*-specific probe.

(A) Lanes contain DNA from strains of *Vibrio spp.*; **1-4**, *V. cholerae* (**1**, 569B; **2**, CA411; **3**, O17; **4**, CA401); **5**, *V. mimicus*; **6**, *V. fluvialis*; **7-11**, *V. cholerae* (**7**, *V. cholerae* O53; **8**, Z17561; **9**, AA14041; **10**, AA13993; **11**, AA14073); **12**, *V. parahaemolyticus*; **13-15**, *V. cholerae* (**13**, C5; **14**, C31; **15**, V911).

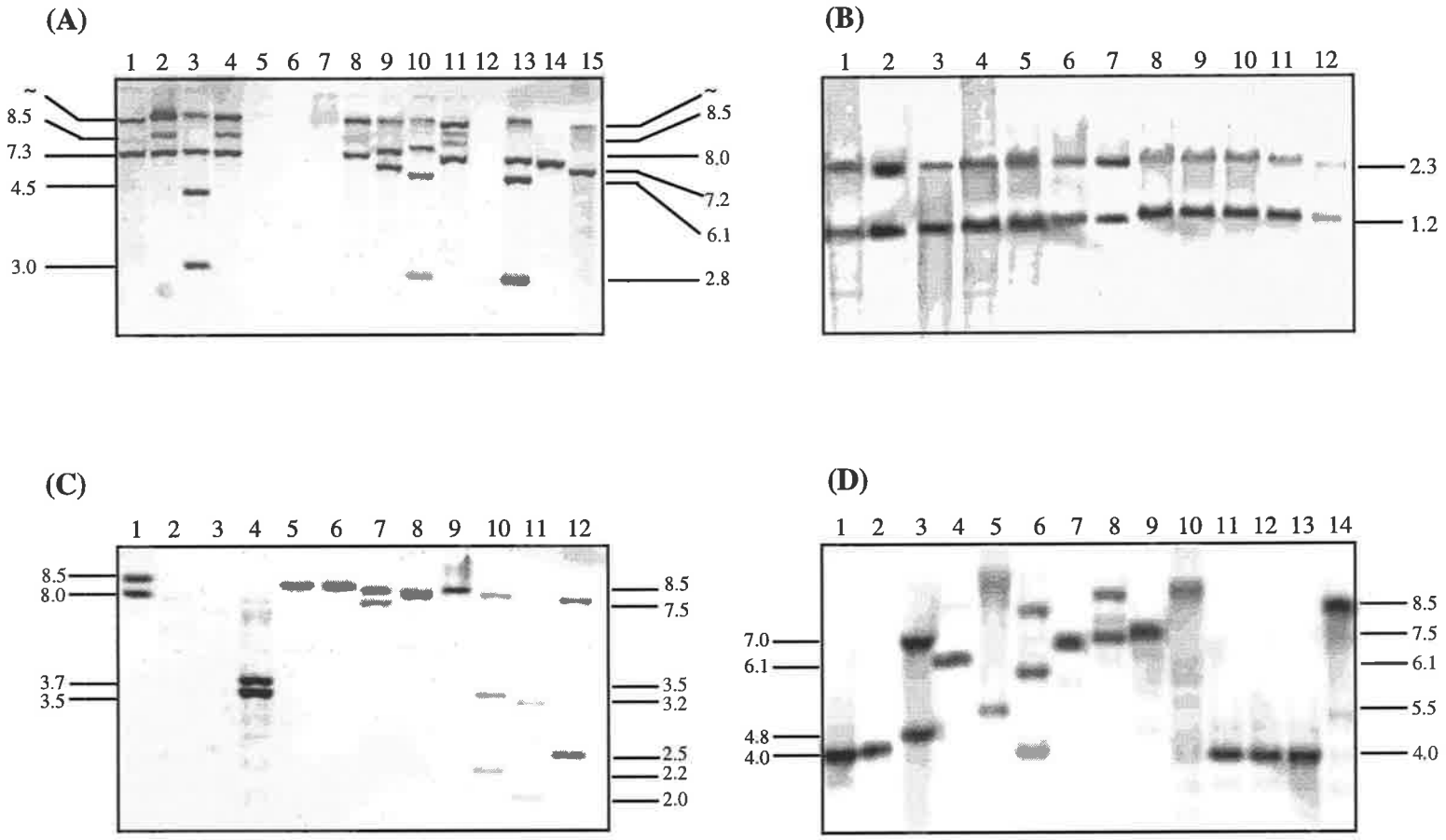
(B) Lanes contain DNA from strains of *V. cholerae* O139; **1**, V911; **2**, V912; **3**, V913; **4**, V914; **5**, V915; **6**, V916; **7**, V917; **8**, V918; **9**, V996; **10**, V997; **11**, V998; **12**, V999.

(C) Lanes contain DNA from strains of *V. cholerae* non-O1; **1**, O2; **2**, O3; **3**, O4; **4**, O11; **5**, O13; **6**, O14; **7**, O16; **8**, O18; **9**, O23; **10**, O25; **11**, O28; **12**, O32.

(D) Lanes contain DNA from strains of *V. anguillarum* : **1**, ATCC43305; **2**, ATCC43306; **3**, ATCC43307; **4**, ATCC43308; **5**, ATCC43309; **6**, ATCC43310; **7**, ATCC43311; **8**, ATCC43312; **9**, ATCC43313; **10**, ATCC43314; **11**, 85-3954-1; **12**, 85-3954-2; **13**, 86-3674; **14**, 89-3748-1.

The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI (in kb). DNA fragment sizes which larger than the markers used are represented by ~.

Figure 3.4


(A)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
~
8.5
7.3
4.5
3.0
~
8.5
8.0
7.2
6.1
2.8
(B)
1 2 3 4 5 6 7 8 9 10 11 12
2.3
1.2
(C)
1 2 3 4 5 6 7 8 9 10 11 12
8.5
8.0
3.7
3.5
8.5
7.5
3.5
3.2
2.5
2.2
2.0
(D)
1 2 3 4 5 6 7 8 9 10 11 12 13 14
7.0
6.1
4.8
4.0
8.5
7.5
6.1
5.5
4.0

Based on the analysis of several genes involved in LPS biosynthesis and the presence of a novel genetic element, Stroeher *et al.*(1998) suggested that *V. cholerae* and *V. anguillarum* are very closely related. In addition, PCR amplification using the internal *vlpA* oligonucleotide primers (#1071 and #1072) showed that the *vlpA* gene is present in all strains of *V. anguillarum*. Therefore, Southern hybridization was performed using the same PCR DIG-labeled *vlpA* specific probe to determine the *vlpA* copy number. The *vlpA* gene is also present as multiple copies in some strains of *V. anguillarum*, with the number ranging from one to three copies (Figure 3.4D).

3.2.2 Cloning fragments containing *vlpA* into the plasmid vector pBluescript.

As *vlpA* is present in *V. cholerae* as multiple copies, a gene library was constructed to clone the different copies of the gene and determine the DNA sequence of the gene and its surrounding DNA. Whole genomic DNA from *V. cholerae* classical strain 569B, El Tor strain O17 and O139 (strain AI-1837 or V911) were partially digested with *Bam*HI, *Eco*RI and *Eco*RV, respectively. These enzymes do not digest within *vlpA*, and the size of DNA fragments containing *vlpA* (as determined from Southern hybridization analysis) is appropriate for insertion into the plasmid vector pBluescript (Stratagene).

Data from Southern hybridization analysis (Figure 3.4A) showed that *V. cholerae* O1 strain 569B (lane 1) contains two hybridizing bands of 6.3 and 4.6-kb using the *vlpA* specific probe when digested with *Bam*HI. Chromosomal DNA was digested with *Bam*HI and ligated into the *Bam*HI site of pBluescript. A 6.3-kb *Bam*HI fragment was previously cloned by Franzon (1988), and shown to contain one copy of *vlpA* (Barker and Manning, 1997). This plasmid was designated pPM471. A 4.6-kb *Bam*HI fragment containing the other copy of *vlpA* was cloned in this study, and identified by colony hybridization using

the PCR DIG-labeled *vlpA* specific probe. Cells from the colonies on the master plate that showed a positive response to the probe were subcultured, and the plasmids isolated. To identify those that carried the appropriate DNA fragment, restriction analysis was used to identify a ca. 4.6-kb *Bam*HI fragment and DNA sequencing was performed to confirm the presence of *vlpA* (see **Section 3.2.3**). This plasmid was designated pPM5001.

A similar approach was used to clone DNA fragments containing *vlpA* from *V. cholerae* El Tor O17. Southern hybridization analysis (Figure 3.4A, lane 3) showed that chromosomal DNA digested with *Eco*RI produced four fragments when probed for *vlpA*. The size of these fragments were 1.0, 4.5, 7.3, and greater than 10-kb. After the digestion with *Eco*RI, these fragments were inserted into the *Eco*RI site of pBluescript. Colonies were screened by colony hybridization using a PCR DIG-labeled *vlpA* probe, and plasmids were isolated. Restriction analysis confirmed whether they contained the *Eco*RI fragments. Attempts to clone these fragments into pBluescript was successful for only one copy. This was the 1.0-kb *Eco*RI fragment which contained only one copy of *vlpA*, and was designated pPM5002. Sequencing was performed to confirm the presence of *vlpA* (see **Section 3.2.3**).

The *vlpA* genes were also cloned from *V. cholerae* O139 (V911). The chromosomal DNA was digested with *Eco*RV and cloned into the *Eco*RV site of pBluescript. Only two hybridization bands are present in V911 (Figure 3.4B). Both *Eco*RV fragments containing *vlpA* were inserted into *Eco*RV sites of pBluescript. Screening for plasmid clones were as described previously for pPM5001 and pPM5002 by colony hybridization using a *vlpA* specific probe. Restriction analysis was used to confirm whether they contained the appropriate *Eco*RV fragments. Both the 2.1-kb *Eco*RV and 1.2-kb *Eco*RV fragments containing *vlpA* were successfully cloned, and designated pPM5003 and pPM5004

respectively. Sequencing was performed to confirm the presence of *vlpA* (see **Section 3.2.3**).

3.2.3 Analysis and variation of the *vlpA* nucleotide sequence in *Vibrio spp.*

3.2.3.1 Nucleotide sequence and analysis of plasmid clones

The entire nucleotide sequence of all the cloned DNA fragments containing the *vlpA* genes (pPM5001, pPM5002, pPM5003 and pPM5004) were determined and the schematic sequence of these fragments are represented in Figure 3.5. The nucleotide sequence of pPM471 has been determined previously (Franzon, 1988), and the DNA sequence has been deposited in the EMBL GenBank database under accession number X64097 (Barker *et al.*, 1994). Interestingly, all the sequenced clones in addition to *vlpA* contained open reading frames flanked by the 124-bp direct repeat sequences (VCRs). The EMBL GenBank database accession numbers for pPM5001 and pPM5002 are AF025662 and AF025663, respectively.

The sequences of all the DNA fragments were compared with entries in the PIR (Release 33) and SWISS-PROT (Release 22) database for similar entries using the BlastX algorithm (Atschul *et al.*, 1997). The 334 amino acids from the first open reading frame (VCO10) in pPM5001 from *V. cholerae* O1 569B showed a significant match with the *mccF* gene on a single copy conjugative *E. coli* plasmid, pMcc7 (56 % identity) (Gonzalez-Pastor, 1995) (Figure 3.6). The *E. coli mccF* is part of an operon that includes the genetic determinants for production, export and immunity to microcin C7 (MccC7), a linear heptapeptide with antibiotic activity (Gonzalez-Pastor, 1995). The *mccF* gene has been designated *EcmccF* for *E. coli*, and *VcmccF* for *V. cholerae* in this study. The amino acid sequences from other open reading frames (named VCO for *V. cholerae* open reading

Figure 3.5 DNA fragments containing *vlpA* from *V. cholerae* O1 569B, El Tor O17 and O139.

(A) pPM471 is a 6.3-kb *Bam*HI DNA fragment containing *vlpA*1 from *V. cholerae* O1 569B (Franzon *et al.*, 1988).

(B) pPM5001 is a 4.8-kb *Bam*HI DNA fragment containing *vlpA*2 from *V. cholerae* O1 569B.

(C) pPM5002 is a 1.0-kb *Eco*RI DNA fragment containing *vlpA*3 from *V. cholerae* El Tor O17.

(D) pPM5003 is a 2.1-kb *Eco*RV DNA fragment containing *vlpA*1 from *V. cholerae* O139 (V911). This fragment shows the similarity of the *vlpA* flanking genes to pPM471.

(E) pPM5003 is a 1.2-kb *Eco*RV DNA fragment containing *vlpA*4 from *V. cholerae* O139 (V911).

The numbered scale indicates the position in kb on the sequence, with the position of some restriction endonuclease recognition sites. ORFs are represented as boxes, designated VCOs with the direction of transcription indicated by horizontal arrows. Hatched boxes represent the 124-bp VCR. The number under the hatched boxes represent the name of VCRs. These DNA fragments have been cloned into the appropriate restriction enzyme sites in the plasmid vector pBluescript.

Figure 3.5

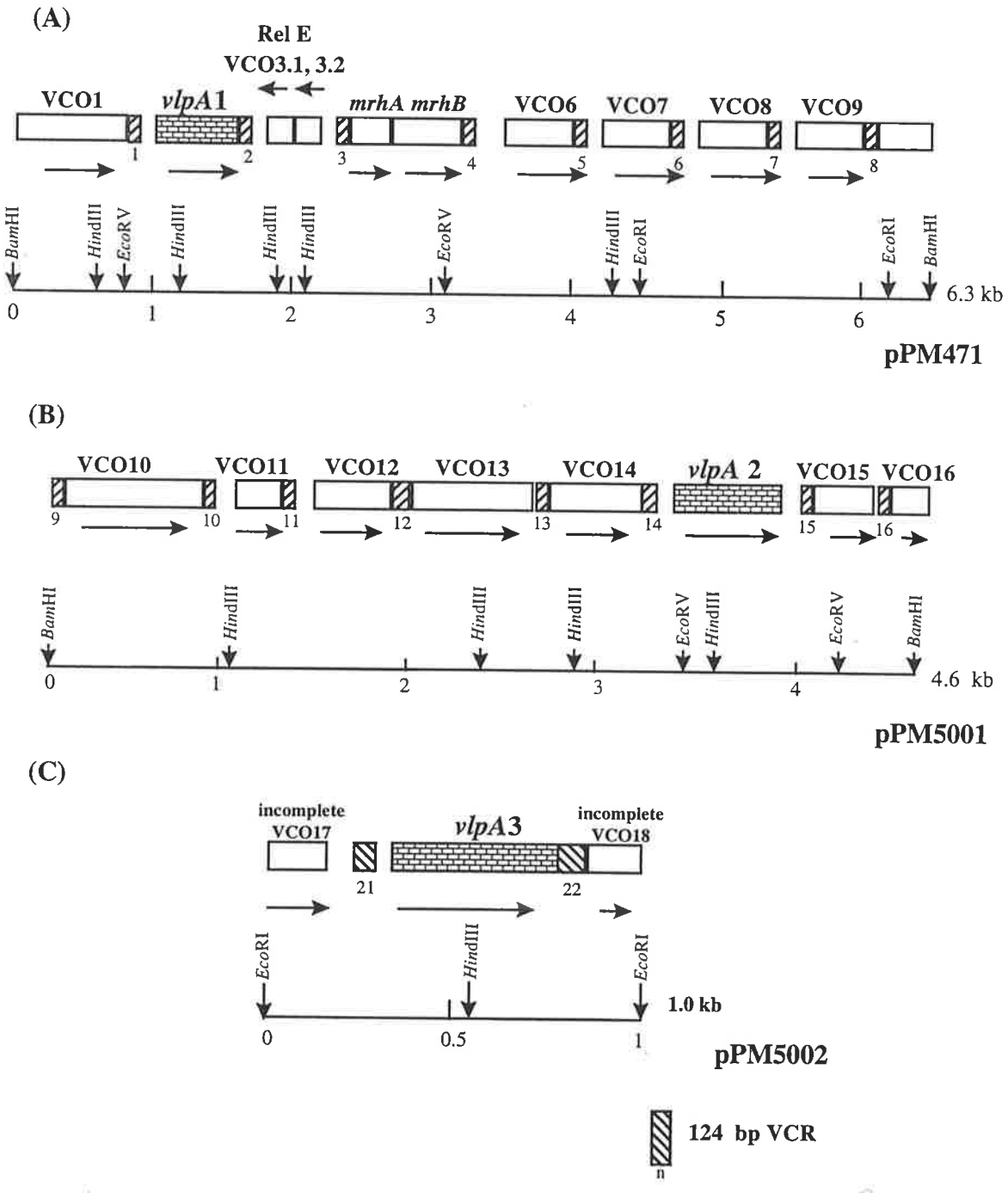

(A)
Rel E
VCO3.1, 3.2
VCO1
vlpA1
mrhA mrhB
VCO6
VCO7
VCO8
VCO9
1
2
3
4
5
6
7
8
BamHI
HindIII
EcoRV
HindIII
HindIII
HindIII
EcoRV
HindIII
EcoRI
EcoRI
BamHI
0
1
2
3
4
5
6
6.3 kb
pPM471
(B)
VCO10
VCO11
VCO12
VCO13
VCO14
vlpA 2
VCO15
VCO16
9
10
11
12
13
14
15
16
BamHI
HindIII
HindIII
HindIII
EcoRV
HindIII
EcoRV
BamHI
0
1
2
3
4
4.6 kb
pPM5001
(C)
incomplete
VCO17
vlpA3
incomplete
VCO18
21
22
EcoRI
HindIII
EcoRI
0
0.5
1
1.0 kb
pPM5002
124 bp VCR
n

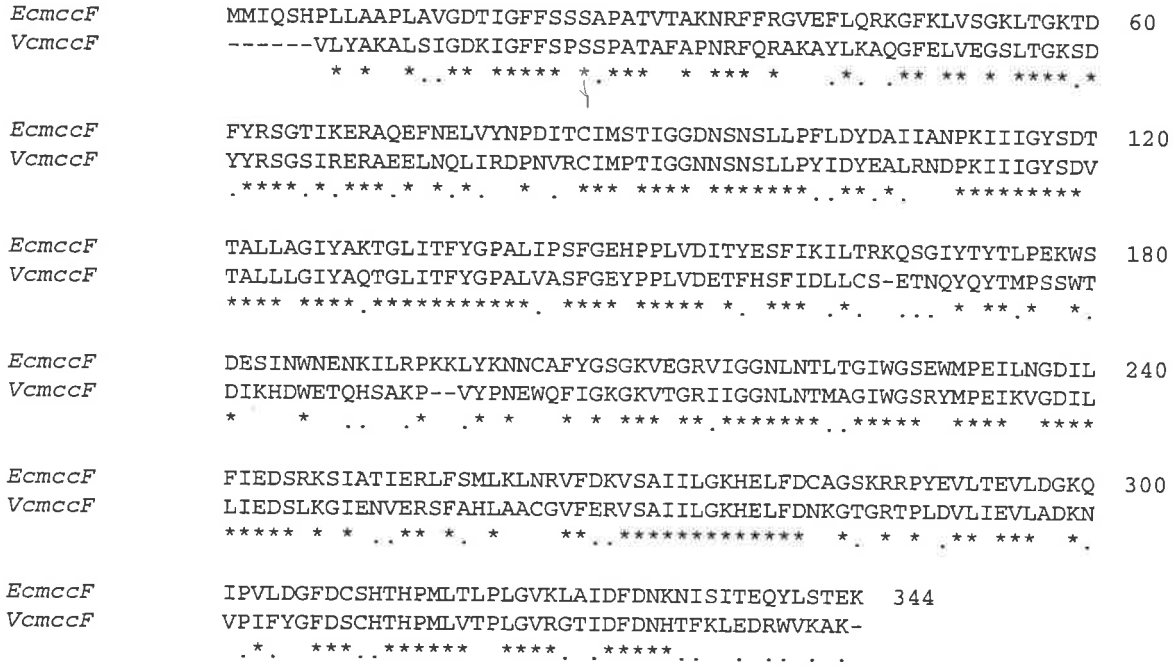
Figure 3.6 Similarity between amino acid from VlpA and MccF.

The amino acid alignment between *EcmccF* gene (*mccF* gene from *E. coli*) and *VcmccF* (ORF1 from pPM5001) using CLUSTAL W program (Thomson *et al.*, 1994; Higgins *et al.*, 1996).

Identical residues are marked with an asterisk, similar residues with a dot.

Identity = 191/336 (56%), Similarity = 244/336, Gaps = 3/336 (0%)

Figure 3.6


EcmccF MMIQSHPLLAAPLAVGDTIGFFSSSAPATVTAKNRFFRGVEFLQRKGFKLVSGKLTGKTD 60
VcmccF ------VLYAKALSIGDKIGFFSPSSPATAFAPNRFQRAKAYLKAQGFELVEGSLTGKSD
EcmccF FYRSGTIKERAQEFNELVYNPDITCIMSTIGGDNSNSLLPFLDYDAIIANPKIIIGYSDT 120
VcmccF YYRSGSIRERAEELNQLIRDPNVRCIMPTIGGNNSNSLLPYIDYEALRNDPKIIIGYSDV
EcmccF TALLAGIYAKTGLITFYGPALIPSFGEHPPLVDITYESFIKILTRKQSGIYTYTLPEKWS 180
VcmccF TALLLGIYAQTGLITFYGPALVASFGEYPPLVDETFHSFIDLLCS-ETNQYQYTMPSSWT
EcmccF DESINWNENKILRPKKLYKNNCAFYGSGKVEGRVIGGNLNTLTGIWGSEWMPEILNGDIL 240
VcmccF DIKHDWETQHSAKP--VYPNEWQFIGKGKVTGRIIGGNLNTMAGIWGSRYMPEIKVGDIL
EcmccF FIEDSRKSIATIERLFSMLKLNRVFDKVSAIILGKHELFDCAGSKRRPYEVLTEVLDGKQ 300
VcmccF LIEDSLKGIENVERSFAHLAACGVFERVSAIILGKHELFDNKGTGRTPLDVLIEVLADKN
EcmccF IPVLDGFDCSHTHPMLTLPLGVKLAIDFDNKNISITEQYLSTEK 344
VcmccF VPIFYGFDSCHTHPMLVTPLGVRGTIDFDNHTFKLEDRWVKAK-

frame), VCO11, VCO12 and VCO13 in pPM5001 showed significant matches with the hypothetical 17.1-kDa protein in the TAG-BISC intergenic region in *E. coli* (33% identity) (GenBank accession number P37664), the sensor histidine kinase in *Deinococcus radiodurans* (strain R1) (21% identity) (GenBank accession number E75617), and the hypothethical protein Rv0911 in *Mycobacterium tuberculosis* (strain H37RV) (GenBank accession number G70581) (22% identity) respectively. No homologies were observed with the remaining VCOs in pPM5001.

The 2.1-kb sequence of pPM5003 from *V. cholerae* O139 was also compared for similar entries using the BlastX algorithm. This sequence showed 98 % similarity to the sequence of pPM471 from *V. cholerae* 569B (GenBank accession number X64097; Manning, 1992; Franzon *et al.*, 1993). The identity of the sequence was observed from the *vlpA*, VCO3.1, VCO3.2, *mrhA* and *mrhB* genes (Figure 3.7A). The sequences were analyzed by using the CLUSTAL W program (Thomson *et al.*, 1994; Higgins *et al.*, 1996) and the alignment is shown in Figure 3.7B. This indicates that there is a similar locus in the *V. cholerae* O139 genome, containing *vlpA* and flanking genes, to that of the *V. cholerae* O1 classical biotype 569B. The 1.0-kb sequence of pPM5002 from *V. cholerae* El Tor O17 did not correspond to any sequence in both *V. cholerae* O1 classical and O139, or to any entries in the database using the BlastX algorithm. This was also case for pPM5004. However, these sequence are present in the *V. cholerae* El Tor strain N16961 of the TIGR database (see **Section 3.2.6** and Figure 3.14).

The oligonucleotide primers #3141 and #3142 were designed from pPM5004 (V911 clone) at each end of ORF VCO19 and VCO20 (Figure 3.5E) to determine if a PCR product could be amplified from the *V. cholerae* O1 chromosome with primers from genes flanking *vlpA*. PCR amplification produced a PCR product of 1.0-kb in *V. cholerae* El Tor

Figure 3.7 Comparison of pPM471 (from *V. cholerae* O1 569B), and pPM5003 (from *V. cholerae* O139).

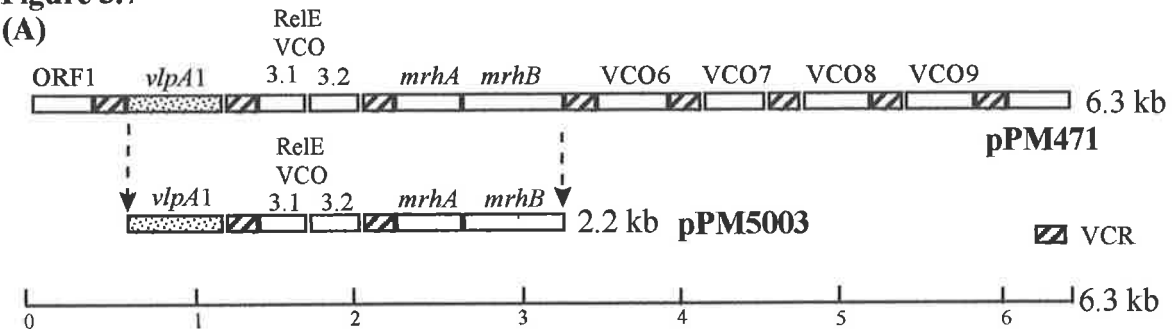
(A) Diagramatic represent of pPM471 and pPM5003. The open reading frames (VCOs) are represented by boxes. Hatched boxes represent the 124-bp VCR.

(B) Homology alignment between pPM471 and pPM5003. Identical sequences are joined by a solid line (|). *vlpA* and VCR sequences are printed in red and blue, respectively. Different bases are marked in bold type and underlined.

I dentities = 2089/2114 (98%), Gaps = 2/2114 (0%).

Figure 3.7

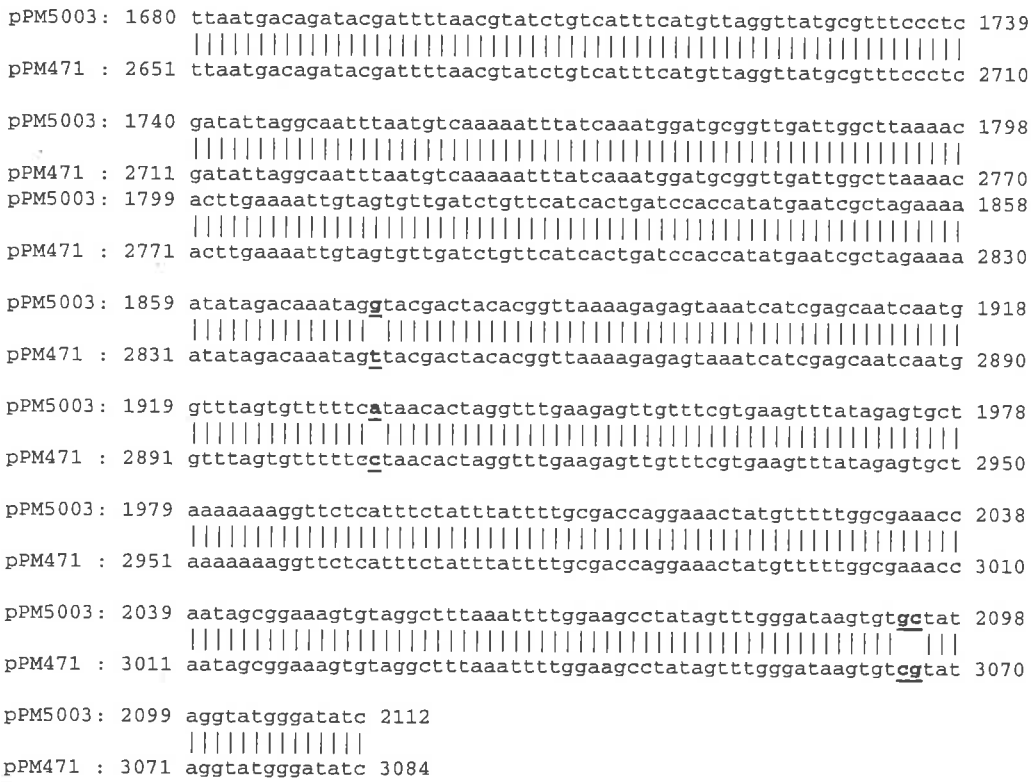
(A)



(B)

```
pPM5003: 1    gatatcatgagagctatctttttgattctttgctctgttttattaaatggctgcttgggc 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 971  gatatcatgagagctatctttttgattctttgctctgttttattaaatggctgcttgggc 1030

pPM5003: 61   atgcccgaatcagtaaaaccagtgtcggattttgaactgaacaactatttaggcaaatgg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1031 atgcccgaatcagtaaaaccagtgtcggattttgaactgaacaactatttaggcaaatgg 1090

pPM5003: 121  tacgaggttgctcgactcgatcactcctttgaaagaggtttaagtcaggttactgcggaa 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1091 tacgaggttgctcgactcgatcactcctttgaaagaggtttaagtcaggttactgcggaa 1150

pPM5003: 181  taccgtgctcgaaatgatggtggtatttcggttcttaatcgtggttattctgaagagaaa 240
                ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1151 taccgtgttcgaaatgatggtggtatttcggttcttaatcgtggttattctgaagagaaa 1210

pPM5003: 241  ggtgagtggaaggaagcggaaggtaaagcttacttttgtgaatggctcaacagatggctat 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1211 ggtgagtggaaggaagcggaaggtaaagcttacttttgtgaatggctcaacagatggctat 1270

pPM5003: 301  ctgaaggtttcatttttttggcccgttttatggctcctacgtagtgtttgagttagaccgt 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1271 ctgaaggtttcatttttttggcccgttttatggctcctacgtagtgtttgagttagaccgt 1330

pPM5003: 361  gaaaactacagttatgcttttgtgtcagggccgaatacagaatatctgtggttactttca 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1331 gaaaactacagttatgcttttgtgtcagggccgaatacagaatatctgtggttactttca 1390

pPM5003: 421  agaacgccgactgtagaacgaggcattctggacaagttcatagaaatgtcgaaagagcgt 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1391 ggttttgatacaaatcggcttattttacgttcagctgcaataaatacttaacaaaggcccc 1450

pPM5003: 541  aagagggactgtcaacgcgtggcgtttccagtcccattgagccgcggtggtttcggttgt 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1511 aagagggactgtcaacgcgtggcgtttccagtcccattgagccgcggtggtttcggttgt 1570

pPM5003: 601  tgtgtttgaattgagtggtaatacgttgccagccccttaggcgggcgttaaacttcatca 660
                ||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1571 tgtgtttgagttagtggtaatacgttgccagccccttaggcgggcgttaaacttcatca 1630

pPM5003: 661  gtcgtctagccgttgcagcgctttagtgtaaacttcgctgcgttctcgtttaccgacagc 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1631 gtcgtctagccgttgcagcgctttagtgtaaacttcgctgcgttctcgtttaccgacagc 1690
```

```
pPM5003: 721  taaaacagtaactacgatgatgtcattttcaacttgatagactaagcgataaccagactg 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1691 taaaacagtaactacgatgatgtcattttcaacttgatagactaagcgataaccagactg 1750

pPM5003: 781  ccgaagttttatttataaatgttctcagcgccagacagtttggctgatggaacatgtgg 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1751 ccgaagttttatttataaatgttctcagcgccagacagtttggctgatggaacatgtgg 1810

pPM5003: 841  attttcaaggcgttctattagtttcttcttgaactgctgttggagtggaacggctaactt 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1811 attttcaaggcgttctattagtttcttcttgaactgctgttggagtggaacggctaactt 1870

pPM5003: 901  tttccattccttcaaagcactctttttgaactcaagcttataggtcatcgatatttaccg 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1871 tttccattccttcaaagcactctttttgaactcaagcttataggtcatcgatatttaccg 1930

pPM5003: 961  atacgctgtcttcacttaagcgttctttggcaatagccagtagctcaagatcttcaagct 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 1931 atacgctgtcttcacttaagcgttctttggcaatagccagtagctcaagatcttcaagct 1990

pPM5003: 1021 tatccatcatgatctcgtaagtgcttgcagggacgcagtaaaaagccggctcatttcggt 1080
              |||||||||||||||||| |||||||||||||||||| |||||||||||||||||||
pPM471 : 1991 tatccatcatgatctcgtaggtgcttgcagggacgcagtaggaaagccggctcatttcggt 2050

pPM5003: 1081 tgagtacagcaacaggagcaccgaaggcaccagttgcaactttcattgggtttgctttaa 1140
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
pPM471 : 2051 tgagtacagcaacaggagcaccgaaggcactagttgcaactttcattgggtttgctttaa 2110

pPM5003: 1141 attcggtaatgcttgctgcaacatcagctaggattctagtggtcatttaatcggtccttt 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2111 attcggtaatgcttgctgcaacatcagctaggattctagtggtcatttaatcggtccttt 2170

pPM5003: 1201 atgtggtctttaattgggtcattatagtcttgaaaatgcagtttaacaaacgcctcaaga 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2171 atgtggtctttaattgggtcattatagtcttgaaaatgcagtttaacaaacgcctcaaga 2230

pPM5003: 1261 gggactgtcaacgcgcggcgtttccagtcccaatgagccgcggtggttgcagtcgttgtg 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
pPM471 : 2231 gggactgtcaacgcgcggcgtttccagtcccaatgagccgcggtggttgcagttgttgtg 2290

pPM5003: 1321 tttgggtttagtgtcatgcgttgccagccccttaggcgggcgttaggcataaaatttata 1380
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2291 tttgggtttagtgttatgcgttgccagccccttaggcgggcgttaggcataaaatttata 2350

pPM5003: 1381 agcaattgcttgatattctccatattgcccatatttaaattatggtcattaaataaaagg 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2351 agcaattgcttgatattctccatattgcccatatttaaattatggtcattaaataaaagg 2410

pPM5003: 1441 agtagaccatatgccagatttaacgaaagctcggtgtccgtgtcgtggatgcaataaagt 1500
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
pPM471 : 2411 agtagaccatatgccagatttaacgaaagctcgctgtccgtgtcgtggatgcaataaagt 2470

pPM5003: 1501 tgcgaacaatgcagaagaggtagaagagctgtttggcctcagaaatatgggtgatggtac 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2471 tgcgaacaatgcagaagaggtagaagagctgtttggcctcagaaatatgggtgatggtac 2530

pPM5003: 1561 gattcgcgtacaatcatattgtcgcgaatgtcgaggcttgcattgtgaagctagtaatcc 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2531 gattcgcgtacaatcatattgtcgcgaatgtcgaggcttgcattgtgaagctagtaatcc 2590

pPM5003: 1621 taaatgtaatcatcatgcctaggtattttgatgatttgttagataagaaaatgatatagg 1679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM471 : 2591 taaatgtaatcatcatgcctaggtattttgatgatttgttagataagaaaatgatatagg 2650
```


pPM5003: 1680 ttaatgacagatacgattttaacgtatctgtcatttcatgttaggttatgcgtttccctc 1739
pPM471 : 2651 ttaatgacagatacgattttaacgtatctgtcatttcatgttaggttatgcgtttccctc 2710
pPM5003: 1740 gatattaggcaatttaatgtcaaaaatttatcaaatggatgcggttgattggcttaaaac 1798
pPM471 : 2711 gatattaggcaatttaatgtcaaaaatttatcaaatggatgcggttgattggcttaaaac 2770
pPM5003: 1799 acttgaaaattgtagtgttgatctgttcatcactgatccaccatatgaatcgctagaaaa 1858
pPM471 : 2771 acttgaaaattgtagtgttgatctgttcatcactgatccaccatatgaatcgctagaaaa 2830
pPM5003: 1859 atatagacaaatagGtacgactacacggttaaaagagagtaaatcatcgagcaatcaatg 1918
pPM471 : 2831 atatagacaaatagTtacgactacacggttaaaagagagtaaatcatcgagcaatcaatg 2890
pPM5003: 1919 gtttagtgtttttcAtaacactaggtttgaagagttgtttcgtgaagtttatagagtgct 1978
pPM471 : 2891 gtttagtgtttttcCtaacactaggtttgaagagttgtttcgtgaagtttatagagtgct 2950
pPM5003: 1979 aaaaaaaggttctcatttctatttattttgcgaccaggaaactatgtttttggcgaaacc 2038
pPM471 : 2951 aaaaaaaggttctcatttctatttattttgcgaccaggaaactatgtttttggcgaaacc 3010
pPM5003: 2039 aatagcggaaagtgtaggctttaaattttggaagcctatagtttgggataagtgtGCtat 2098
pPM471 : 3011 aatagcggaaagtgtaggctttaaattttggaagcctatagtttgggataagtgtCGtat 3070
pPM5003: 2099 aggtatgggatatc 2112
pPM471 : 3071 aggtatgggatatc 3084

strain O17 (Figure 3.8A). This PCR product was then sequenced and analyzed using the CLUSTAL W program. The alignment between sequences from pPM5004 and the PCR product revealed a similarity of 97% (Figure 3.8B). This result shows that at least one chromosomal copy of *vlpA* in *V. cholerae* O139 and El Tor strain O17 have the same sequence and flanking genes.

3.2.3.2 The *vlpA* nucleotide sequence variation in *V. cholerae* and *V. anguillarum*.

The *vlpA* DNA sequences from pPM471, pPM5001, pPM5002, pPM5003 and pPM5004 were analyzed using the CLUSTAL W program to compare the identity between *vlpA* in different strains of *V. cholerae* (Figure 3.9). A proposed consensus sequence was derived from the alignment of the nine copies of *vlpA* sequenced from these plasmid clones, *V. cholerae* non-O1 and *V. anguillarum* (Figure 3.10). The overall percentage of the identity between *vlpA* nucleotide consensus and *vlpA* nucleotide sequences from plasmid clones is greater than 96%, with the identity within strains varied from 97-99% (see Table 3.3A) indicating that this gene is highly conserved.

In order to determine the variation in *vlpA* DNA sequences between *V. cholerae* serogroups and other species, the PCR products were amplified in different strains of *V. cholerae* O1, O139, non-O1, and *V. anguillarum* using the internal *vlpA*-specific oligonucleotide primers (#1071 and #1072). The nucleotide sequences of the 516-bp coding region obtained from the PCR products were analyzed using the CLUSTAL W program. The multiple alignment is shown in Figure 3.10. Interestingly, the sequence of *vlpA* in all strains of *V. cholerae* (O1 and non-O1) and *V. anguillarum* tested showed an overall identity of more than 94% to each other. The average pairwise percentage

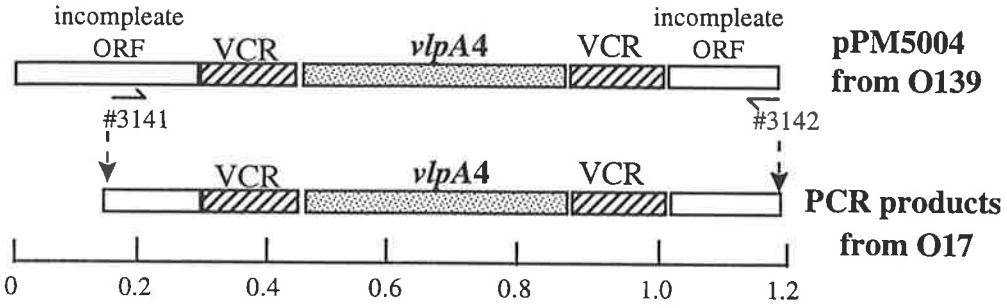
Figure 3.8 Similarity between the *vlpA* sequence from *V. cholerae* El Tor O17 and pPM5004 from *V. cholerae* O139 (V911).

(A) Comparison of the sequence of PCR product from *V. cholerae* O17 using oligonucleotide primers (#3141 and #3142) generated from pPM5004. The open reading frames (VCOs) are represented by boxes. Hatched boxes represent the 124-bp VCR.

(B) Homology alignment between the PCR product generated from *V. cholerae* O17 and pPM5004. Identical sequences are joined by a solid line (|). (97.0% identity in 785 bp overlap).

Figure 3.8

(A)



(B)

```
O17          CGTATGAAAATGCA-GCTATCAGCATGAAGGTATCAAAAGAGAATCGGGCTATCGCAATG
             ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
pPM5004      CGTATGAAAATGCAGGCTATCAGCATGAAGGTATCAAAAGAAAATCGGGCTATCGCAATG
                 290        300        310        320        330        340
             60         70         80         90         100        110
O17          GCCGATTCATGGACAAAGTCCAAATGTCCGTTTTATCGCGAGAGTGGCCGGCCACATAAC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM5004      GCCGATTCATGGACAAAGTCCAAATGTCCGTTTTATCGCGAGAGTGGCCGGCCACATAAC
                 350        360        370        380        390        400
             120        130        140        150        160        170
O17          AAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGTGGT
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM5004      AAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGTGGT
                 410        420        430        440        450        460
             180        190        200        210        220        230
O17          TTCGGTTGTTGAGGTTGAGTTTGGTGTTAATGCGTTGCCAGCCCCTTAGGCGGGCGTTAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM5004      TTCGGTTGTTGAGGTTGAGTTTGGTGTTAATGCGTTGCCAGCCCCTTAGGCGGGCGTTAG
                 470        480        490        500        510        520
             240        250        260        270        280        290
O17          GTATTTGGAGACATCATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM5004      GTATTTGGAGACATCATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGC
                 530        540        550        560        570        580
             300        310        320        330        340        350
O17          TGCTTGGGCATGCCCGAATCAGTAAAACCGGTGTCGGATTTTGAACTGAACAACTATTTA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pPM5004      TGCTTGGGCATGCCCGAATCAGTAAAACCGGTGTCGGATTTTGAACTGAACAACTATTTA
                 590        600        610        620        630        640
             360        370        380        390        400        410
O17          GGCAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAAGTTTAAGTCAGGTT
             |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
pPM5004      GGCAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTT
                 650        660        670        680        690        700
             420        430        440        450        460        470
O17          ACTGCGGAATACCGTGTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCT
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
pPM5004      ACTGCGGAATACCGTGTTGGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCT
                 710        720        730        740        750        760
             480        490        500        510        520        530
O17          GAAGAGAAAGGTGAGTGGAAGGAA-CGGAAGGCAAAGCTTACTTTGTGAATGGCTCAACA
             ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
pPM5004      GAAGAGAAAGGTGAGTGGAAGGAAGCGGAAGGCAAAGCTTACTTTGTGAATGGCTCAACA
                 770        780        790        800        810        820
```

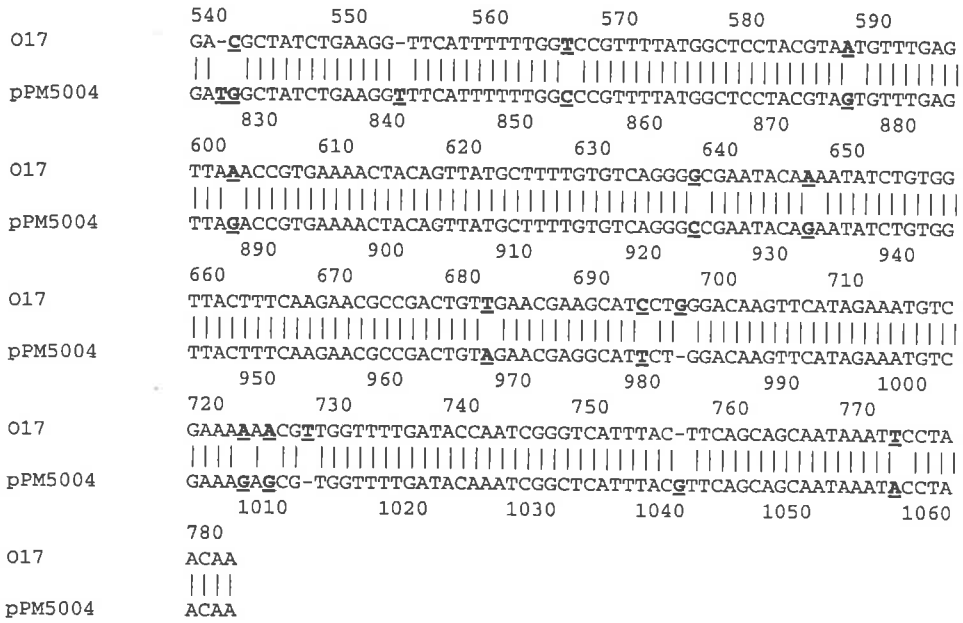

540 550 560 570 580 590
O17 GA-CGCTATCTGAAGG-TTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAATGTTTGAG
pPM5004 GATGGCTATCTGAAGGTTTCATTTTTTGGCCCGTTTTATGGCTCCTACGTAGTGTTTGAG
830 840 850 860 870 880
600 610 620 630 640 650
O17 TTAAACCGTGAAAACTACAGTTATGCTTTTGTGTCAGGGGCGAATACAAAATATCTGTGG
pPM5004 TTAGACCGTGAAAACTACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGG
890 900 910 920 930 940
660 670 680 690 700 710
O17 TTACTTTCAAGAACGCCGACTGTTGAACGAAGCATCCTGGGACAAGTTCATAGAAATGTC
pPM5004 TTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCT-GGACAAGTTCATAGAAATGTC
950 960 970 980 990 1000
720 730 740 750 760 770
O17 GAAAAAACGTTGGTTTTGATACCAATCGGGTCATTTAC-TTCAGCAGCAATAAATTCCTA
pPM5004 GAAAGAGCG-TGGTTTTGATACAAATCGGCTCATTTACGTTCAGCAGCAATAAATACCTA
1010 1020 1030 1040 1050 1060
780
O17 ACAA
pPM5004 ACAA

Figure 3.9 CLUSTAL W analysis (Thomson *et al.*, 1994; Higgins *et al.*, 1996) of *vlpA* in different strains of *V. cholerae* (pPM471 and pPM5001 from *V. cholerae* O1 strain 569B, pPM5002 from El Tor strain O17, pPM5003 and pPM5004 from O139). Identical sequences are marked with an asterisk, similar residues with a dot, and different residues are printed in red.

Figure 3.9

```
pPM471          ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCC  60
pPM5001         ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCC  60
pPM5002         ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTT-TTAAATGGCTGCTTGGGCATGCCC  60
pPM5003         ATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGCTGCTTGGGCATGCCC  60
pPM5004         ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCC  60
                ************************** ******* ** ********************

pPM471          GAATCAGTAAAACCAGTGTCGGATTTTGAACTGAACAACTATTTAGGCAAATGGTACGAG  120
pPM5001         GAATCAGTAAAACCAGTGTCGGATTTTGAACTGAACAACTATTTAGGTAAATGGTACGAA  120
pPM5002         GAATCAGTAAAACCAGTGTCGGATTTTGAACTGAACAACTATTTAGGTAAATGGTACGAA  120
pPM5003         GAATCAGTAAAACCGGTGTCGGATTTTGAACTGAACAACTATTTAGGCAAATGGTACGAG  120
pPM5004         GAATCAGTAAAACCAGTGTCGGATTTTGAACTGAACAACTATTTAGGCAAATGGTACGAG  120
                ************** ******************************** ***********

pPM471          GTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
pPM5001         GTCGCTCGACTCGATCACTCCTTTGAAAAAGGTTTAAGTCAGGTTACAGCGGAATACCGT  180
pPM5002         GTCGCTCGACTCGATCACTCCTTTGAAAAAGGTTTAAGTCAGGTTACAGCGGAATACCGT  180
pPM5003         GTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
pPM5004         GTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
                ** ************************* ****************** ************

pPM471          GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAGGTGAG  240
pPM5001         GTTCGAAATGATGGTGGCATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAGGTGAG  240
pPM5002         GTTCGAAATGATGGTGGCATTTTGGTTCTTAATCGTGGTTATTCTGAAGAGAAAGGTGAG  240
pPM5003         GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAGGTGAG  240
pPM5004         GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAGGTGAG  240
                ***************** **** ************************************

pPM471          TGGAAGGAAGCGGAAGGTAAAGCTTACTTTGTGAATGGCTCAACAGATGGCTATCTGAAG  300
pPM5001         TGGAAGGAAGCCGAAGGCAAAGCTTACTTTGTGAATGGCTCAACAGATGGCTATCTGAAG  300
pPM5002         TGGAAGGAAGCCGAAGGCAAAGCTTACTTTGTGAATGGCTCAACAGACGGCTATCTGAAG  300
pPM5003         TGGAAGGAAGCGGAAGGCAAAGCTTACTTTGTGAATGGCTCAACAGACGGCTATCTGAAG  300
pPM5004         TGGAAAGAAGCGGAAGGCAAAGCTTACTTTGTGAATGGCTCAACAGACGGCTATCTGAAG  300
                ***** ***** ***** ***************************** ************

pPM471          GTTTCATTTTTTGGCCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
pPM5001         GTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
pPM5002         GTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
pPM5003         GTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
pPM5004         GTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
                ************** ********************************************

pPM471          TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACG  420
pPM5001         TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACG  420
pPM5002         TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACG  420
pPM5003         TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACG  420
pPM5004         TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACG  420
                ************************************************************

pPM471          CCGACTGTAGAACGAGGCATTCTGGACAAGTTCATAGAAATGTCGAAAGAGCGTGGTTTT  480
pPM5001         CCGACTGTAGAACGAGGCATTCTGGACAAGTTCATAGAAATGTCGAAAGAGCGTGGTTTT  480
pPM5002         CCGACTGTAGAACGAGGCATTCTGGACAAGTTCATAGAAATGTCGAAAGAGCGTGGTTTT  480
pPM5003         CCGACTGTAGAACGAGGCATTCTGGACAAGTTCATAGAAATGTCGAAAGAGCGTGGTTTT  480
pPM5004         CCGACTGTAGAACGAGGCATTCTGGACAAGTTCATAGAAATGTCGAAAGAGCGTGGTTTT  480
                ************************************************************

pPM471          GATACAAATCGGCTTATTTACGTTCAGCTGCAATAA  516
pPM5001         GATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
pPM5002         GATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
pPM5003         GATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
pPM5004         GATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
                ************** ************ *******
```

Table 3.3 (A) Identity between *vlpA* nucleotide consensus (from Figure 3.10) and *vlpA* sequences in *V. cholerae* O1 (from pPM471, pPM5001, pPM5002), and O139 (from pPM5003, pPM5004).

Sources of *vlpA*	*vlpA* consensus	pPM471 (569B-1)	pPM5001 (569B-2)	pPM5002 (O17)	pPM5003 (O139-1)	pPM5004 (O139-2)
vlpA consensus	100%	98.64%	99.03%	98.44%	99.41%	98.64%
pPM 471 (569B-1)		100%	97.86%	99.80%	97.67%	97.86%
pPM 5001 (569B-2)			100%	97.86%	98.64%	98.64%
pPM 5002 (O17)				100%	98.06%	97.48%
pPM 5003 (O139-1)					100%	99.22%
pPM 5004 (O139-2)						100%

Table 3.3 (B) Identity between *vlpA* nucleotide consensus (from Figure 3.10) and *vlpA* sequences in *V. cholerae* non-O1 (O14 and O38) and *V. anguillarum* (ATCC43306, ATCC43307, 85-3954-1).

Sources of *vlpA*	*vlpA* consensus	*V. cholerae* O14	*V. cholerae* O38	ATCC 43306	ATTCC 43307	85-3954-1
vlpA consensus	100%	98.25%	98.83%	96.51%	95.34%	99.22%
V. cholerae O14		100%	98.83%	94.76%	93.99%	98.44%
V. cholerae O38			100%	95.73%	94.76%	98.83%
ATCC 43306				100%	97.28%	95.93%
ATCC 43307					100%	94.96%
85-3954-1						100%

Figure 3.10 Clustal W multiple sequence alignment of the nucleotide sequence of *vlpA* in *Vibrio spp*.

The alignment was performed with the multiple-alignment program CLUSTAL. Bases which are not identical to the corresponding base in a proposed consensus sequence (derived from these nine copies of *vlpA*) are in red letters, and bases identical among all sequenced *vlpA* are indicated by an asterisk. [vlpA1-2= *V.cholerae* non-O1 no.O14 and O38; vlpA3-5 = *V. Anguillarum* ATCC 43306; ATCC 43307 and 85-3954-1; vlpA6-8 = *V.cholerae* O1, 6 = O17, 7 = 569B (from pPM471); 8 = 569B (from pPM5001); vlpA 9 = *V. cholerae* O139 (from pPM5003)].

Figure 3.10 The *vlpA* nucleotide sequence identity in *Vibrio spp.*

```
vlpA1       ATGAGAGCTATCTTTTCGATTCTTTGCTCTGTTCTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA2       ATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA3       ATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA4       ATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA5       ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA6       ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA7       ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA8       ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTAAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90
vlpA9       ATGAGAGCTATCTTTTTGATTCTTTGCGCTGTTTTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCGGTGTCGGATTTTGAA  90
            ********** ********** ***** **** ********************************** **********
CONSENSUS   ATGAGAGCTATCTTTTTGATTCTTTGCTCTGTTTTATTGAATGGCTGCTTGGGCATGCCCGAATCAGTAAAACCAGTGTCGGATTTTGAA  90

vlpA1       CTGAACAACTATTTAGGCAAATGGTACGAAGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGGATACCGT  180
vlpA2       CTGAACAACTATTTAGGTAAATGGTACGAAGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
vlpA3       CTGAACAACTATTTAGGTAAATGGTACGAGGTTGCTCGACTTGATCACTCTTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
vlpA4       CTGAACAACTATTTAGGTAAATGGTACGAGGTTGCTCGACTTGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
vlpA5       CTGAACAACTATTTAGGCAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
vlpA6       CTGAACAACTATTTAGGTAAATGGTACGAAGTCGCTCGACTCGATCACTCCTTTGAAAAAGGTTTAAGTCAGGTTACAGCGGAATACCGT  180
vlpA7       CTGAACAACTATTTAGGTAAATGGTACGAAGTCGCTCGACTCGATCACTCCTTTGAAAAAGGTTTAAGTCAGGTTACAGCGGAATACCGT  180
vlpA8       CTGAACAACTATTTAGGCAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
vlpA9       CTGAACAACTATTTAGGCAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180
            ********** ********** ** ******** ******** ******* ********** ********** **** *******
CONSENSUS   CTGAACAACTATTTAGGTAAATGGTACGAGGTTGCTCGACTCGATCACTCCTTTGAAAGAGGTTTAAGTCAGGTTACTGCGGAATACCGT  180

vlpA1       GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGACGCGGAAGGCAAAGCTTACTTT  270
vlpA2       GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGATAAAAGGTGAGTGGAAGGACGCGGAAGGCAAAGCTTACTTT  270
vlpA3       GTTCGAAATGATGGCGGTGTTTCGGTTCTTAATCGTGGTTATTCCGAAGAGAAAAGGTGAGTGGAAAGAGGCTGAAGGCAAAGCTTACTTT  270
vlpA4       GTTCGAAATGATGGCGGCGTTTCGGTTCTTAATCGTGGTTATTCCGAAGAGAAAAGGTGAGTGGAAAGAGGCTGAAGGCAAAGCTTACTTT  270
vlpA5       GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCCGAAGGCAAAGCTTACTTT  270
vlpA6       GTTCGAAATGATGGTGGCATTTTGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCCGAAGGCAAAGCTTACTTT  270
vlpA7       GTTCGAAATGATGGTGGCATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCCGAAGGCAAAGCTTACTTT  270
vlpA8       GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCGGAAGGTAAAGCTTACTTT  270
vlpA9       GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCGGAAGGCAAAGCTTACTTT  270
            ********** **  *** ********** ***** ********** ** ** ***** **********
CONSENSUS   GTTCGAAATGATGGTGGTATTTCGGTTCTTAATCGTGGTTATTCTGAAGAGAAAAGGTGAGTGGAAGGAAGCGGAAGGCAAAGCTTACTTT  270

vlpA1       GTGGATGGCTCAACAGATGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA2       GTGGATGGCTCAACAGATGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA3       GTAAATGGTCCAACAGACGGTTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCTTACGTGGTGTTTGAGTTAGAGCGTGAAAAC  360
vlpA4       GTAAATGGTCCAACAGACGGTTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGAGCGTGAAAAC  360
vlpA5       GTGAATGGCTCAACAGACGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA6       GTGAATGGCTCAACAGACGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA7       GTGAATGGCTCAACAGATGGCTATCTGAAGGTTTCATTTTTTGGCCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA8       GTGAATGGCTCAACAGATGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
vlpA9       GTGAATGGCTCAACAGACGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
            **  **** ******* ** ********************** ********** ***** ********** *********
CONSENSUS   GTGAATGGCTCAACAGACGGCTATCTGAAGGTTTCATTTTTTGGTCCGTTTTATGGCTCCTACGTAGTGTTTGAGTTAGACCGTGAAAAC  360
```

Figure 3.10 The *vlpA* nucleotide sequence identity in *Vibrio spp* (continued).

```
vlpA1       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA2       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA3       TACAGTTATGCCTTCGTATCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCTACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA4       TACAGTTATGCCTTCGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGAAAGTCGCACCAGACATTCTGGAGAAGT  450
vlpA5       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA6       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA7       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA8       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
vlpA9       TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450
            ********** ** ** ****************************************** *  ** * ** ** ********* ****
CONSENSUS   TACAGTTATGCTTTTGTGTCAGGGCCGAATACAGAATATCTGTGGTTACTTTCAAGAACGCCGACTGTAGAACGAGGCATTCTGGACAAGT  450

vlpA1       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
vlpA2       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
vlpA3       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
vlpA4       TCATTCAAATGTCGAAAGACCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
vlpA5       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
vlpA6       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
vlpA7       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCAGCAATAA  516
vlpA8       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTTATTTACGTTCAGCTGCAATAA  516
vlpA9       TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
            **** ************************************ ************ *******
CONSENSUS   TCATAGAAATGTCGAAAGAGCGTGGTTTTGATACAAATCGGCTCATTTACGTTCAGCTGCAATAA  516
```

difference is less than 5% (Table 3.3B). It is strongly indicated that the *vlpA* DNA sequence is highly conserved within *Vibrio spp.*

3.2.3.3 Nucleotide sequence and analysis of VCRs

Sequence analysis in **Section 3.2.3.1** revealed that the open reading frames (VCOs) of the cloned DNA fragments in plasmids pPM471, pPM5001, pPM5002, pPM5003, and pPM5004 are flanked by VCRs as shown in Figure 3.5. In order to determine the similarity of each copy of VCR, twenty-two VCR sequences from these cloned DNA fragments were analyzed by using the CLUSTAL W program, and the multiple alignment is shown in Figure 3.11A. The VCR sequences are present in the same orientation and share specific bases at the beginning and the terminus of VCR sequences (5′ and 3′ ends). These sequences represent the most conserved features of the 59-be and have been termed the core sites which can be found at both ends of integrated gene cassette sequences (Recchia and Hall, 1995). The core sites from all VCR sequences in this study consist of RYYYAAC at the 5′ end and GTTRRRY at the 3′end in 55% and 64% of cases respectively (R = purine and Y = pyrimidine) (Figure 3.11B). The ratio of the core site sequences is present in Table 3.4. In addition, the VCR core-sites also show similarity to other 59-bes as shown in Figure 3.11C.

3.2.4 PCR amplification using the VCR and *vlpA* - specific primers.

To test for the association of *vlpA* with VCRs, PCR was performed in plasmid clones (pPM471, pPM5001, pPM5002), and both serogroups of *V. cholerae* O1, *V. cholerae* O139, *Vibrio spp.*, and *Aeromonas spp.* The VCR-specific primers (#922 and #923) which read outwards of both conserved ends of VCRs were used, with the *vlpA*-

Figure 3.11 VCR nucleotide sequence alignment.

(A) CLUSTAL W multiple sequence alignment of VCR nucleotide sequences in cloned DNA fragments from different strains of *V. cholerae*. Bases identical among all sequenced VCR are indicated by an asterisk. (VCR1-8, VCR8-16, VCR17-18, VCR19-20, and VCR 21-22 are the VCR sequences from pPM471, pPM5001, pPM5003, pPM5004, and pPM5002 respectively). (R = purine, Y = pyrimidine).

(B) The seven-base core sites at the 5′ and 3′ ends of VCR which are the most conserved features of VCR sequences, converted from Figure 3.11 (A).

(C) Comparison of the 7-bp core site (bold type) at both ends of 59-be (from pDGO100, R6-5(Tn*21*), pSa, and 59-be consensus) with the VCR consensus sequence.

Note : References of 59-be sequences above are :

pDGO100	accession number X04555	Cameron *et al.* (1986)
R6-5(Tn*21*)	accession number X12870	Sunstorm *et al.* (1988)
pSa	accession number M11444	Tait *et al.* (1985)
59-be consensus		Hall *et al.* (1991)

Figure 3.11 (A)

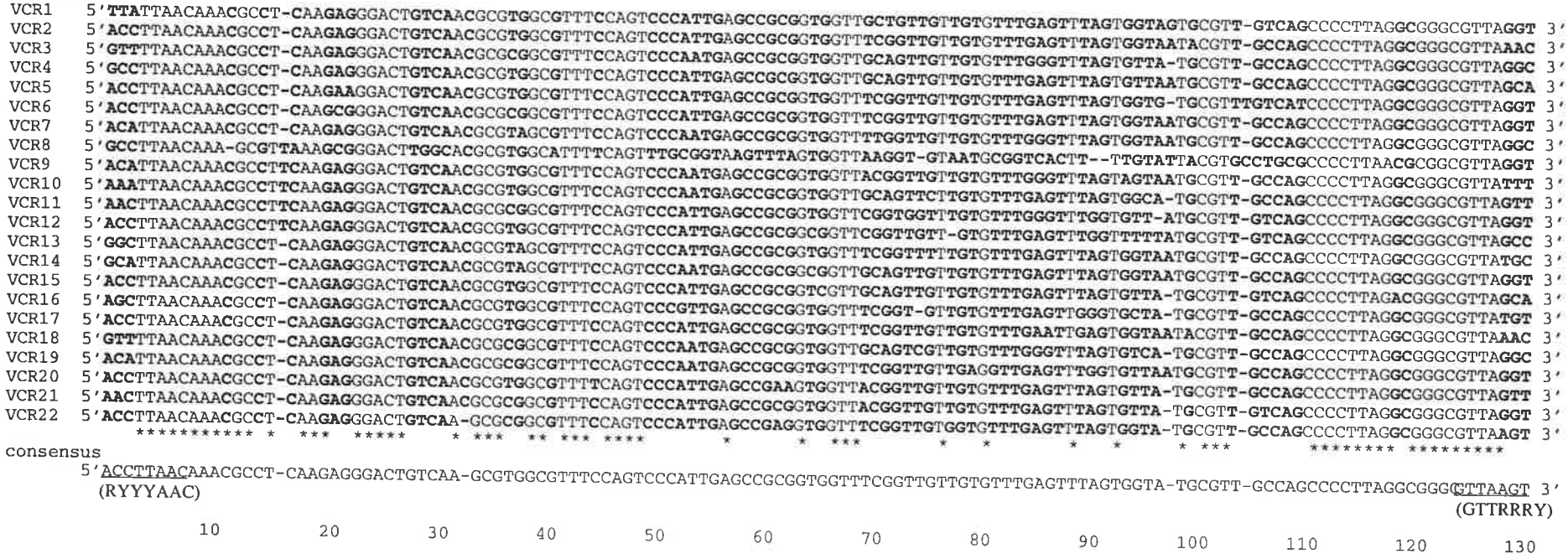

VCR1 5'TTATTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTGCTGTTGTTGTGTTTGAGTTTAGTGGTAGTGCGTT-GTCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR2 5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTCGGTTGTTGTGTTTGAGTTTAGTGGTAATACGTT-GCCAGCCCCTTAGGCGGGCGTTAAAC 3'
VCR3 5'GTTTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTGCAGTTGTTGTGTTTGGGTTTAGTGTTA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGC 3'
VCR4 5'GCCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTGCAGTTGTTGTGTTTGAGTTTAGTGTTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGCA 3'
VCR5 5'ACCTTAACAAACGCCT-CAAGAAGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTCGGTTGTTGTGTTTGAGTTTAGTGGTG-TGCGTTTGTCATCCCCTTAGGCGGGCGTTAGGT 3'
VCR6 5'ACCTTAACAAACGCCT-CAAGCGGGACTGTCAACGCGCGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTTGGTTGTTGTGTTTGAGTTTAGTGGTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR7 5'ACATTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTAGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTTTGGTTGTTGTGTTTGAGTTTAGTGGTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR8 5'GCCTTAACAAA-GCGTTAAAGCGGGACTTGGCACGCGTGGCATTTTCAGTTTGCGGTAAGTTTAGTGGTTAAGGT-GTAATGCGGTCACTT--TTGTATTACGTGCCTGCGCCCCTTAACGCGGCGTTAGGT 3'
VCR9 5'ACATTAACAAACGCCTTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTACGGTTGTTGTGTTTGGGTTTAGTAGTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTATTT 3'
VCR10 5'AAATTAACAAACGCCTTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTGCAGTTCTTGTGTTTGAGTTTAGTGGCA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGTT 3'
VCR11 5'AACTTAACAAACGCCTTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTCGGTGGTTTGTGTTTGGGTTTGGTGTT-ATGCGTT-GTCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR12 5'ACCTTAACAAACGCCTTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGCGGTTCGGTTGTT-GTGTTTGAGTTTGGTTTTTATGCGTT-GTCAGCCCCTTAGGCGGGCGTTAGCC 3'
VCR13 5'GGCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTAGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTCGGTTTTTGTGTTTGAGTTTAGTGGTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTATGC 3'
VCR14 5'GCATTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTAGCGTTTCCAGTCCCAATGAGCCGCGGCGGTTGCAGTTGTTGTGTTTGAGTTTAGTGGTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR15 5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTCGTTGCAGTTGTTGTGTTTGAGTTTAGTGTTA-TGCGTT-GTCAGCCCCTTAGACGGGCGTTAGCA 3'
VCR16 5'AGCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCGTTGAGCCGCGGTGGTTTCGGT-GTTGTGTTTGAGTTGGGTGCTA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTATGT 3'
VCR17 5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTCGGTTGTTGTGTTTGAATTGAGTGGTAATACGTT-GCCAGCCCCTTAGGCGGGCGTTAAAC 3'
VCR18 5'GTTTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTGCAGTCGTTGTGTTTGGGTTTAGTGTCA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGC 3'
VCR19 5'ACATTAACAAACGCCT-CAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGTGGTTTCGGTTGTTGAGGTTGAGTTTGGTGTTAATGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR20 5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGTGGCGTTTTCAGTCCCATTGAGCCGAAGTGGTTACGGTTGTTGTGTTTGAGTTTAGTGTTA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAGTT 3'
VCR21 5'AACTTAACAAACGCCT-CAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTACGGTTGTTGTGTTTGAGTTTAGTGTTA-TGCGTT-GTCAGCCCCTTAGGCGGGCGTTAGGT 3'
VCR22 5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAA-GCGCGGCGTTTCCAGTCCCATTGAGCCGAGGTGGTTTCGGTTGTGGTGTTTGAGTTTAGTGGTA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAAGT 3'
consensus
5'ACCTTAACAAACGCCT-CAAGAGGGACTGTCAA-GCGTGGCGTTTCCAGTCCCATTGAGCCGCGGTGGTTTCGGTTGTTGTGTTTGAGTTTAGTGGTA-TGCGTT-GCCAGCCCCTTAGGCGGGCGTTAAGT 3'
(RYYYAAC)
(GTTRRRY)
10 20 30 40 50 60 70 80 90 100 110 120 130

Figure 3.11 (B)

```
VCR1    5'YYRYAAC......124-bp..............GTTRRRY 3'
VCR2    5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR3    5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR4    5'RYYYAAC......124-bp..............GTTRRYR 3'
VCR5    5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR6    5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR7    5'RYRYAAC......124-bp.............. GTTRRRY 3'
VCR8    5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR9    5'RYRYAAC......124-bp..............GTTRYYY 3'
VCR10   5'RRRYAAC......124-bp..............GTTRRYY 3'
VCR11   5'RRYYAAC......124-bp..............GTTRRRY 3'
VCR12   5'RYYYAAC......124-bp..............GTTRRYY 3'
VCR13   5'RRYYAAC......124-bp..............GTTRYRY 3'
VCR14   5'RYRYAAC......124-bp..............GTTRRRY 3'
VCR15   5'RYYYAAC......124-bp..............GTTRRYR 3'
VCR16   5'RRYYAAC......124-bp..............GTTRYRY 3'
VCR17   5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR18   5'RYYYAAC......124-bp..............GTTRRRY 3'
VCR19   5'RYRYAAC......124-bp..............GTTRRRY 3'
VCR20   5'RYYYAAC......124-bp..............GTTRRYY 3'
VCR21   5'RRYYAAC......124-bp..............GTTRRRY 3'
VCR22   5'RYYYAAC......124-bp..............GTTRRRY 3'
```

Figure 3.11 (C)

pDGO100	*aadB*	5′	**GCCTAAC**AA-TTCGTCCAAGCCGACGCCGC-T-TC-GCGGC-GCGGCTTAACTC-AGGT**GTTAGAT**	3′
R6-5(Tn*21*)	*aadA1*	5′	**GTCTAAC**AA-TTCGTTCAAGCCGACGCCGC-T-TC-GCGGC-GCGGCTTAACTC-AAGC**GTTAGAT**	3′
pSa	*aadA2*	5′	**GTCTAAC**AA-TTCGTTCAA-CCGACTCA-----TC-GCGCCGGCGGCTTAACTC-CAGC**GTTAGAT**	3′
59-be consensus		5′	**GTCTAAC**AA-TTCGTTCAAGCCGACGCCGC-T-TC-GCGGC-GCGGCTTAACTC-AAGC**GTTAGAT**	3′
VCR consensus		5′	**ACCTAAC**AAA.....113nt.....**GTTAGCT**	3′

Table 3.4 The ratio of the 7-bp conserved nucleotide bases at the VCR core site sequences (R = purine, Y = pyrimidine).

bases at the 5′ end of VCR		bases at the 3′ end of VCR	
bases	%	bases	%
RYYYAAC	55	GTTRRRY	64
RYRYAAC	18	GTTRRRY	14
RRYYAAC	18	GTTRRYR	9
RRRYAAC	4.5	GTTRYRY	9
YYRYAAC	4.5	GTTRYYY	4

specific primers (#1071 and #1072) in combination, i.e. #1071 with #923, and #1072 with #922 (see diagram from Figure 3.12A). Different sizes of PCR products were detected in the plasmid clones as shown in Figure 3.12B, as these plasmids contain many copies of VCR. Some strains of *V. cholerae* O1, O139, non-O1 and *V. anguillarum* tested, amplified a specific product of 520-bp while others amplified two specific products of 520-bp and 1.1-kb (Figures 3.12C, D). PCR amplifications in *V. anguillarum* in lane 2 (Figures 3.12C, D) show different size products from other strains. It is possible that the sequence between VCR and *vlpA* in this strain may be also different. It is indicated that there are many copies of VCR which can give many PCR products in different sizes as shown in diagram in Figure 3.12A. However, no product was detected for *V. mimicus*, *V. fluvialis*, *V. parahaemolyticus* and *Aeromonas spp.* implying that the *vlpA* gene is not present in these strains, or alternatively that sufficient divergence has occurred in *vlpA* such that the primers are unable to bind, or that the VCR is not associated with *vlpA* in these strains. Analysis by PCR-amplification using the *vlpA* primers (#1071 and #1072) and Southern hybridization in **Section 3.2.1** confirms that the *vlpA* was not detected in these *Vibrio spp.* further supports this.

3.2.5 Long range PCR to map the distance between *vlpA* copies in *V. cholerae* 569B, O17 and O139.

Long range PCR which can amplify target fragments of up to 35-kb (Barnes, 1994) was performed in order to determine the distance between each copy of the *vlpA* genes in *V. cholerae* O1 569B, El Tor O17 and O139 chromosomes. The specific oligonucleotide primers were designed with a minimum length of 33 bases from both ends of pPM471 (#2480, #2481), pPM5001 (#2476, #2477) and pPM5002 (#2478, #2479). The primers

Figure 3.12 PCR amplification using VCR and *vlpA*- specific primers.

(A) Schematic representation of PCR-amplification across VCR and *vlpA* using primers (#1071 with #923) and (#1072 with #922). The boxes represent the *vlpA* and VCR sequences. The primers are represented by arrows and number (not to scale).

(B) PCR products of plasmid clones using primers #923 and #1071. Lanes contain PCR products from : **1,2**, *V. cholerae* 569B and O139 as positive controls; **3**, pPM471; **4**, pPM5001; **5**, pPM5002; **6**, pPM5003; **7**, *E. coli* DH5α as a negative control.

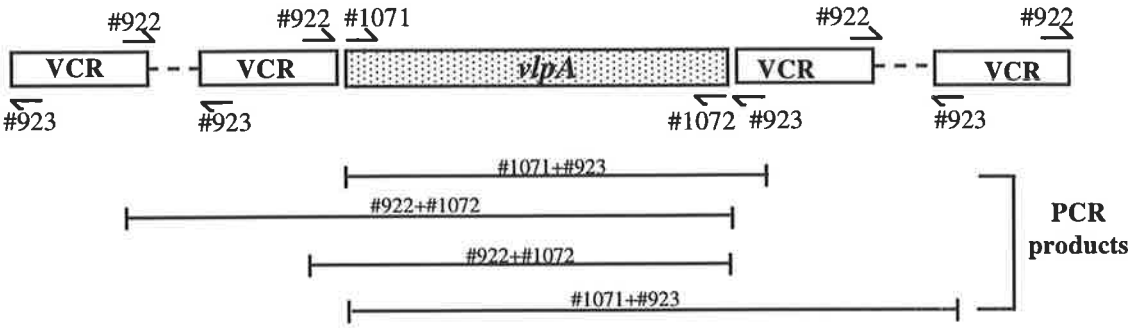
(C) PCR products from genomic DNA of *Vibrio spp.* using primers #922 and #1072. Lanes contain PCR products from strains: **1**, *V. cholerae* O139 (V911); **2-3**, *V. anguillarum* (**2**, ATCC43307; **3**, O1-85-3954-1); **4-6**, *V. cholerae* non-O1 (**4**, O55; **5**, O38; **6**, O14); **7**, *V. cholerae* O139 (AI-1837); **8-15**, *V. cholerae* O1 (**8**, AA14041; **9**, CA411; **10**, CA401; **11**, Z17561; **12**, 569B; **13**, O17; **14**, AA14073; **15**, AA13993).

(D) PCR products of *Vibrio spp.* using primers #923 and #1071. Lanes contain PCR products from strains: **1**, V911; **2-3**, *V. anguillarum* (**2**, ATCC43307; **3**, O1-85-3954-1); **4-6**, *V. cholerae* non-O1 (**4**, O55; **5**, O38; **6**, O14); **7**, *V. cholerae* O139 (AI-1837); **8-15**, *V. cholerae* O1 (**8**, AA14041; **9**, CA411; **10**, CA401; **11**, Z17561; **12**, 569B; **13**, O17; **14**, AA14073; **15**, AA13993).

The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.

Figure 3.12

(A)



(B)

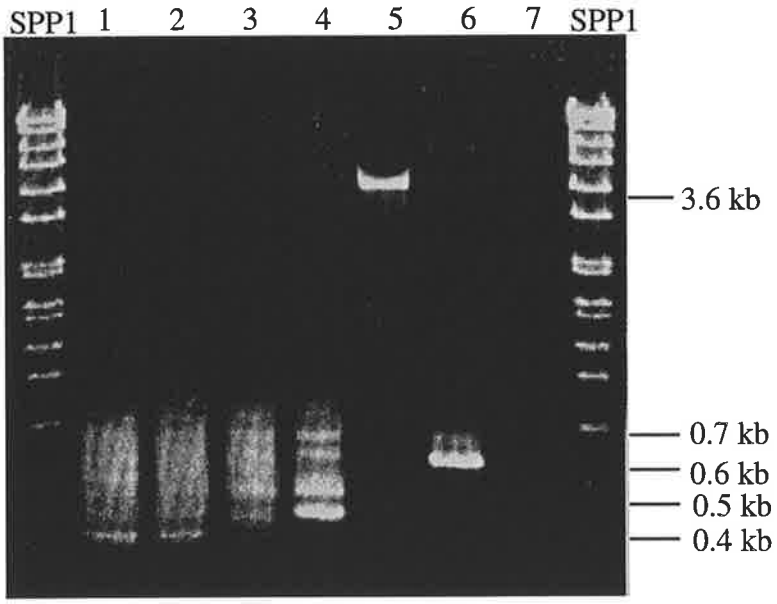
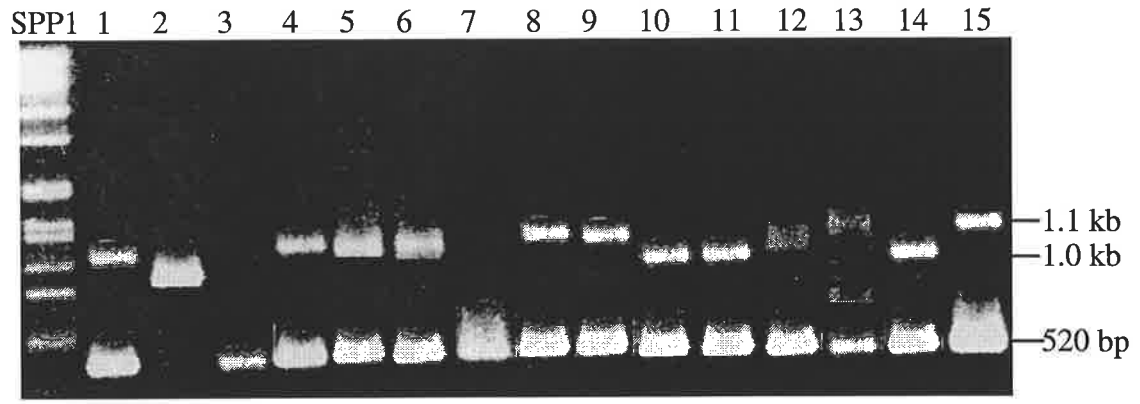
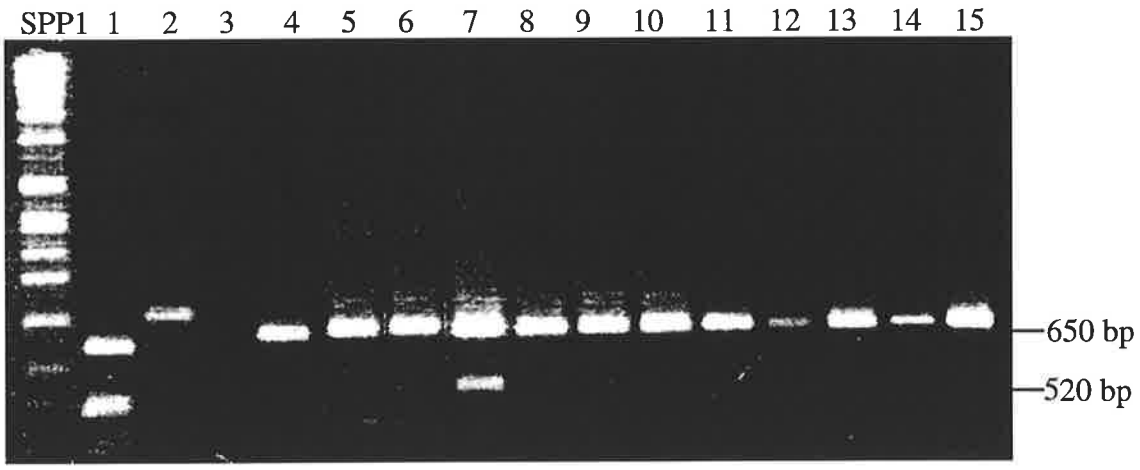


Figure 3.12

(C)



(D)



were not designed for pPM5003 and pPM5004, which are the plasmid clones from *V. cholerae* O139, as these plasmids have the same sequence to pPM471 and pPM5002, respectively. It was assumed that each copy of *vlpA* transcribes in the same direction. The outward reading primers were used to amplify across the end of two *Bam*HI fragments in *V. cholerae* classical 569B (see diagram from Figure 3.13A). Primer #2476 was used with #2481, and #2477 with #2480 to see if the fragment from pPM5001 links to that in pPM471. In addition, primer #2476 was used with #2479 and #2477 with #2478 to see if any sequence identity exists in *V. cholerae* El Tor O17 with *V. cholerae* classical 569B (Figure 3.13B). The primers #2479 was also used with #2480, and #2481 with #2478 to see if any sequence identify exists in *V. cholerae* 569B with *V. cholerae* El Tor O17 (Figure 3.13C). No PCR product was detected in any PCR amplification. It is possible that the distance between each copies of *vlpA* is too far for efficient amplification, even with the use of the long range PCR. In addition, it is possible that the *vlpA* flanking genes in different strains are not identical, and in this case the primers may not be useful in amplifying the fragment.

3.2.6 Analysis of the distance between each copy of *vlpA* using the Institute for Genomic Research (TIGR) Database.

As the long range PCR strategy was unsuccessful in amplifying between each copy of *vlpA*, a physical map was constructed from the *V. cholerae* genome sequence database (http://www.tigr.org). At this stage, the fully completed genome sequences had not been released or published (Heidelberg *et al.*, 2000). Sequence analysis of *V. cholerae* O1 El Tor strain N16961 was conducted by overlapping of the contiguous regions (contigs) surrounding the *vlpA* genes and VCRs. Analysis of genome sequence from this strain

Figure 3.13 Schematic representation of primers used in long range PCR to amplify the fragment between each copy of *vlpA* in *V. cholerae*.

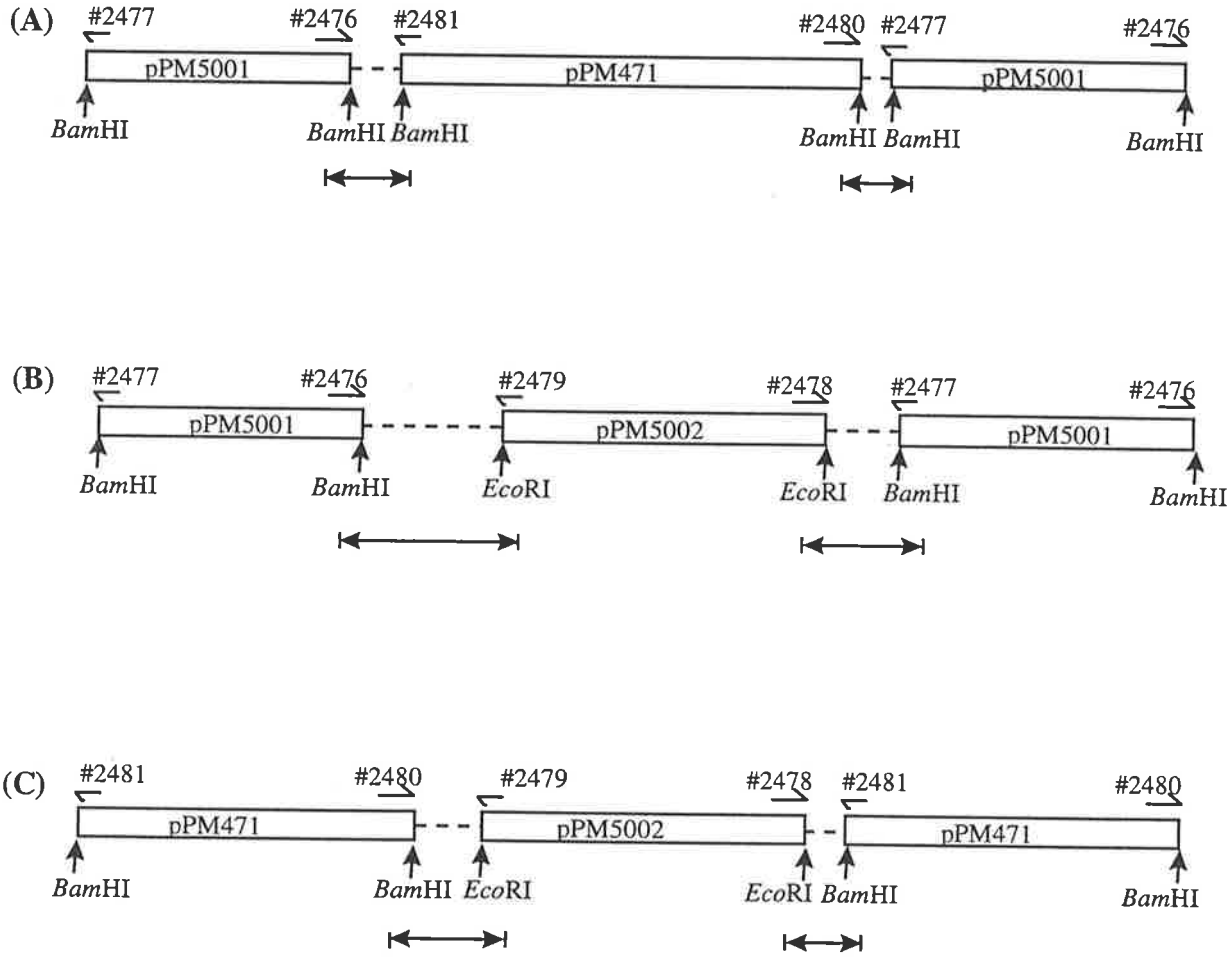
The boxes represent the cloned fragments and the primers are represented by arrows and numbers. Bipolar arrows represent the PCR products to be amplified (not to scale) .

(A) Primers used to amplify between two *Bam*HI fragment (pPM5001 and pPM471) in *V. cholerae* 569B. Primers used are (#2476 / #2481), and (#2480 / #2477), respectively.

(B) Primers used to amplify between the *Bam*HI fragment (pPM5001) and *Eco*RI fragment (pPM5002) in *V. cholerae* O17. Primers used are (#2476 / #2479), and (# 2478 / # 2477), respectively.

(C) Primers used to amplify between the *Bam*HI fragment (pPM471) and *Eco*RI fragment (pPM5002) in *V. cholerae* O17. Primers used are (#2480 / #2479), and (#2478 / #2481), respectively.

Figure 3.13


(A)
#2477
#2476
#2481
#2480
#2477
#2476
pPM5001
pPM471
pPM5001
BamHI
BamHI
BamHI
BamHI
BamHI
BamHI
(B)
#2477
#2476
#2479
#2478
#2477
#2476
pPM5001
pPM5002
pPM5001
BamHI
BamHI
EcoRI
EcoRI
BamHI
BamHI
(C)
#2481
#2480
#2479
#2478
#2481
#2480
pPM471
pPM5002
pPM471
BamHI
BamHI
EcoRI
EcoRI
BamHI
BamHI

showed the first VCR to be present on the smaller of the two *V. cholerae* genomes (i.e. chromosome 2). This VCR starts at nt 311,316 which is immediately upstream of the *intI*4 gene (GenBank accession number AF055586; Mazel *et al.*, 1998), and the last VCR ending at nt 435,032. This region contains at least 155 copies of VCR in arrays with the distance approximately 124-kb (from nt 311,316 to nt 435,032), which corresponds to 12% of the 1.0-Mb smaller *V. cholerae* chromosome (Trucksis *et al.*, 1998; Heidelberg *et al.*, 2000) (Figure 3.14).

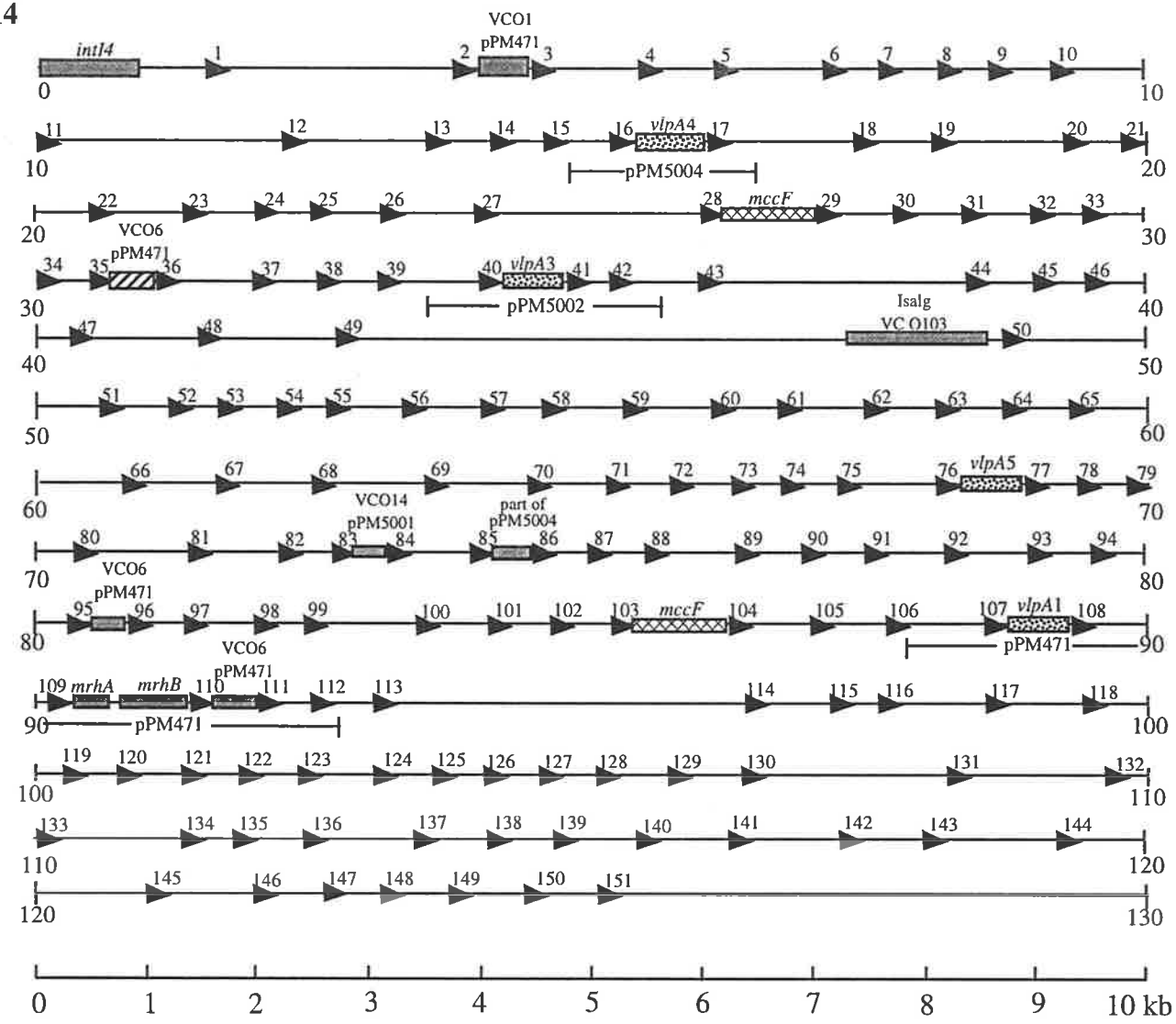
The *V. cholerae* O1 El Tor strain N16961 genome contains four copies of *vlpA* gene (Figure 3.14) This result correlates well with the data in **Section 3.2.1**, which showed by Southern hybridization that *V. cholerae* El Tor strain O17 also contained four copies of *vlpA*. The first, second, third, and fourth *vlpA* genes were found beginning at nt 325,224, 343,996, 378,081 and 398,664, respectively. Therefore, the distance between each copy of *vlpA* varies by approximately 19, 21 and 35-kb, respectively. The first *vlpA* (*vlpA*4) and flanking genes are similar to the DNA cloned fragment in pPM5004. This is in agreement with **Section 3.2.3.1** where sequence data from the PCR product of *V. cholerae* O1 El Tor strain O17 (using the primers #3141 and #3142) also showed identity with pPM5004. The second *vlpA* (*vlpA*3) and flanking genes are similar to the DNA cloned fragment in pPM5002. The flanking genes of the third *vlpA* (*vlpA*5) does not show similarity to any DNA cloned sequence presented in this thesis. The last *vlpA* (*vlpA*1) and flanking genes were similar to the DNA cloned fragment in pPM471. Interestingly, the sequence of the *vlpA*2 flanking genes from DNA cloned in pPM5001 was not homologous to any fragment from the TIGR *V. cholerae* genomic sequence. However, some open reading frames from pPM5001 in addition to *vlpA*, for example, the *mccF* gene and VCO14 (VCA0415 and VCA0423 from GenBank accession no AE004376) were found. In the case of *mccF*, this

Figure 3.14 Physical map of the *vlpA* / VCR containing region of the *V. cholerae* El Tor strain N16961.

The comparison with the data generate here, the TIGR *V. cholerae* genomic database (http://www.tigr.org) was used to investigate the copy number of VCRs and the distance between each copy of *vlpA* in the El Tor strain N16961. The *vlpA* and flanking genes in *V. cholerae* El Tor are similar to pPM471 from *V. cholerae* 569B , and pPM5004 from *V. cholerae* O139. ORFs are represented as boxes, termed VCOs.

$\blacktriangleright^{n}$ repesents VCR

Figure 3.14


VCO1
pPM471
intI4
vlpA4
pPM5004
mccF
VCO6
pPM471
vlpA3
pPM5002
IsaIg
VC 0103
vlpA5
VCO14
pPM5001
part of
pPM5004
VCO6
pPM471
mccF
vlpA1
pPM471
VCO6
pPM471
mrhA
mrhB
pPM471
0 1 2 3 4 5 6 7 8 9 10 kb

gene was found present as two copies in the genome. Although VCO14 is next to the *vlpA* gene in *V. cholerae* O1 strain 569B, this VCO is not associated with the *vlpA* gene in the *V. cholerae* O1 El Tor strain N16961.

Many VCOs from *V. cholerae* O1 strain 569B and O139 were also present in the El Tor strain N16961 genome, although not in the same relative positions. These VCOs are always adjacent to a VCR and present as more than one copy in the integron, for example, VCO6 from pPM471 (VCA0410, VCA0413, VCA0428 from GenBank accession no AE004372, AE004376, AE004378 respectively) present as three copies. Some part of pPM5004 (from *V. cholerae* O139) present twice in the integron. In addition, VCO from other *Vibrio spp.* is also present in *V. cholerae* El Tor N16961, for example an inserted sequence in nqr6 of Nap-I mutant (Isalg) which functions as the sodium pump-defective from the marine *V. alginolyticus* (GenBank accession number GI1565218; Hayashi *et al.*, 1997), is found located at the nt 357173 to nt 358432. Isalg is also found in *V. cholerae* serogroup O103 (GenBank accession number GI6014464). Interestingly, no VCRs are present in approximately 4.3-kb (from nt 352,815 to 357,173) in the region which contains an insertion element Isalg, although most of VCOs are found adjacent to VCRs.

3.3 Discussion

The distribution of the *vlpA* genes in *Vibrio* and related genera was analyzed by colony hybridization, Southern hybridization, PCR amplification and DNA sequencing. Colony hybridization was the preliminary detection method used and the data were confirmed by Southern hybridization using a PCR DIG-labeled *vlpA* specific probe (derived from the 5′ and 3′ ends internal to the open reading frame). Individual PCR amplification was performed for each strain using the same specific internal *vlpA* primers.

Analysis of Southern hybridization data has revealed that the numbers of copies of the *vlpA* genes varies from one to four. While the *vlpA* genes are present in all strains of *V. cholerae* O1 and O139 tested in this study, this is not the case for all strains of *V. cholerae* non-O1 serogroups. *vlpA* was not detected in *V. mimicus*, *V. parahaemolyticus*, and *V. fluvialis*, whereas it was present (up to four copies) in the genome of different *V. anguillarum* serotypes. In addition, *vlpA* was not detected in *Aeromonas spp*. The presence of *vlpA* as multiple copies in *V. cholerae* O1, O139, some strains of *V. anguillarum* and *V. cholerae* non-O1 poses the question of how the gene copies were generated.

It is not surprising that the *vlpA* gene was not detected in *Aeromonas spp*. since there is increasing evidence to separate them from *Vibrio spp*. Their growth characteristics are different: Aeromonads are usually inhibited on thiosulfate-citrate-bile salt-sucrose (TCBS) agar, a selective medium used for the isolation of *V. cholerae* (Baumann and Schubert, 1984). Although members of the genus *Aeromonas* are grouped in the family *Vibrionaceae*, their phylogeny now appears more closely related to the *Enterobacteriaceae*. The Aeromonads may therefore have to be removed from the family *Vibrionaceae* and given a new family designation (Baumann and Schubert, 1984). Indeed, Colwell *et al.* (1986) suggested that the Aeromonads should be classified within their own family, the *Aeromonadaceae*, because they are closely related to both the *Enterobacteriaceae* and the *Vibrionaceae* (Pemberton *et al.*, 1997).

One particularly interesting observation to emerge from this study is that the *vlpA* genes are spread among *V. cholerae* O1, O139, some of the non-O1 serogroups, and *V. anguillarum* but not in *V. mimicus*, *V. parahaemolyticus*, and *V. fluvialis*. The fact that the *vlpA* gene has not been found in *V. parahaemolyticus* and *V. fluvialis* is not altogether surprising since they differ from *V. cholerae* both by rRNA sequencing and DNA

hybridization (Kita *et al.*, 1993). In addition, both *V. parahaemolyticus* and *V. fluvialis* are more distantly related species to *V. cholerae* than *V. mimicus*, and belong to other lines of descent in the *Vibrio* genus (Kita *et al.*, 1993).

Surprisingly, *vlpA* has not been detected in the *V. mimicus* strains tested in this study, although *V. cholerae* and *V. mimicus* are very closely related, and at least three VCR cassettes were reported in *V. mimicus* (Mazel *et al.*, 1998). Strains of *V. mimicus* were formerly calssified as in the *V. cholerae* non-O1 serogroup (Sanyal, 1992). Biochemical and physiological tests (Molitoris *et al.*, 1989) have shown *V. mimicus* and *V. cholerae* to be very similar. The susceptibility pattern to antibiotics is also very similar, except for colistin, to which most *V. cholerae* strains show resistance (Davis *et al.*, 1981). Some *V. mimicus* isolates also have a high percentage DNA relatedness (80% DNA-hybridization) to *V. cholerae* (Desmarchelier and Reichelt, 1981). *V. mimicus* is indistinguishable from *V. cholerae* by serology (Sakazaki, 1992), and also by either rRNA or DNA hybridization. The toxin produced by *V. mimicus* appears to be identical to cholera toxin including its receptor site is GM1 ganglioside of gastro-intestinal epithelial cells (Sanyal, 1992). However, biotypically *V. mimicus* is distinguishable from *V. cholerae* in many characteristics (Davis *et al.*, 1981; Baumann and Schubert, 1984). Strains of *V. mimicus* are sucrose negative, whereas all *V. cholerae* strains are sucrose positive and this is the most useful diagnostic test in separating these two species. However, *V. mimicus* is also differentiated from *V. cholerae* by its negative Voges-Proskauer and Jordan tartrate reactions, and by its sensitivity to polymyxin (Davis *et al.*, 1981). Most strains of *V. mimicus* are typable with antisera produced against *V. cholerae* (Davis *et al.*, 1981). It is possible that the *vlpA* gene in association with VCR may be disseminated among *Vibrio*

spp. via horizontal gene transfer and may be acquired or lost by some strains during evolution.

In an attempt to determine the relationship between *V. cholerae* and *V. anguillarum*, Stroeher *et al.* (1998) analyzed several genes involved in LPS biosynthesis and the presence of a novel genetic element. Based on this analyses, it appears that *V. cholerae* and *V. anguillarum* are very closely related. The study of an association of IS*1358* with *rfb* and capsule loci in these two species revealed that IS*1358* is associated with the *rfb* region in *V. cholerae* O1 and O139 and in *V. anguillarum* O1 and O2 (Stroeher *et al.*, 1998).

V. anguillarum is a pathogen of many fish species and causes a bacteraemia in salmonid fish resulting in internal haemorrhage and often death. *V. anguillarum* is very similar to *V. cholerae* with O-antigen biosynthesis (Norqvist and Wolf-Watz, 1993). Interestingly, like *V. cholerae* only a few of the ten to eleven recognized *V. anguillarum* serogroups cause the majority of the disease. *V. anguillarum* is genetically closely related to *V. cholerae*. Many of the novel genetic elements such as VCR (Barker *et al.*, 1994), and IS*1358* (Stroeher *et al.*, 1998) are found in both species, but do not appear to be present in other *Vibrio* species. Furthermore, the RecA proteins are very closely related (Stroeher *et al.*, 1994). Stroeher *et al.* (1998) suggested that it is possible to envisage that there is or has been an exchange of genetic material between these species. This genetic exchange may well also take place between many of the serogroups of *V. cholerae* and *V. anguillarum.* The presence of IS*1358* in seven *V. anguillarum* serogroups and approximately seventy *V. cholerae* serogroups (Stroeher *et al.*, 1998) as well as the presence of the VCR element (C.A. Clark, personal communication) is suggestive of such exchanges.

When *vlpA* is analyzed, the similarity of the sequences in *V. cholerae* and *V. anguillarum* is almost identical (more than 94 %, see **Section 3.2.3.2**) and it seems feasible to suggest that this gene represents genetic material that has been exchanged between these two species. In addition, *vlpA* in both species is also flanked by a VCR element. The VCR element has features common to gene capture systems termed integrons (Hall, 1995). VCRs possess a 7-bp core site at both ends, which in the case of 59-bes function as recombination sites (Recchia and Hall, 1995; Mazel *et al.*, 1998). The presence of these VCR elements is suggestive of such an exchange. An increasing number of genes in prokaryotes have now been reported to be acquired by horizontal transfer (Hall, 1998). Thus, it is possible that the *vlpA* gene in association with a VCR could also be exchanged between *Vibrio* species via horizontal gene transfer using such a gene capture system. Therefore, while it appears that horizontal transfer may have occurred between these species, or a high level of recombination occurs between these species, it also indicates that *V. cholerae* and *V. anguillarum* are very closely related. It is feasible to suggest that *V. cholerae* and *V. anguillarum* may share the same pool of cassettes and that VCR cassettes are spread amongst these two species.

Screening of 155 serogroups of *V. cholerae* non-O1, O1 - O155, has revealed that approximately 26 % of all strains contain the *vlpA* gene. The result from **Section 3.2.1** (Table 3.1) indicates that *vlpA* was not detected in some *V. cholerae* non-O1 serogroups (approximately 5%) by PCR amplification but could be detected by Southern hybridization when probed with a PCR DIG-labeled *vlpA* specific probe at low stringency. This may due to mutations in the primer target region and a subsequent loss of PCR product, although these strains may still have the *vlpA* genes. The absence of a PCR product and detection at

low stringency suggests that there may be some nucleotide sequence variation in the *vlpA* gene especially at the ends of the gene.

Although a virulence-associated distribution in *vlpA* has not been found, the feature of the cassette gene indicates that *vlpA* may be a mobile gene and can move within both *V. cholerae* and other closely related *Vibrio spp.* **Section 3.2.6** demonstrates the identity of chromosomal DNA fragments between *V. cholerae* O1 classical, El Tor, and O139. A number of loci on the *V. cholerae* O1 El Tor strain N16961 genome contain the *vlpA* and flanking genes from *V. cholerae* O1 strain 569B (pPM471 and pPM5001) and O139 (pPM5003 and pPM5004). In addition, the insertion element Isalg from *V. alginolyticus* (Hayashi *et al.*, 1997) was also found in this area of the genome. Furthermore, the *vlpA* nucleotide sequence identity in *V. cholerae* O1, O139 and non-O1 is more than 90% (**Section 3.2.3.2**). These results strongly indicate that the *vlpA* gene cassette and other VCOs in associated with the VCR element can move within *Vibrio spp.* A further discussion on this matter will be presented in **Chapter7**.

Cloning and sequencing the DNA fragments containing *vlpA* and flanking regions (**Section 3.2.2**) has indicated that the nucleotide sequence in some loci from the different strains of *V. cholerae* genomes are similar. Homologous sequences are found in *V. cholerae* O1 strain 569B (pPM471) and *V. cholerae* O139 (pPM5003). In addition, PCR amplification using the primers generated from pPM5004, which contain DNA cloned from the *V. cholerae* O139 chromosome, amplified a PCR product from *V. cholerae* El Tor O17. DNA sequencing of this PCR product showed identity between the *V. cholerae* O1 El Tor strain O17 and O139 (pPM5004). Sequence analysis of the TIGR database of *V. cholerae* El Tor N16961 in **Section 3.2.6** also supports this result.

All of the VCOs in pPM471, pPM5001, pPM5002, pPM5003 and pPM5004 are associated with VCRs, and encode mostly unknown functions. No well-defined antibiotic resistance genes appear within these clusters, although until now most of the gene cassettes reported in other integrons have commonly been antibiotic resistance genes. Some VCOs are in the opposite transcriptional orientation compared to the majority of the VCOs, for example, VCO 3.1 and VCO 3.2. Interestingly, some VCOs appear to be virulence related, such as the *mrhA* gene in pPM471 and pPM5003. This gene encodes MFRHA (mannose-fucose-resistant hemagglutination), and the MFRHA - negative mutant has been assessed for virulence in the infant mouse cholera model. This mutant showed a marked defect in its ability to persist in the infant mouse gut and is incapable of competing with the wild-type organism (Franzon *et al.*, 1993).

The amino acid sequence of VCO1 in pPM5001 displays 56% identity with *EcmccF* gene (*EcmccF* refers to the *E. coli mccF* gene in this study). This gene is located at a 6.2-kb region of the *E. coli* plasmid pMccC7. This region contains six ORFs, *mcc ABCDE*, which are transcribed in the same direction and *mccF*, adjacent to *mccE*, which is transcribed in the opposite direction. Five genes, *mcc ABCDE* are required to produce mature extracellular microcin, and *mccF* gene encodes specific self-immunity (Gonzalez-Pastor *et al.*, 1995). Like the *vlpA* gene, the occurrence of *VcmccF* gene in *V. cholerae* indicates that *VcmccF* is also flanked by a VCR. Purins (1997) reported that Southern hybridization of the *V. cholerae* O1 569B chromosome with a *VcMccF* probe showed two hybridization bands which correlates well with the sequence of *V. cholerae* El Tor N16961 which also has two copies (Section 3.2.6, Figure 3.14).

VlpA is homologous to the bacterial lipocalins, i.e. Blc from *E. coli* (SWISS PROT accession no. P39281) as well as Blc from *C. freundii* (SWISS PROT accession no.

C freundii

Q46306). These genes have now been designated *Ecblc*, and *Ctblc* respectively. A repeat region which is similar to a VCR element was reported in *E. coli* K-12 chromosomal region from 92.8 through 100.1 minutes (GenBank accession no. U14003). However, this repeat sequence is not found to be associated with the *Ecblc* gene (GenBank accession no. GI3132840), and the *Ctblc* gene (GenBank accession no. GI3056569). Bishop and Weiner (1998) have suggested that the bacterial lipocalins were acquired by horizontal transmission from eukaryotes. Flower (1995) reported that it might be argued that the existence of one prokaryotic lipocalin can be explained by the mechanism of horizontal transfer, but with several examples from different species this seems unlikely. If the prokaryotic genes have not arisen by horizontal transfer, then the gene capture system could be considered (Flower, 1995).

In conclusion, the spread of *vlpA* as multiple copies in the *V. cholerae* chromosome still needs to be explained. Two hypotheses can be put forward: (1) horizontal gene transfer and integration into in to the chromosome by a non-specific mechanism, or (2) a site specific gene capture system catalyzed by an integron integrase. It is noted that *vlpA* including all VCO's are associated with VCRs. If VCRs function as specific recombination sites, it is possible that the multiple copies of *vlpA* genes occur by gene capture via an integron system. Further analysis is presented in chapter 4.

Table 3.5 Plasmids used in this chapter

Plasmid	Characteristics	Source/Reference
pPM471	6.3-kb *Bam*HI fragment from *V. cholerae* 569B cloned into pBR322	Franzon and Manning (1986)
pPM5001	4.6-kb *Bam*HI fragment from *V. cholerae* 569B cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5002	1.0-kb *Eco*RI fragment from *V. cholerae* El Tor O17 cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5003	2.1-kb *Eco*RV fragment from *V. cholerae* O139 cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5004	1.2-kb *Eco*RV fragment from *V. cholerae* O139 cloned into pBC-SK	This study (**Section 3.2.2**)

Chapter 4

The *vlpA* gene is a part of a *Vibrio cholerae* chromosomally-encoded integron

4.1 Introduction

The relatedness of the VCR consensus sequence to the recombination sites found in gene cassettes was first recognized by Recchia and Hall (1997) who observed that the outer ends of VCRs contain sequences similar to the core and inverse core sites of a 59-base element (59-be). Mazel *et al.* (1998) recently showed that VCRs can function as 59-bes and are indeed recognized by the integrase IntI1. The integrase gene (*int*) located in the 5′ conserved segment of the integron structure, encodes the recombinase, which is a member of the tyrosine recombinase or λ integrase family (Stokes and Hall, 1989; Martinez *et al*, 1990; Abremski *et al.*, 1992; Collis and Hall, 1995; Ouellette and Roy, 1997). The chromosomal integron is located on chromosome 2, the smaller of the two *V. cholerae* chromosomes (Heidelberg *et al.*, 2000).

Another *V. cholerae* integrase has been reported by Kovach and colleagues (1996) but located on the larger chromosome. This integrase is present at the 5′ end of the ToxR-regulated TCP-ACF gene cluster on the *V. cholerae* pathogenicity island (VPI) (Kirby *et al.*, 1994; Kovach *et al.*, 1996). The integrase is located downstream of the *ssrA* gene, encoding the 10S small stable RNA. Numerous base pair substitutions, insertions and deletions have occurred in the DNA sequence between the *ssrA* and *int* genes in *V. cholerae* O1 El Tor and O139 (Kovach *et al.*, 1996).

Clark *et al.* (1997) together with recently published results of the genome sequence of the *V. cholerae* O1 El Tor strain N16961 (Heidelberg *et al.*, 2000) (see also Figure 3.15) show the presence of a chromosomal integron present as a 125-kb element containing the *intI*4 gene, four copies of *vlpA*, and at least 152 copies of VCRs. The *V. cholerae* N16961 integron has 216 ORFs and 95% of them have unknown functions. The *V. cholerae* Integrase IntI4 has 45 - 50% identity with the three known integrases, i.e. IntI1, IntI2 and IntI3 from class 1, 2 and 3 integrons, respectively (Mazel *et al.*, 1998). While recombinase activity has been demonstrated for these integrases, function for IntI4, as well as the more recently identified IntI5 from *V. mimicus* (Rowe-Magnus *et al*, 1999) which shows 94% amino acid identity with IntI4, has not yet been established.

Deletion, duplication, and rearrangement of cassette genes require the presence of an integrase, and this has been best studied using IntI1 isolated from *P. aeruginosa*, as a model system (Martinez and de la Cruz, 1988, 1990; Hall *et al.*, 1991; Collis and Hall, 1992a, 1992b; Collis *et al.* 1993; Stokes *et al.*, 1997). The recombination cross-over sites catalyzed by integrase IntI1 occur close to the 3′ end of 59-bes, within a conserved core site, defined as the consensus site GTTRRRY (Collis and Hall, 1992a; Stokes *et al.*, 1997). Interestingly, the outer ends of VCR sequences in *V. cholerae* are found to be significantly related to the consensus sequences of 59-bes (Recchia and Hall, 1997). In addition, the *vlpA* gene is flanked by VCRs and possesses all the features of a mobile gene cassette. However, it remains to be established experimentally whether *vlpA* exists as a gene cassette which can be excised from, or inserted into a new location in the *V. cholerae* chromosome and whether this can be catalyzed by an integron integrase, in particular IntI4, the integrase which defines this fourth class of integrons.

4.2 Results

4.2.1 Analysis of the *vlpA* cassette gene

Cloning and nucleotide sequence analysis of DNA fragments containing *vlpA* in Chapter 3 (**Section 3.2.2**) showed that all of the *vlpA* genes in plasmids pPM471, pPM5001, pPM5002, pPM5003 and pPM5004 (Table 4.5) contained a 516-bp coding region flanked by VCRs. The outer ends of the VCRs share common features with 59-be conserved core sites in both 5′ and 3′ ends (RYYYAAC and GTTRRRY) (see Chapter 3, **Section 3.2.3.3**, and Table 3.4). In addition, the *vlpA* stop codons are within the 5′ end of the downstream VCR, while the *vlpA* start codons are within 30 bp of the 3′ end of the upstream VCR. The conserved triplet GTT, which is part of the core site GTTRRRY found at the 3′ end of the 59-be, is also found at both ends of the *vlpA*-VCR region (Figure 4.1A), thereby showing the main features of integrated cassettes (Recchia and Hall, 1995; Stokes *et al.*, 1997) (Figure 4.1B).

All of the *vlpA* coding regions found in the plasmid clones (pPM471, pPM5001, pPM5002, pPM5003 and pPM5004) have the same orientation (Chapter 3, Figure 3.5) which correlates well to the direction of transcription found in the four copies of *vlpA* present in *V. cholerae* O1 El Tor strain N16961 (Heidelberg *et al.*, 2000). This feature is also found in other integrated cassettes that are always in the same orientation, with the 5′ end of the gene closest to the integron segment which contains the *intI* gene (Recchia and Hall, 1995). The presence of a VCR element, which functions as a recombination site, is an essential feature of the proposed *vlpA* gene cassette. Therefore, the first six bases of the proposed *vlpA* cassette would be derived from the 3′ end of a VCR and have the sequence TTRRRY, while the last base should be a conserved G residue (Figure 4.1B). The recombination crossover site has been localized to a unique position between the adjacent

Figure 4.1 Nucleotide sequence (A), and boundary (B) of the *vlpA* gene cassette.

(A) Nucleotide sequence of the *vlpA* cassette from pPM5004 contains the *vlpA* coding region (red) with VCRs (blue) flanking both ends. The seven-base core sites are shown in bold type (yellow highlight).

(B) Boundary of the *vlpA* gene cassette from (A). Vertical arrows indicate the position of the recombination cross-over. The core sites (GTTRRRY) found at each end of the *vlpA* cassette are shown. The inverse core site (RYYYAAC) at the 5′ end of the VCR is also shown. The boundary of a gene cassette in a linear array is shown to be formed into a circular gene cassette.

Figure 4.1

(A)

7-bp core site
AACTAACAAA CGCCTCAAGA GGACTGTCA ACGCGCGGCG TTTCCAGTCC CATTGAGCCG 1

CGGTGGTTAC GGTTGTTGTG TTTGAGTTTA GTGTTATGCG TTGTCAGCCC CTTAGGCGGG 61

7-bp core site
CGTTAGGTAT TTGGAGATAT start codon
ATCTTTTTGA TTCTTTGCTC TGTTTTATTA 121

AATGGCTGCT TGGGCATGCC CGAATCAGTA AAACCAGTGT CGGATTTTGA ACTGAACAAC 181

TATTTAGGTA AATGGTACGA AGTCGCTCGA CTCGATCACT CCTTTGAAAA AGGTTTAAGT 241

CAGGTTACAG CGGAATACCG TGTTCGAAAT GATGGTGGCA TTTTGGTTCT TAATCGTGGT 301

TATTCTGAAG AGAAAGGTGA GTGGAAGGAA GCCGAAGGCA AAGCTTACTT TGTGAATGGC 361

TCAACAGACG GCTATCTGAA GGTTTCATTT TTTGGTCCGT TTTATGGCTC CTACGTAGTG 421

TTTGAGTTAG ACCGTGAAAA CTACAGTTAT GCTTTTGTGT CAGGGCCGAA TACAGAATAT 481

CTGTGGTTAC TTTCAAGAAC GCCGACTGTA GAACGAGGCA TTCTGGACAA GTTCATAGAA 541

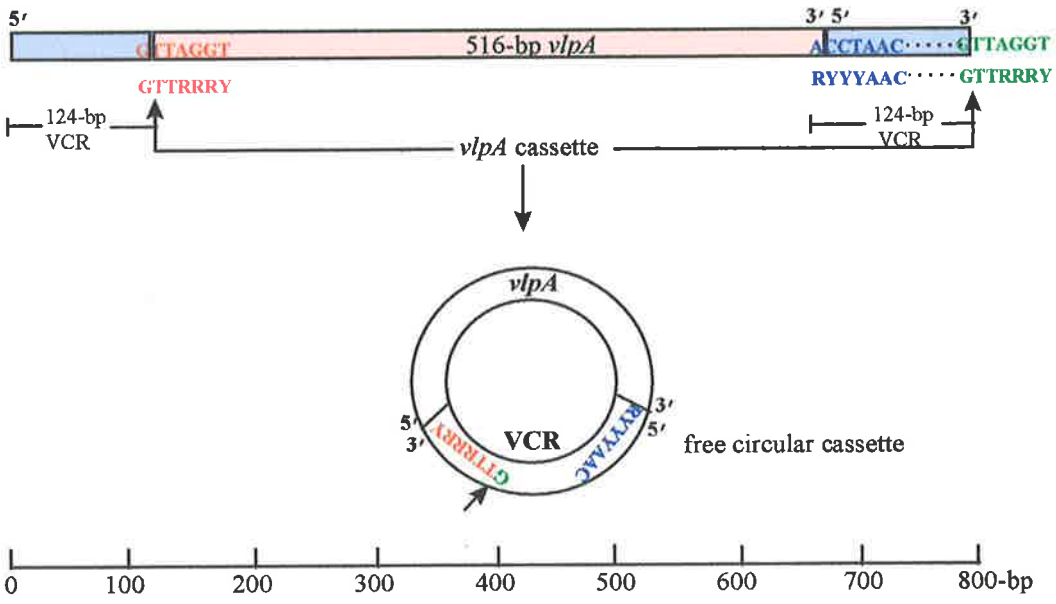
ATGTCGAAAG AGCGTGGTTT TGATACAAAT CGGCTCATTT ACGTTCAGCA GCAATAAATA 601

7-bp core site
CCTAACAAAC GCCTCAAGAG GGACTGTCAA GCGCGGCGTT TCCAGTCCCA TTGAGCCGAG 661
stop codon

GTGGTTTCGG TTGTGGTGTT TGAGTTTAGT GGTATGCGTT GCCAGCCCCT TAGGCGGGCG 721

7-bp core site
TTAAGT 781

(B)



G and the first residue of the proposed *vlpA* cassette as demonstrated for other gene cassettes by Recchia and Hall (1995).

4.2.2 Cassette insertion mediated by integrase.

To support the hypothesis that both the deletion and insertion of gene cassettes can occur in the presence of an integron-encoded DNA integrase, the ability of cassettes to insert into VCRs was initially attempted. Hall and colleagues (1991) reported that before the initial insertion of a particular gene can occur, further events are likely to be required. These are (i) the association of a 59-be with the 3′ end of an antibiotic resistance gene, and (ii) the circularization of this unit. To study the movement of a *vlpA* cassette, an insertionally inactivated copy of this gene with an antibiotic resistance gene was used. A *vlpA* circular cassette was constructed as described below to enable cassette insertion experiments to be undertaken.

4.2.2.1 Construction of a plasmid containing an insertionally inactivated *vlpA* gene with a kanamycin resistance cartridge (Km^R).

The 1.2-kb *Eco*RV DNA fragment in pPM5004 derived from the *V. cholerae* O139 (V911) chromosome (Table 4.5) was linearized by *Hin*dIII digestion. This restriction enzyme cleaves at approximately the center of *vlpA* (at nt 260 of the 516-bp *vlpA* gene). A 1.1-kb *Hin*dIII fragment carrying the Km^R from the plasmid pUWEKT (Table 4.5) was then ligated to pPM5004. The resultant plasmid was transformed into *E. coli* DH5α, and Ap^R Km^R colonies were obtained. Restriction analysis of plasmids isolated from Ap^R Km^R colonies were performed to confirm that the Km^R cartridge was inserted in the correct orientation in the unique *Hin*dIII site such that the transcription of the Km^R cartridge was in

the same direction as that of the *vlpA* gene. This plasmid was designated pPM5013 (Figure 4.2A).

4.2.2.2 Construction of a circular *vlpA*::Km^R gene cassette.

To construct a circular gene cassette, plasmid pPM5013 was restricted with *Bsu*36I. DNA fragments were separated in a 1.0% agarose gel and an approximately 1.7-kb *Bsu*36I fragment was isolated and purified using a QIA-quick Gel Extraction Kit (QIAGEN). Digestion of pPM5013 with *Bsu*36I leads to cleavage at nt 107 of VCR19 located at the 5′end of *vlpA* and at nt 106 of VCR20 located at the 3′end of *vlpA* (Figure 4.2B). The *Bsu*36I linear fragment was recircularized by ligation using T4 DNA ligase (8 units in 200µl reaction). The ligation reaction was performed in a large volume (200µl) to increase the chances of cassette recircularization, while reducing the level of concatamer formation. This recircularized cassette contained a *vlpA*::Km^R gene cassette and a fully reconstituted VCR. The ligation products were run in a 1.5% agarose gel to confirm the presence of the circular cassettes. However, the ligation products included dimers and multimers (Figure 4.2C), but these species would not interfere in the recombination as they were unable to replicate independently. It is well-established that linear DNA fragments are susceptible to degradation by intracellular nucleases, and covalent circularization increases their stability (Collis *et al.*, 1993). As a result, the ligation products were then treated with 150 units exonuclease III (New England Biolabs Inc.) to remove all species except covalently closed circular cassettes.

Figure 4.2 Construction of pPM5013 (A), circular gene cassettes (B), and analysis of the *vlpA* :: Km^R circular cassette (C) .

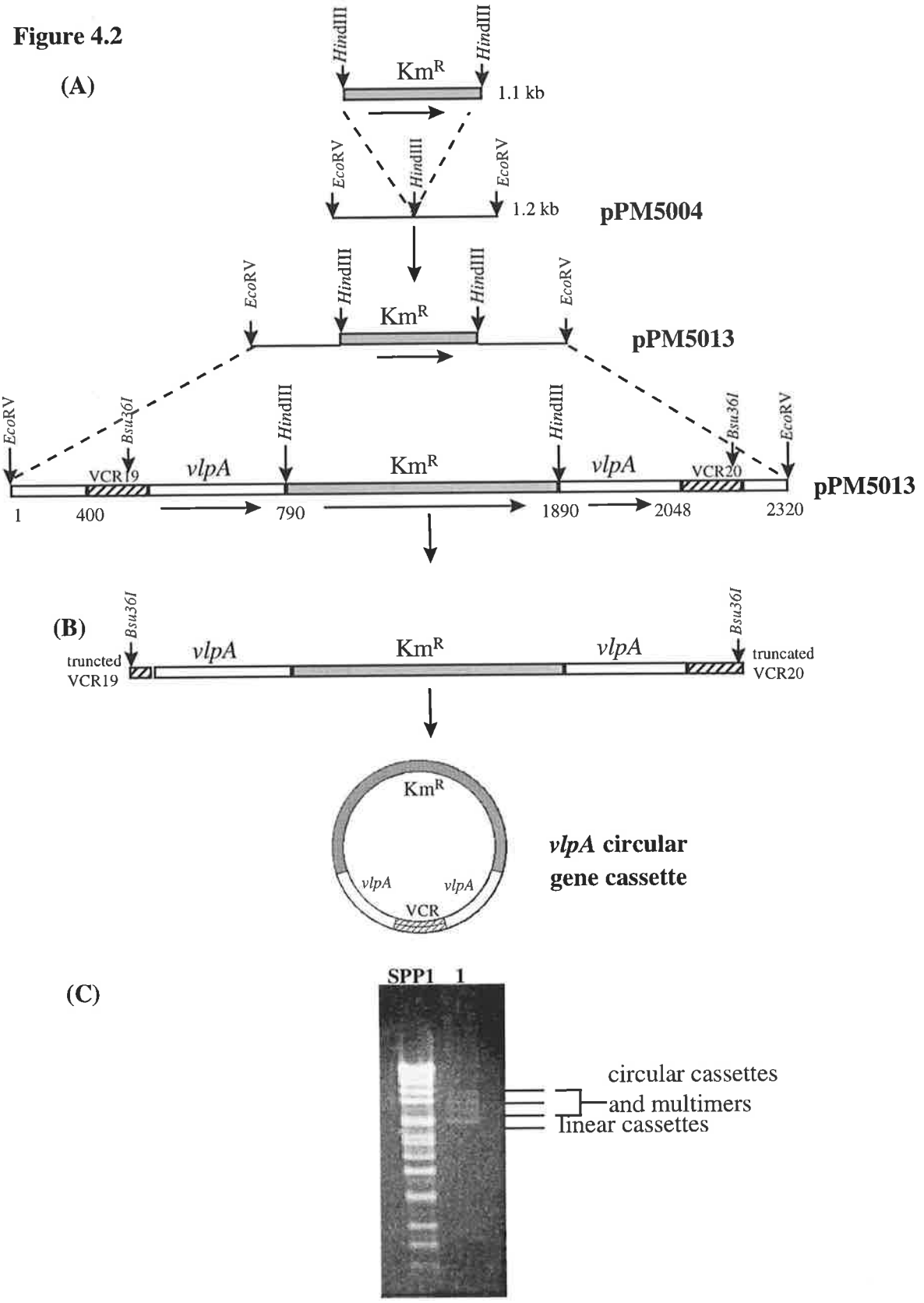
(A) A 1.1-kb Km^R cartridge from pUWEKT digested with *Hin*dIII is represented by gray boxes. The Km^R cartridge was inserted at the unique *Hin*dIII site in pPM5004 to generate pPM5013.

(B) The linear *Bsu*36I fragment containing *vlpA*::Km^R and VCR from pPM5013 was circularized by ligation at both ends of the linear cassette.

Horizontal arrows represent the direction of transcription. The expanded region shows the DNA fragment in greater detail. Numbers refer to the distance from the proximal end of the clone in bp.

(C) The *vlpA*::Km^R gene casstte was recircularized with T4 DNA ligase resulting in a fully reconstructed VCR. This ligated product (lane **1**) was analyzed on a 1.5% agarose gel. Different sizes of cassettes are due to multimers as well as circularized cassettes. (The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.).

Figure 4.2


(A)
HindIII
KmR
HindIII
1.1 kb
EcoRV
HindIII
EcoRV
1.2 kb
pPM5004
EcoRV
HindIII
KmR
HindIII
EcoRV
pPM5013
EcoRV
Bsu36I
HindIII
HindIII
Bsu36I
EcoRV
VCR19
vlpA
KmR
vlpA
VCR20
pPM5013
1
400
790
1890
2048
2320
(B)
Bsu36I
truncted VCR19
vlpA
KmR
vlpA
Bsu36I
truncated VCR20
KmR
vlpA
vlpA
VCR
vlpA circular gene cassette
(C)
SPP1 1
circular cassettes
and multimers
linear cassettes

4.2.2.3 Transformation of circular cassettes into *E. coli* DH5α.

The procedures to introduce circular cassettes into *E. coli* DH5α were performed as described by Collis *et al.* (1993). Initially, the circular cassettes were electroporated into *E. coli* DH5α containing the plasmids pSU2056 and pPM5021 (E5242). pSU2056 is a 1.1-kb *intI*1 gene from Tn*21* (Tnp) cloned into the *Bam*HI-*Rsa*I sites of pUC9 (Martinez and de la Cruz, 1988, 1990). pPM5021 contains two VCRs and *attI*4 from *V. cholerae* 569B cloned into the unique *Hin*dIII site of pSU23, thereby providing a potential recombination site for the circles (Figure 4.3A). *attI*4 is an attachment site, located at the beginning of the first cassette in the *V. cholerae* integron reported by Mazel *et al.* (1998).

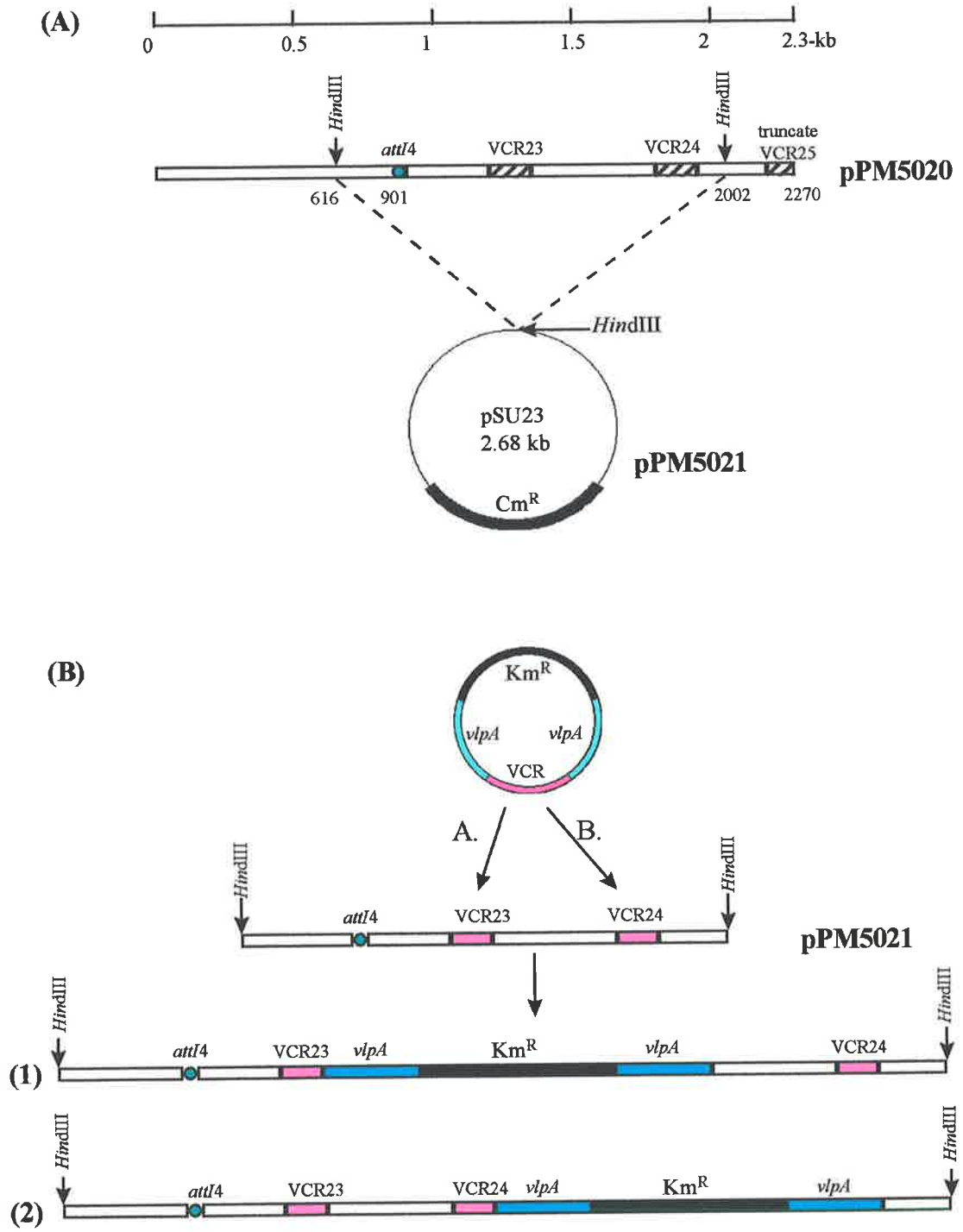
Electroporated cells were incubated with 1 ml of SOC medium and grown for two hours at 37°C with shaking. Aliquots were plated on Luria agar plates containing 50 μg/ml ampicillin for pSU2056, 25 μg/ml chloramphenicol for pSU23, and 50 μg/ml kanamycin to select for the incoming cassette, i.e. the integrated *vlpA*::Km^R circular cassette. A further 200 μl of the transformation mix was diluted 25-fold into Luria broth containing ampicillin, chloramphenicol and kanamycin to a final volume of 5ml. This culture was grown at 37°C overnight shaking, then concentrated by centrifugation. Finally, this culture was plated onto NA with kanamycin added. Because circularized cassettes cannot replicate, Km^R colonies could only arise if circular cassettes integrated into the chromosome (Figure 4.3B). Furthermore, *E. coli* DH5α is a *recA* mutant, therefore homologous recombination cannot occur. However, no colonies were observed in this study after a number of attempts, and therefore another strategy was used to test for the integrase mediated movement of VCR containing gene cassettes.

Figure 4.3 Construction of pPM5021 (A), and the possible recipient DNA structures generated after cassette insertion event (B).

(A) A 1.4-kb *Hin*dIII DNA fragment from pPM5020 was inserted into the unique *Hin*dIII site in pSU23 to generate pPM5021. Numbers refer to the distance from the proximal end of the clone in bp.

(B) Insertion of circular cassettes in pPM5021. The recipient DNA may receive the cassette at different VCR site. This model predicts possibilities 1 and 2 at different VCRs (not to scale).

Figure 4.3


(A)
0
0.5
1
1.5
2
2.3-kb
HindIII
HindIII
attI4
VCR23
VCR24
truncate VCR25
pPM5020
616
901
2002
2270
HindIII
pSU23
2.68 kb
Cm^R
pPM5021
(B)
Km^R
vlpA
vlpA
VCR
A.
B.
HindIII
HindIII
attI4
VCR23
VCR24
pPM5021
HindIII
HindIII
(1)
attI4
VCR23
vlpA
Km^R
vlpA
VCR24
HindIII
HindIII
(2)
attI4
VCR23
VCR24
vlpA
Km^R
vlpA

4.2.3 Cassette deletion mediated by integrase.

The integrase IntI1 from *P. aeruginosa* (from a class 1 integron) recognizes both the *attI* site in integrons and 59-bes (Collis *et al.*, 1992a, b). The integration or excision of gene cassettes catalyzed by IntI1 has been demonstrated with different antibiotic resistance genes (Collis and Hall, 1992a; Collis and Hall, 1992b; Hall and Collis, 1995).

Deletion of gene cassettes can be detected readily if an antibiotic resistance gene is used to tag the cassette (Collis and Hall, 1992a). A plasmid construct carrying a *vlpA*::Km^R gene with flanking VCRs should thus be a useful tool to assess whether such a gene cassette can delete site-specifically (at either *attI* or VCR recombination sites) in the presence of an integrase. Therefore, cassette deletion in both *E. coli* and *V. cholerae* was tested in this study by using both IntI1 and IntI4. Plasmid clones containing the *intI*1 or *intI*4 gene, and the *attI*1 or *attI*4 site were constructed.

4.2.3.1 Construction of plasmid clones to monitor cassette movement via site-specific deletion events.

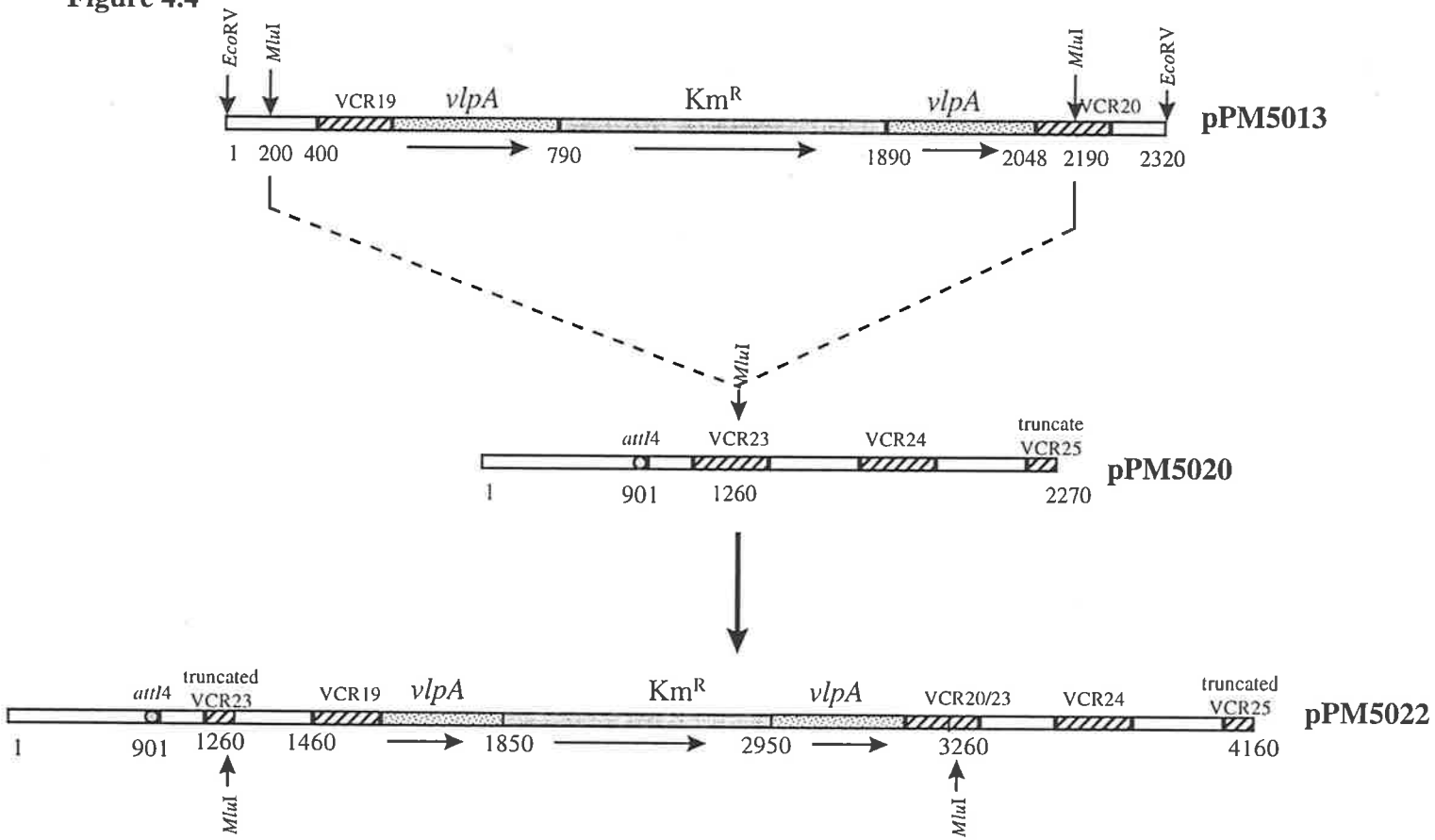
4.2.3.1.1 Construction of pPM5022

Plasmid pPM5013 containing two *Mlu*I sites at each end of the *vlpA*::Km^R was digested with this enzyme to release a 2.0-kb *Mlu*I DNA fragment. This fragment was then ligated into the unique *Mlu*I site of pPM5020 (Figure 4.4). Plasmid pPM5020 is a 2.2-kb *Sph*I-*Sal*I fragment containing an N-terminal truncated *intI*4 gene with the adjacent *attI*4 site and two copies of VCR cloned from *V. cholerae* 569B in the plasmid vector pSU23 (Bartelomé *et al.*, 1991). The resulting plasmid was transformed into *E. coli* DH5α and Km^R Cml^R colonies were obtained. Restriction analysis and DNA sequencing were used to confirm that the *Mlu*I fragment was inserted in the correct orientation such transcription of

Figure 4.4 Construction of pPM5022.

A 2.0-kb *Mlu*I DNA fragment from pPM5013 was inserted into the unique *Mlu*I site in pPM5020 to generate pPM5022. Horizontal arrows represent the direction of transcription. Numbers refer to the distance from the proximal end of the clone in bp.

Figure 4.4


EcoRV
MluI
VCR19
vlpA
Km^R
vlpA
MluI
VCR20
EcoRV
pPM5013
1
200
400
790
1890
2048
2190
2320
MluI
attI4
VCR23
VCR24
truncate
VCR25
pPM5020
1
901
1260
2270
attI4
truncated
VCR23
VCR19
vlpA
Km^R
vlpA
VCR20/23
VCR24
truncated
VCR25
pPM5022
1
901
1260
1460
1850
2950
3260
4160
MluI
MluI

the *vlpA*::Km^R cassette and VCRs from pPM5020 was in the same direction to the *lac* promoter in the vector. This plasmid was designated pPM5022, and *E. coli* DH5α containing this plasmid was called E5243.

4.2.3.1.2 Construction of pPM5023

Plasmid pPM5013 was also digested with *Eco*RV to yield a 2.3-kb DNA fragment containing *vlpA*::Km^R. This fragment was then ligated into pRMH313 provided by Dr. Christina Collis (CSIRO Division of Biomolecular Engineering, Sydney) (Table 4.5) at a unique *Bam*HI site after end filling with Klenow. pRMH313 contains a 263-bp DNA fragment including *attI*1 and the 3′ conserved fragment from Tn*1696* cloned at the *Eco*RV-*Taq*I sites of pACYC184. The ligation mix was transformed into *E. coli* DH5α, and Ap^R Km^R colonies were obtained. Restriction analyses of plasmids isolated from Ap^R Km^R colonies were performed to confirm that the 2.3-kb *Eco*RV fragment was inserted in the correct orientation at the unique Klenow end-filled *Bam*HI / *Eco*RV site in pRMH313, such that the transcription of the *vlpA*::Km^R cassette and VCRs was in the same direction as the *lac* promoter from the plasmid vector. This plasmid was designated pPM5023 (Figure 4.5). *E. coli* DH5α containing this plasmid was called E5247.

4.2.3.1.3 Construction of *E. coli* strains containing integrase and *attI*

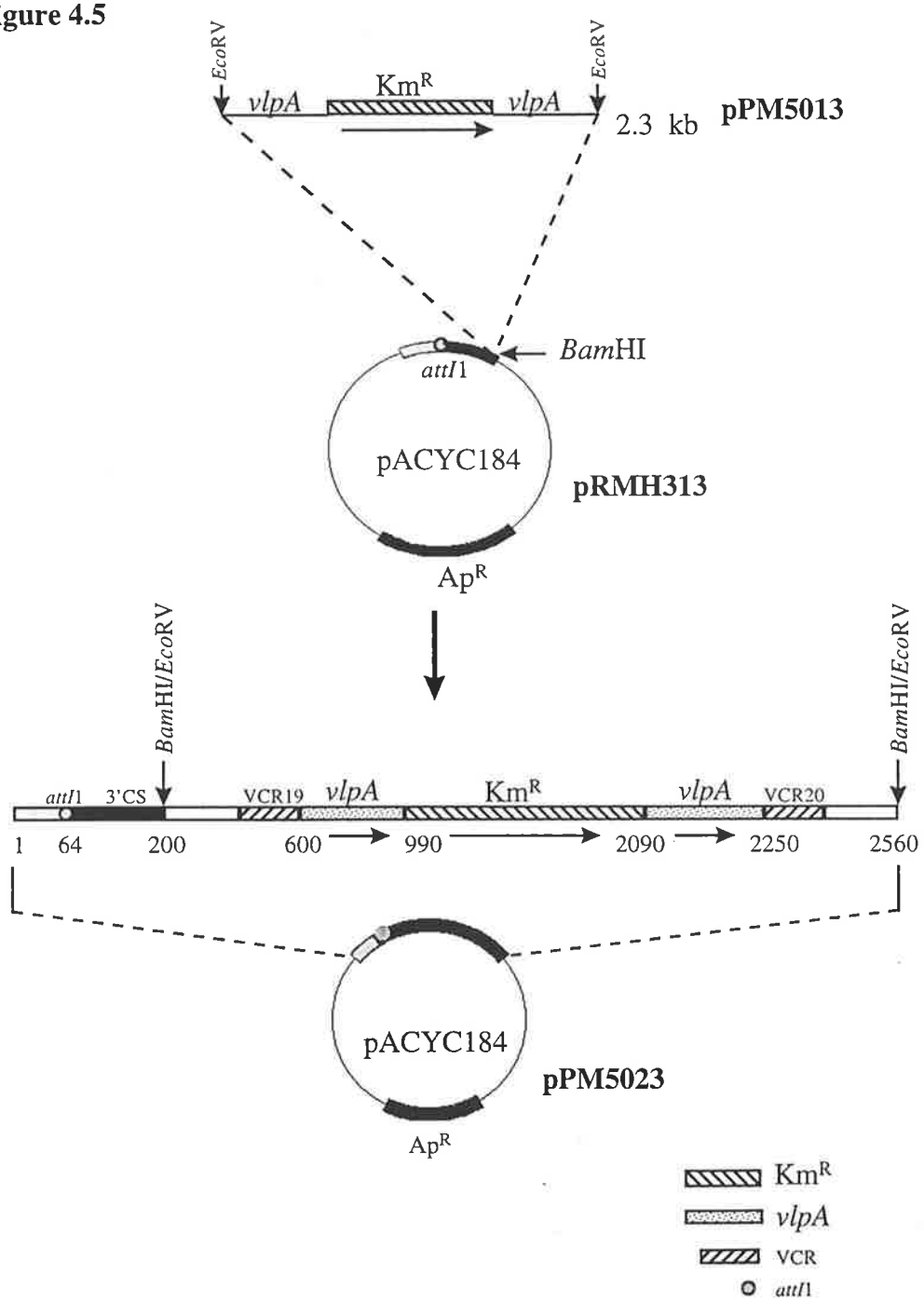
To test for cassette deletions using *vlpA*::Km^R in the presence of integrase IntI1 or IntI4, plasmids pPM5022 and pPM5023 were then introduced separately into *E. coli* containing either IntI1 or IntI4 as described below.

E. coli containing IntI1 on pSU2056 (E5191) was transformed with either pPM5022 or pPM5023. Plasmid DNA preparations were made out from the Ap^R Km^R

Figure 4.5 Construction of pPM5023.

A 2.3-kb *Eco*RV fragment from pPM5013 was inserted into the *Bam*HI site in pRMH313 containing *attI1* and part of the 3′CS, to generate pPM5023. Horizontal arrows represent the direction of transcription. The expanded region shows the DNA fragment in more detail. Numbers refer to the distance from the proximal end of the clone in bp (not to scale).

Figure 4.5


EcoRV
EcoRV
vlpA
KmR
vlpA
2.3 kb
pPM5013
BamHI
attI1
pACYC184
pRMH313
ApR
BamHI/EcoRV
BamHI/EcoRV
attI1
3'CS
VCR19
vlpA
KmR
vlpA
VCR20
1
64
200
600
990
2090
2250
2560
pACYC184
pPM5023
ApR
KmR
vlpA
VCR
attI1

Cml^R colonies. Restriction analysis confirmed that both pSU2056 and pPM5022 or pSU2056 and pPM5023 were retained. *E. coli* DH5α containing pSU2056 together with pPM5022, and pSU2056 with pPM5023 were designated E5246 and E5252, respectively.

Either pPM5022 or pPM5023 was also introduced into *E. coli* containing the *intI*4 gene on pPM5029 (E5227). pPM5029 was constructed by Mr. Christopher Clark (Department of Microbiology and Immunology, The University of Adelaide) and contains a 963-bp fragment of the *intI*4 gene encoding IntI4 cloned into pBAD24 (Guzman *et al.*, 1995). The *intI*4 gene fragment was originated from PCR amplification of the *V. cholerae* 569B chromosomal DNA using the oligonucleotide primers #2596 and #2609 which bind at both ends of *intI4*. The pBAD series of vectors contain the arabinose promoter. In the presence of arabinose, transcription from the pBAD promoter is turned on, and in its absence, transcription occurs only at very low levels (Guzman *et al.*, 1995). Therefore, 0.2% arabinose would be used to induce the expression of the *intI*4 gene from the pBAD promoter in pPM5029. The *E. coli* strains containing either pPM5029 and pPM5022 (E5245) or pPM5029 and pPM5023 (E5260) were obtained from Ap^R Km^R Cml^R colonies. Restriction analysis was performed to establish that both pPM5029 and pPM5022 or pPM5029 and pPM5023 were retained in *E. coli*.

4.2.3.2 IntI1-mediated *vlpA* gene cassette deletion in *E. coli*

In order to detect the deletion event catalyzed by integrase IntI1, either E5246 containing pSU2056 and pPM5022, or E5252 containing pSU2056 and pPM5023 were initially grown in Luria broth supplemented with the appropriate antibiotics overnight shaking at 37 °C. The culture was then diluted 2000-fold into 2 ml of fresh Luria broth containing only chloramphenicol, i.e. to retain the cassette containing plasmid, but with no

direct selection for the cassette itself, i.e. Km^R. Cells were grown to stationary phase, and the culture was again diluted 2000-fold into 2 ml of fresh Luria broth with selection for chloramphenicol. After final overnight incubation with shaking at 37°C, plasmid DNA was isolated and resuspended to a final volume of 50 µl.

DNA was diluted 1 in 10, and 10 µl of this was electroporated into *E. coli* DH5α. This dilution was performed to ensure that fewer than 1 in 100 transformants would receive more than one DNA molecule. Transformants were selected on chloramphenicol plates. The isolation of colonies with the deleted *vlpA*::Km^R gene cassette was screened by plating for kanamycin sensitive (Km^S) colonies. Two thousands colonies from each *E. coli* DH5α strain (E5243, E5246, E5247, E5252) were then patched onto separate nutrient agar plates containing either chloramphenicol, kanamycin or ampicillin, respectively. Plating on nutrient agar containing ampicillin was used to check if cells still retained the integrase. Only six Km^S colonies were observed from the patching of E5252 (IntI1 and *attI*1) and three of these were also Ap^R (Table 4.1). Plasmid DNA from the six Km^S colonies was isolated. Cassette deletion was analyzed by restriction analysis, and PCR amplification using the oligonucleotide primers flanking both ends of the *vlpA*::Km^R cassette (#3142 and #3148). The primer #3142 binds downstream of the VCR20, and #3148 binds at nt 1952 of the plasmid vector pACYC184 and reads forward through the cassette (Figure 4.7A). PCR amplification showed different size of the PCR products (Figure 4.6A). Lane 1 shows a 340-bp fragment, while lanes 2, 3 and 4 show a 1.1-kb PCR product, suggesting that the deletion in the first plasmid (lane 1) was larger than those in the others (lanes 2, 3 and 4). To determine the precise nature of the deletion event, dye-terminator DNA sequencing was performed using the oligonucleotide primers #3142 and #3148. The nucleotide sequences

Table 4.1 **IntI1 and IntI4 mediated cassette deletion in plasmids carrying *vlpA*::Km^R.**

Strain number	Characteristics	Integrase and 59-be (VCR)	No. of transformants screened	No. with phenotype Cml^R/Km^S/Ap^R	No. with phenotype Cml^R/Km^S/Ap^S	Frequency of Km^S
E5243	DH5 α [pPM5022]	*attI*4/VCR	2000	0	0	-
E5246	DH5α [pPM5022&pSU2056]	IntI1/ *attI*4/VCR	2000	0	0	-
E5247	DH5 α [pPM5023]	*attI*1/VCR	2000	0	0	-
E5252	DH5α [pPM5023&pSU 2056	IntI1/ *attI*1/VCR	2000	3	3	3.0×10^{-3}
E5243	DH5 α [pPM5022]	*attI*4/VCR	1500	0	0	-
E5245	DH5α [pPM5022& pPM5029	IntI4/ *attI*4/VCR	1500	0	2	1.33×10^{-3}
E5247	DH5 α [pPM5023]	*attI*1/VCR	1500	0	0	-
E5260	DH5α [pPM5023&pPM5029]	IntI4 *attI*1/VCR	1500	0	0	-

Figure 4.6 PCR amplification of IntI-mediated cassette deletions

(A) IntI1-mediated cassette deletions in pPM5023 (E5247).

PCR amplification using primers #3148 and #3142 to detect gene cassette movement deletions. A 340-bp band in lane **1** indicates that more than one gene cassettes has deleted. Lanes **2-4** show a 1.1-kb band due to a single cassette deletion. Lane **5** shows a 516-bp *vlpA* band representing th gene cassette with no Km^R inserted, and lane **7** is a PCR product of *vlpA*:: Km^R before cassette deletion. Lanes contain **1**, pPM5026 ; **2**, pPM5027 (No. 1) ; **3**, pPM5027 (No. 2);**4**, pPM5027 (No. 3) ; **5**, V911; **6**, negative control; **7**, pPM5023 (*vlpA*::Km^R).

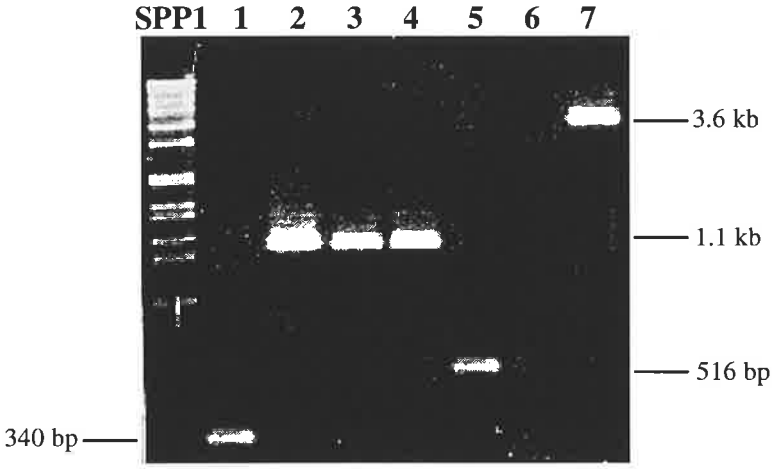
(B) IntI4-mediated cassette deletions in pPM5023.

PCR amplification using primers #2926 and #2993 shows 2.0-kb bands were observed in lanes 2 and 3 resulting from cassette deletions, whereas lane **1** shows a 4.2-kb band from pPM5022 before cassette deletion. Lanes contain **1**, pPM5022 ; **2**, pPM5028 (No. 1); **3**, pPM5028 (No. 2).

(The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.)

Figure 4.6

(A)



(B)

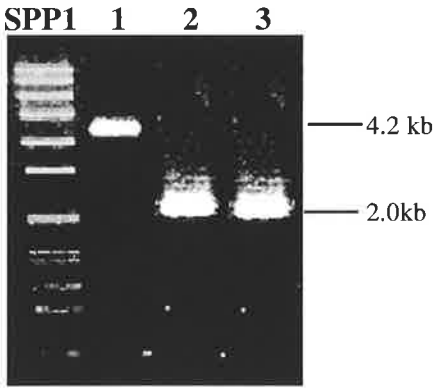
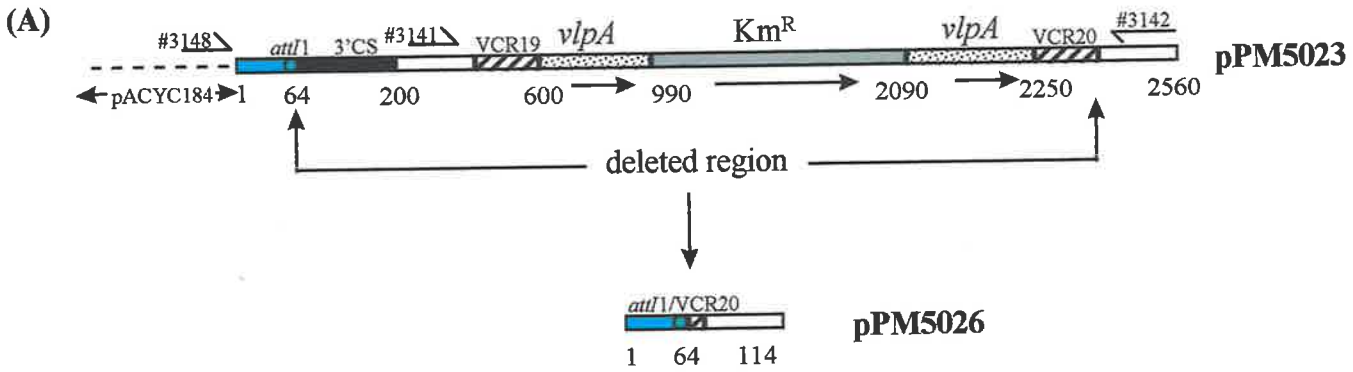


Figure 4.7 IntI1-mediated *vlpA*::Km^R cassette deletion in *E. coli*.

(A) Schematic representation of IntI1-mediated *vlpA*::Km^R cassette deletion in pPM5023, resulting in pPM5026.

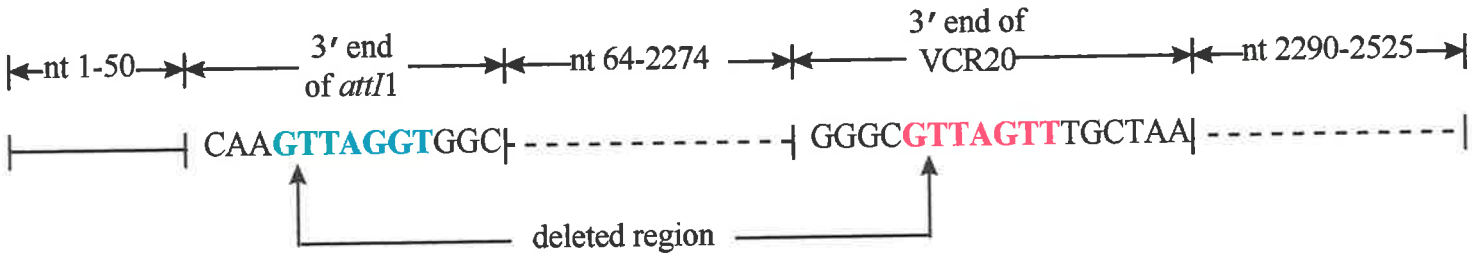
(B) Comparison between nucleotide sequences of pPM5023 and pPM5026. The recombination site occurred at the G and T of the seven-base core-site of both *attI*1 and VCR20. The new junction sequence in pPM5026 is a hybrid between *attI*1 and VCR20.

Figure 4.7

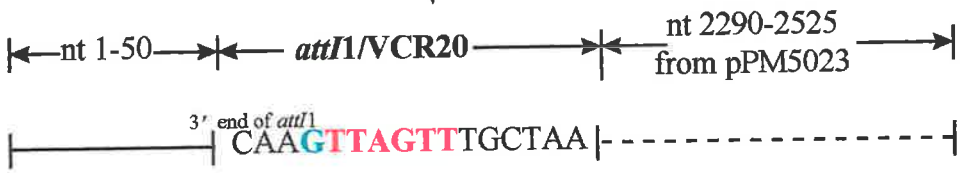


(B)

Before deletion (pPM5023)



After deletion (pPM5026)



of the deletion points in the plasmids obtained from the Km^S colonies is present in **Section 4.2.3.4**.

Interestingly, Km^S colonies were obtained in the presence of pSU2056 (IntI1) with pPM5023 containing the *attI*1 site (E5252), but did not occur in the presence of pSU2056 (IntI1) with pPM5022 containing the *attI*4 site (E5246). This suggests that deletion of the *vlpA*::Km^R cassette either requires the same source of integrase and the *attI* site to be present, (i.e. IntI1 and *attI*1) as for site-specific recombination to occur, or at least such a reaction is more efficient. In the absence of pSU2056 (IntI1), no Km^S colonies were observed, i.e. *E. coli* containing only pPM5022 (E5243) or pPM5023 (E5247). The frequency of Km^S colonies was 3×10^{-3}.

4.2.3.3 IntI4-mediated *vlpA* gene cassette deletion in *E. coli*

A similar approach was used for the isolation and analysis of *vlpA* gene cassette deletions catalyzed by integrase IntI4. E5245 containing pPM5029 (IntI4) and pPM5022 or E5260 containing pPM5029 (IntI4) and pPM5023 were cultured in a similar manner to that described in **Section 4.2.3.2**, except that 0.2% arabinose was added initially to induce *intI*4 expression under control of the pBAD promoter. Transformants were selected on chloramphenicol plates. One thousand five hundred colonies from each *E. coli* strain (E5243, E5245, E5247, E5260) were then patched onto separate nutrient agar plates containing either chloramphenicol, kanamycin or ampicillin. Only two Km^S colonies were observed from the patching of E5245 (IntI4 and *attI*4), and both were Ap^S (Table 4.1). Plasmids from these Km^S colonies were analyzed by restriction analysis and PCR amplification using the T7 and SP6 promoter primers (#2926 and #2993, respectively) which bind to the plasmid vector immediately adjacent to both ends of the *vlpA*::Km^R

cassette. PCR analysis showed that the same size deletion occurred in both Km^S colonies, with the expected size of 2.0-kb if deletion occurred between VCR19 and VCR20 (Figure 4.6B). DNA sequencing was performed to determine the precise location of the deletion event (**Section 4.2.3.4**).

Table 4.1 shows IntI4 (pPM5029) mediated deletion of the *vlpA*::Km^R cassette. It is interesting that deletion of this cassette was only observed in the presence of pPM5029 (IntI4) with pPM5022 (*attI*4) in E5245. No Km^S colonies were detected in E5260 containing pPM5029 (IntI4) and pPM5023 (*attI*1). This result is similar to the deletion mediated by IntI1 in which the same source of integrase and *attI* site, (i.e. IntI4 and *attI*4) seems to be most efficient. The frequency of Km^S colonies was 1.3×10^{-3}.

4.2.3.4 Nucleotide sequence analysis of plasmids resulting from cassette deletion.

To localize the recombination cross-over point more accurately, plasmid DNA from six putative Km^S colonies resulting from IntI1 mediated deletion (**Section 4.2.3.2**) were isolated, and then sequenced. Two types of cassette deletion were found, as determined by PCR amplification analysis, and a representative example of each plasmid was designated pPM5026 and pPM5027. The sequences of the PCR product from pPM5026 and pPM5027 using primers #3142 and #3148 (**Section 4.2.3.2**) were determined. DNA sequencing of the PCR product from pPM5026 revealed that two cassettes had deleted between the *attI*1 and VCR20 site (Figures 4.7A, B). However, sequencing of the PCR product from pPM5027 showed that the deletion point of the *vlpA*::Km^R cassette had occurred between two copies of VCR (VCR19 and VCR20) as shown in Figures 4.8A and B. The sequence data are consistent with recombination occurring between the G and first T in the seven base core-site (GTTRRRY) of both *attI*1 and VCR sites, which is the point of crossover determined

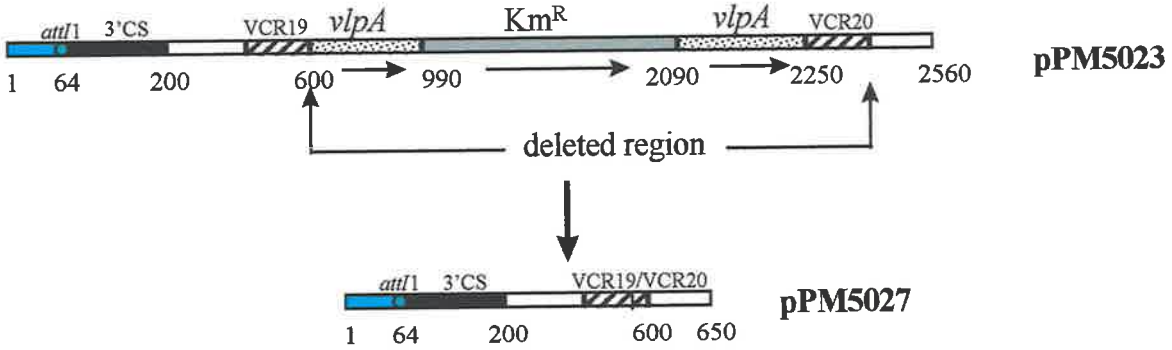
Figure 4.8 IntI1-mediated *vlpA*::Km^R cassette deletion.

(A) Schematic representation of IntI1 mediated the *vlpA*::Km^R cassette deletion in pPM5023, resulting in pPM5027.

(B) Comparison between nucleotide sequences of pPM5023 and pPM5027. Recombination occurred at the VCR19 and VCR20 sites between the G and T of the seven-base core-site of VCR19 and VCR20. The new junction sequence in pPM5027 is the hybrid between VCR19 and VCR20.

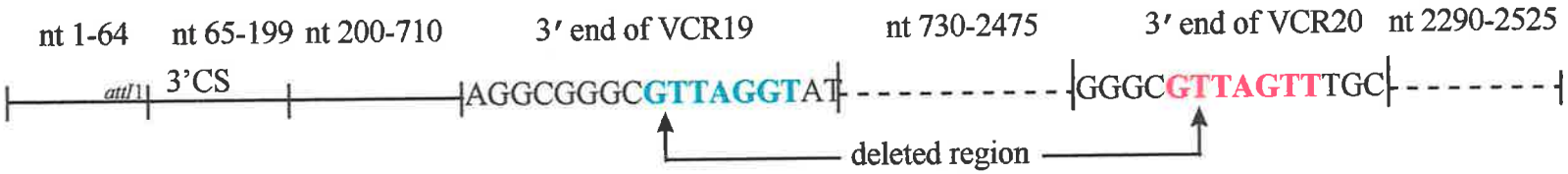
Figure 4.8

(A)

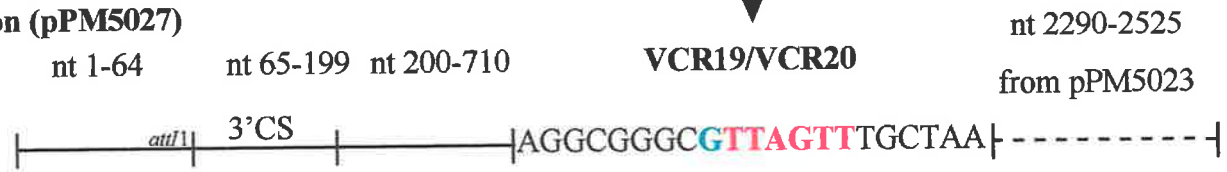


(B)

Before deletion (pPM5023)



After deletion (pPM5027)



by Collis and Hall (1992a, b). DNA sequencing revealed the absence of the gene cassettes, and the formation of a hybrid VCR site between the G residue from *attI*1 with the first T of VCR20 (Figure 4.7B), and between the G residue from VCR19 and the first T of VCR20 (Figure 4.8B).

Plasmids from two kanamycin sensitive derivatives mediated by IntI4 (pPM5029) (**Section 4.2.3.3**) revealed only one type of deletion by restriction analysis and PCR amplification, and a representative plasmid was designated pPM5028. DNA sequencing of the PCR product from pPM5028 using the primers #2926 and #2993 was determined. Sequencing analysis confirmed that the *vlpA*::Km^R cassette had deleted between two copies of VCR (VCR19 and VCR24) as shown in Figures 4.9A and B. Recombination occurred at the seven-base core site, i.e. between the G and first T, in the same manner as observed for IntI1-mediated deletion. Although deletions at the *attI*4 site were not observed in this study, further screening of Km^S isolates may be required to determine whether this can occur.

4.2.3.5 Integrase-mediated cassette deletion in the *V. cholerae* O139 chromosome.

It has been previously mentioned (**Section 4.2.3**) that an array of gene-VCR cassettes is found upstream of *intI*4 in the *V. cholerae* chromosome, and the variation of these genes in the genome of different strains of *V. cholerae* is thought to represent indirect evidence that cassette movement is indeed occurring or has occurred. To facilitate such an event, IntI1 (pSU2056) and IntI4 (pPM5029) were introduced separately into *V. cholerae* O139. It was hypothesized that the overexpression of integrase IntI1 from an alternative promoter should increase the frequency of cassette movement within the chromosome.

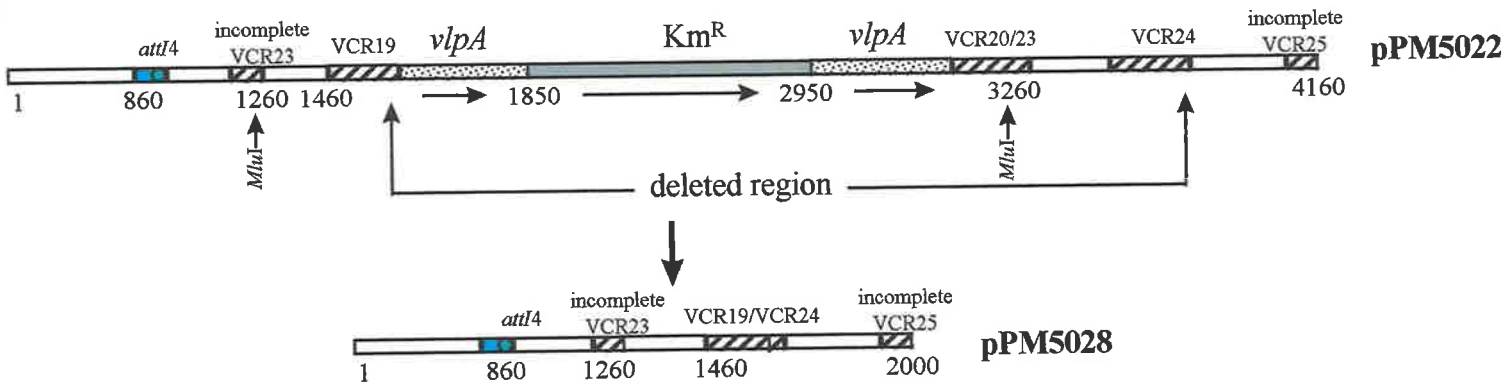
Figure 4.9 IntI4-mediated *vlpA*::Km^R cassette deletion.

(A) Schematic representation of IntI4-mediated the deletion of *vlpA*::Km^R and flanking gene cassettes in pPM5022, resulting in pPM5028.

(B) Comparison between nucleotide sequences of pPM5022 and pPM5028. Recombination occurs at the G and T of seven-base core-site of VCR19 and VCR24. The new junction sequence in pPM5028 is the hybrid between VCR19 and VCR24.

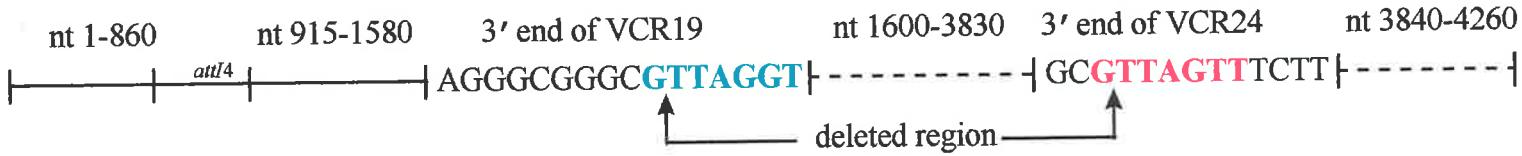
Figure 4.9

(A)



(B)

Before deletion (pPM5022)



After deletion (pPM5028)

4.2.3.5.1 Integrase-IntI1-mediated cassette deletion in the *V. cholerae* O139 chromosome.

Strain V1271 is *V. cholerae* O139 (V911) in which has one chromosomal copy of *vlpA*, i.e. *vlpA*4 (see Chapter 5 for *vlpA* designation) has a kanamycin resistance gene inserted in its coding region and designated V1271 (see Chapter 5 for construction of this *vlpA* mutant). IntI1 was supplied in *trans* via plasmid pSU2056. pSU2056 was introduced into V1271 by electroporation selecting for Km^R Ap^R colonies, and the resulting strain was designated V1285. This strain was initially grown overnight at 37°C with shaking in nutrient broth supplemented with ampicillin and kanamycin. The culture was diluted 2000-fold into 2 ml of fresh Luria broth containing only ampicillin to maintain plasmid pSU2056. Cells were grown at 37°C with shaking to stationary phase and the culture was again diluted 2000-fold into 2 ml of the same fresh media. This culture was grown for two hours, then diluted 10000-fold, and 100 µl of this dilution was plated onto ampicillin plates. Colonies obtained on plates were patched separately onto nutrient agar containing ampicillin, and kanamycin to screen for Km^S colonies. From one thousand colonies patched, three were Km^S and therefore represented potential deletion events. A control strain (V1271) in the absence of IntI1 was treated in a similar manner except for the use of the antibiotics. Km^S colonies were not observed in this strain (Table 4.2).

In order to determine whether the Km^S phenotype was due to a specific cassette deletion at *vlpA*::Km^R, PCR amplification, DNA sequencing and Southern hybridization were performed with the three Km^S *V. cholerae* isolates obtained from colony patching. One of these (V1287) produced a 340-bp PCR product (data not shown) using oligonucleotide primers #3141 and #3142 (see diagram in Figure 4.10A). Conversely, the other two colonies (V1288 and V1289) did not produce an amplified PCR product when

Table 4.2 IntI1 and IntI4 mediated gene cassette deletion in the *V. cholerae* O139 chromosome.

Strain number	Characteristics	Integrase	No. of colony screened	No. with phenotypes		Frequency of
				Km^R Ap^R	Km^S Ap^R	Km^S
V1271	O139 with *vlpA*4::Km^R	-	1000	1000	-	$<1x10^{-3}$
V1285	O139 with *vlpA*4::Km^R and pSU2056	IntI1	1000	997	3	3×10^{-3}
V1286	O139 with *vlpA*4::Km^R and pPM5029	IntI4	1000	1000	-	$-<1x10^{-3}$

Figure 4.10 Southern hybridization and PCR amplification to show Integrase IntI1-mediated cassette deletion in the *V. cholerae* O139 chromosme.

(A) Schematic representation of primers used to amplify PCR product surrounding the deleted region resulting from cassette movement (V1287, V1292, V1309).

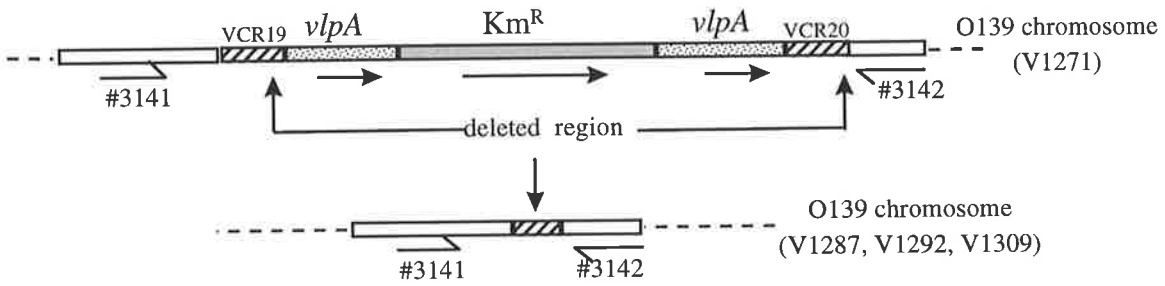
(B) Southern hybridization to confirm that IntI1-mediated cassette deletion in the *V. cholerae* O139 chromosome. Chromosomal DNA from *V. cholerae* strains were restricted with *Eco*RV and probe with PCR DIG-labeled *vlpA* specific probe. Lanes contain chromosome DNA from strains; **1**, V911 (wild-type); **2**, V1269 (*vlpA*1:: Km^R) ; **3**, V1271 (*vlpA*2::Km^R) ; **4**, V 1287; **5**, V1288; **6**, V1289. In all cases (lanes **4-6**) the 2.4-kb (*vlpA*4::Km^R) band is no longer present.

(C) PCR amplification of *vlpA*4 flanking genes in *V. cholerae* strains resulting from IntI1-mediated deletion of cassettes using primers #3141 and #3142. PCR products were not detected in lanes **1, 3, 4, 6, 8, 9** and **11** due to deletion occurring at more distant VCR sites. Lanes contain **1**, V1291 ; **2**, V1292; **3**, V1293; **4**, V1294; **5**, V1295; **6**, V1296; **7**, V1297; **8**, V1298; **9**, V1299; **10**, V1300; **11**, V1301; **12**, V911 (wild-type); **13**, V1271 (*vlpA*1::Km^R); **14**, V1279 (*vlpA* double mutant).

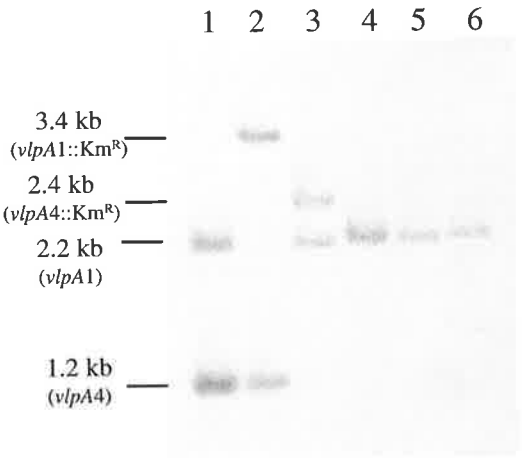
(The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.).

Figure 4.10

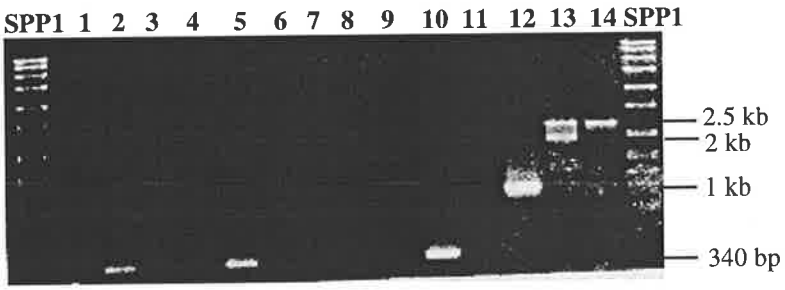
(A)



(B)



(C)



using the same oligonucleotide primers, suggesting that the deletion in these strains occurred beyond the *vlpA* gene cassette at more distant VCR sites, resulting in the loss of the *vlpA* flanking genes. The size of these deletions could not be determined as no DNA sequence data is currently available for the gene cassette in this region of the chromosome. Although the genomic sequence surrounding the *vlpA* gene from *V. cholerae* El Tor strain N16961 is present in the TIGR database, the *vlpA* flanking genes compared to *V. cholerae* O139 strain used in this study are different in this region.

The 340-bp PCR product from V1287 was then sequenced to define the nature of the deletion. Dye terminator sequencing using oligonucleotide primers #3141 and #3142 showed that the deletion occurred between two VCRs (Figure 4.10A). The sequence of the VCR junction in the V1287 chromosome is identical to that of the deletion which was observed in pPM5027 (**Section 4.2.3.4** and Figure 4.8), confirming that the recombination cross-over point lies within the VCR sequences.

Southern hybridization was performed for all three Km^S colonies in order to confirm the loss of *vlpA*::Km^R particularly V1288 and V1289 which did not amplify any PCR products. Chromosomal DNA from V1287, V1288 and V1289 along with V911 (*V. cholerae* O139 wild-type contains *vlpA*1 and *vlpA*4), V1269 (*vlpA*1::Km^R) and V1271 (*vlpA*4::Km^R) (see chapter 5 for construction of this *vlpA* mutant) were restricted with *Eco*RV, and probed with a PCR DIG-labeled specific *vlpA* probe (#1071 and #1072). Hybridization confirmed that the *vlpA*4::Km^R fragment from the three strains (V1287, V1288, V1289) was no longer present, whereas the *vlpA*1 copy was still present in the chromosome (Figure 4.10B).

IntI1 mediated cassette deletion in *V. cholerae* O139 was repeated using the strain V1279 which has an insertional mutation in each of the two *vlpA* copies. The first copy of

vlpA has a tetracycline resistance gene inserted (*vlpA*1::Tet^R) while the second copy has a kanamycin resistance gene (*vlpA*4::Km^R). The integrase IntI1 was also supplied in *trans* via pSU2056. pSU2056 was introduced into V1279 by electroporation selecting for Km^R Tet^R Ap^R colonies, and designated V1290. The *V. cholerae* O139 strain V1279 was used as a control. The procedures were similar to those used in **Section 4.2.3.5.1**.

Screening for Km^S, Tet^S and Km^S Tet^S colonies in V1290 was carried out in a similar manner to that described previously. From one thousand two hundred colonies patched, eleven were Km^S, and only one was Tet^S (Table 4.3). Sensitivity to only one of the antibiotics was observed for each colony (i.e. either Km^R or Tet^R), and no two antibiotic resistance genes were lost in the same colony. No Tet^S or Km^S colonies were observed in V1279 without pSU2056 that did not contain IntI1. Interestingly, the frequency of Km^S colonies (9.2×10^{-3}) was higher than for Tet^S (8.3×10^{-4}).

To see whether the Km^S and Tet^S phenotypes were due to a specific cassette deletion at either *vlpA*::Km^R or *vlpA*::Tet^R, PCR amplification and DNA sequencing were used to confirm the site of excision. Eight of the eleven Km^S colonies did not produce a PCR product using the oligonucleotide primers #3141 and #3142 (Figure 4.10C lanes 1, 3, 4, 6, 7, 8, 9 and 11), suggesting that the deletion of cassettes in these strains occurred at more distant VCR sites. However, the three remaining Km^S colonies produced a 340-bp PCR product using the same primers (Figure 4.10C lanes 2, 5 and 10). DNA sequencing using the primers #3141 and #3142 revealed that only one cassette, *vlpA*4::Km^R, was absent from these strains, and a representative isolate was designated V1292. The sequence obtained is identical to Figure 4.8B.

In addition, the Tet^S colony was analyzed as described above. PCR-amplification was performed using the primers #2905 and #3199 which bind at each end of the

*vlpA*1::Km^R cassette i.e. upstream and downstream of both VCRs flanking this cassette. However, no PCR product was observed using these primers. This is also thought to be due to the deletion occurring at more distant VCR sites resulting in the loss of the *vlpA*-flanking genes as in **Section 4.2.3.5.1**. This strain was referred to V1302, and was then used to screen for the deletion of *vlpA*::Km^R. From one thousand colonies patched, two were Km^S Tet^S. PCR-amplification using primers #3141 and #3142 was performed to confirm that the deletion occurred in these two colonies. PCR product was obtained from only one colony, indicating that the deletion occurred only at the *vlpA*::Km^R cassette. The second colony with no PCR amplified product appeared also to have had a deletion occurring at more distant VCR sites. These strains were designated V1306 and V1307, respectively.

4.2.3.5.2 Integrase IntI4 mediated cassette deletion in the *V. cholerae* O139 chromosome.

Plasmid pPM5029 encoding IntI4 was introduced into V1271 (which contains a Km^R cartridge inserted into the coding region of *vlpA*4) to generate V1286 (see **Section 4.2.3.5.1**). A similar approach to detect cassette movement was used as described in **Section 4.2.3.5.1**. However, no Km^S colonies were obtained from over a thousand colonies patched. One explanation is the expression of IntI4 from the plasmid vector promoter, P_{BAD}, did not produce high enough levels of integrase for function in *V. cholerae*, even though it was sufficiently expressed for cassette movement in *E. coli* as shown in **Section 4.2.3.3**.

Therefore, another plasmid containing the *intI4* gene was constructed by PCR amplification from the *V. cholerae* O139 chromosome, using oligonucleotide primers

#3200 and #3201 (designed from GenBank AF055586). These primers bind at both ends of the *intI*4 gene. The PCR product was cloned into plasmid pGEM-T Easy to generate pPM5034 and the inserts were screened for integrase IntI4 orientated for expression from the *lac* promoter, (as is the case for IntI1 in pRMH313). The integrase IntI4 was supplied in *trans* via pPM5034 by electroporation of this plasmid into V1279, selecting for Km^R Tet^R Ap^R colonies. This strain was designated V1305. The *V. cholerae* O139 strain V1279 was used as a control. Screening for Km^S or Tet^S colonies in V1305 was carried out in a similar manner to that described in **Section 4.2.3.5.1**.

Two Km^S colonies were isolated from one thousand two hundred colonies patched in the IntI4 over expressing strain (Table 4.3). In order to confirm that the Km^S phenotype was due to a specific cassette deletion at *vlpA*::Km^R, PCR amplification and DNA sequencing were performed. Two types of deleted cassettes were observed using PCR analysis. One corresponded to the deletion of a single cassette and produced a 340-bp PCR product using oligonucleotide primers #3141 and #3142 (*vlpA*::Km^R) (Figure 4.11A lane 3) This strain was designated V1309. Another Km^S colony could not detect a PCR product using the same primers (lane 4), and was designated V1308. DNA sequencing using the same primers revealed that only one cassette, *vlpA*4::Km^R, was absent from V1309. The sequence obtained is identical to Figure 4.8B and that the recombination occurred between the G and the first T in the seven-base core-sites (GTTRRRY) of two VCR sites and is similar to the deletion that occurred in plasmids pPM5022 and pPM5023 (**Section 4.2.3.3** and Figure 4.5). However, the other deletion in V1308 involves a larger DNA fragment. While the first deletion in V1309 was confirmed by DNA sequencing of the PCR fragment, the extent of the larger deletion in V1308 could not be determined as no DNA sequence data is currently available for this region of the *V. cholerae* O139 chromosome. The

Table 4.3 IntI1- and IntI4-mediated gene cassette deletion in *V. cholerae* O139 double-mutant (V1279)* chromosome.

Strain number	Plasmids	Integrase	No. of colony screened	No. with phenotypes					Frequency of Km^S or Tet^S	
				Km^R Ap^R	Km^S Ap^R	Tet^R Ap^R	Tet^S Ap^R	Km^S Tet^S	Km^S	Tet^S
V1279	-	-	1,200	1,200	-	1,200	-	-	$<8.3 \times 10^{-4}$	$<8.3 \times 10^{-4}$
V1290	pSU2056	IntI1	1,200	1,189	11	1,199	1	-	9.2×10^{-3}	8.3×10^{-4}
V1305	pPM5034	IntI4	1,200	1,200	2	1,200	-	-	1.6×10^{-3}	$<8.3 \times 10^{-4}$

* V1279 has an insertional mutation in each of the two *vlpA* copies. (see Chapter 5 for construction of this *vlpA* mutant).

Figure 4.11 IntI4-mediated *vlpA*::Km^R cassette deletion in the *V. cholerae* O139 chromosome

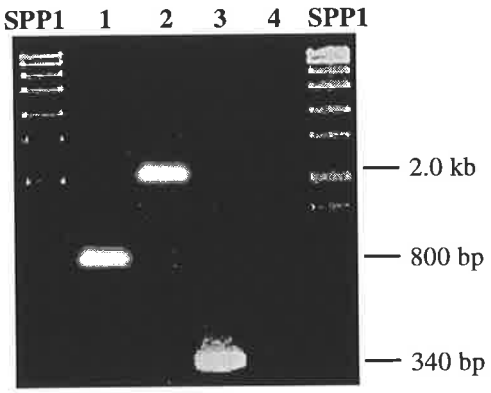
(A) PCR-amplification of *vlpA*4 flanking genes in *V. cholerae* strains resulting from IntI4-mediated deletion of cassettes using primers #3141 and #3142. PCR product was not detected in lane **4** due to a longer deletion occurring at more distant VCR sites, whereas a 340-bp was observed in lane **3**. Lanes contain **1**, V911 (wild-type); **2**, V1279 (*vlpA* double mutant); **3**, V1309; **4**, V1308.

(B) Southern hybridization to confirm the IntI4-mediated cassette deletion in the *V. cholerae* O139 chromosome. Chromosomal DNA from *V. cholerae* strains were restricted with *Eco*RV and probed with PCR DIG-labeled *vlpA* specific probe. Lanes contain strains; **1**, V1308 which *vlpA*4 is not observed in this strain; **2**, V911 (wild type); **3**, V1269 (*vlpA*1:: Km^R) ; **4**, V1271 (*vlpA*2::Km^R).

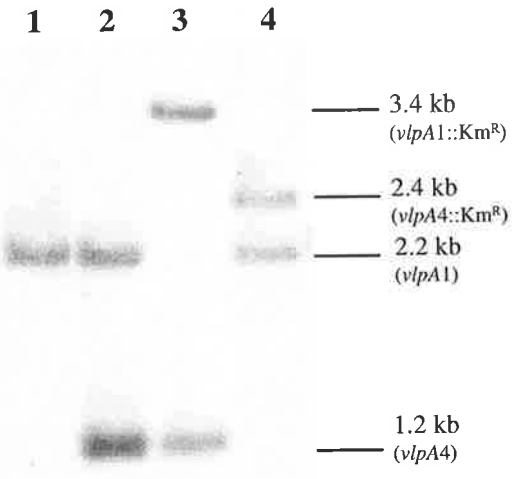
(The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.).

Figure 4.11

(A)



(B)



genomic sequence surrounding the *vlpA* gene in *V. cholerae* El Tor strain N16961 is different in this region.

Southern hybridization was performed on the V1308 chromosome to confirm the loss of *vlpA*4::Km^R. This was of particular interest as no PCR products could be amplified, implying the loss of more than one cassette, and probing by Southern hybridization would confirm the deletion of *vlpA*4. As described previously (**Section 4.2.3.5.1**), chromosomal DNA from V1308, V911 (wild-type), V1269 (*vlpA*1::Km^R) and V1271 (*vlpA*4::Km^R) were restricted with *Eco*RV, and probed with a DIG-labeled specific *vlpA* PCR probe (#1071 and #1072). Hybridization was not observed in *vlpA*4::Km^R, confirming that this cassette was no longer present, whereas the *vlpA*1 copy was still present in the V1308 chromosome (Figure 4.11C).

4.2.3.6 Excised *vlpA* cassettes are covalently closed circular molecules.

It has been reported that, upon excision, the gene cassettes exist as covalently closed circles (Collis and Hall, 1992b), although, unlike plasmids, they are capable of autonomous replication. Therefore, the detection of these circles was attempted from stationary phase stage cells of V1285 and V1305. Primers were designed to read out from the middle of *vlpA* (Figure 4.12A) and could not amplify a PCR product unless it existed in a circularized form, i.e. a circle gene cassette. A 640-bp PCR product was detected on a 1.0% agarose gel. However, this band was indistinct because these excised circular cassettes do not contain an origin of replication, and their numbers are usually very low, although each excision gives rise to a single molecule of the circular cassette (Collis and Hall, 1992b). To increase some sensitivity, this PCR product was then reamplified using

Figure 4.12 Analysis of the putative *vlpA* circular cassette excised from the *V. cholerae* O139 chromosome.

(A) Schematic representation of primers used to amplify across the circular cassette excised from the *V. cholerae* O139 chromosome. The oligonucleotide primers #2816 and #2817 read out from the centre of *vlpA* and can amplify the PCR product only when *vlpA*::Km^R is present in a circular form.

(B) PCR products from the excised cassette using primers #2186 and #2187. A 640-bp product resulting from amplification of circular cassettes is present in lanes **6-8**. Lanes contain **1**, V911 (wild-type); **2**, V1280 (*vlpA*1::Tet^R); **3**, V1271 (*vlpA*4::Km^R); **4**, pPM5013 (pBS::*vlpA*::Km^R); **5**, pPM5007 (pGEM-T:: *vlpA*); **6**, circular cassette no 1; **7**, circular cassette no 2; **8**, circular cassette no 3.

(C) Southern blot of DNA from the PCR product of circular cassettes (lane**1**), compared with DNA from pPM5007 cleaved with *Eco*RV (lane**2**) , pPM5013 cleaved with *Sac*I and *Sac*II (lane**3**) , and pGEM5Zf+ cleaved with *Eco*RV as a negative control (lane**4**). Hybridization was performed using a DIG-labeled *vlpA* specific PCR probe. An approximately 1-kb band is shown due to the excision of further VCR sites.

(The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.).

Figure 4.12

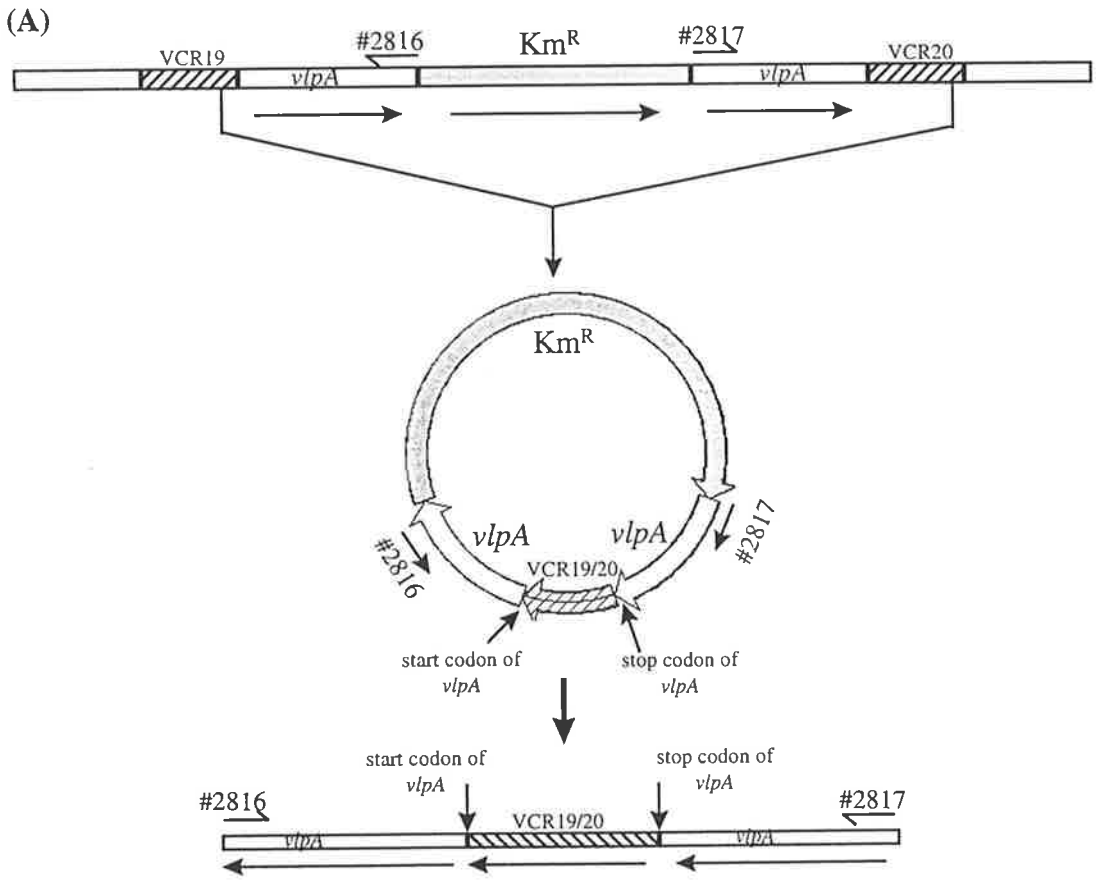
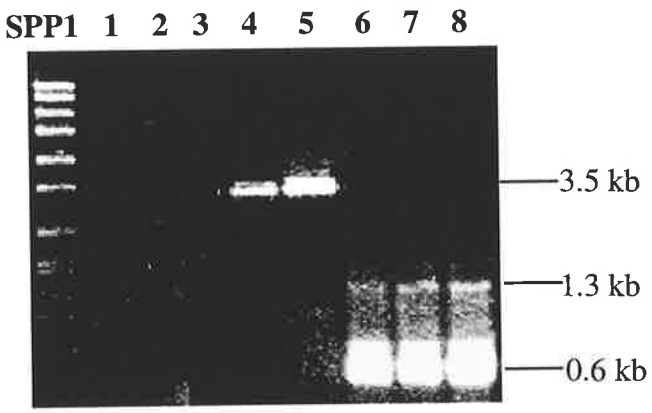
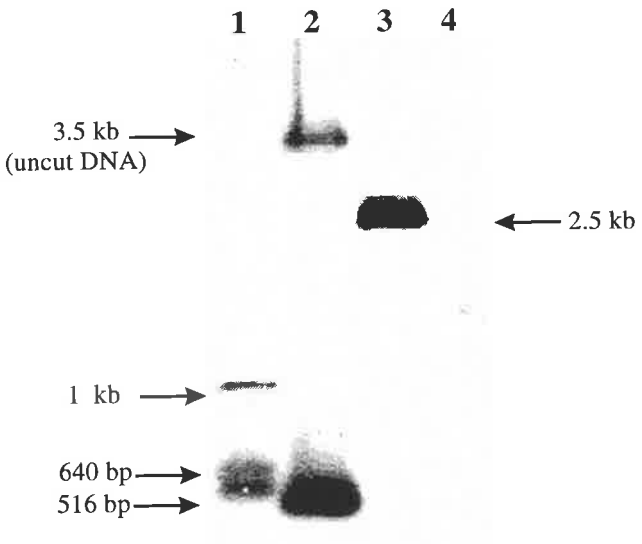

(A)
VCR19
#2816
Km^R
#2817
VCR20
vlpA
vlpA
Km^R
vlpA
vlpA
#2816
#2817
VCR19/20
start codon of vlpA
stop codon of vlpA
start codon of vlpA
stop codon of vlpA
#2816
#2817
VCR19/20
vlpA
vlpA

Figure 4.12

(B)



(C)



the same oligonucleotide primers #2816 and #2817 to obtain sufficient product to enable sequence analysis (Figure 4.12B).

Southern hybridization was performed using the DIG-labeled specific-*vlpA* PCR probe (#1071 and #1072) to confirm that these PCR products did indeed contain *vlpA*. The reamplified PCR product was compared with the positive controls pPM5007 and pPM5013 (see Table 4.5 and Figure 4.2), and the pGEM-5Zf(+) vector alone as a negative control. A 640-bp band was detected as shown in lane 1 in Figure 4.12C. However, an additional larger band of 1.1-kb also hybridized to the *vlpA* probe. This possibly represent circles containing more than one gene cassette, as larger deletions have been found in the *V. cholerae* chromosome (see **Sections 4.2.3.2 and 4.2.3.3**). In addition, DNA sequencing was performed from these PCR products using primers #2816 and #2817 which would read across the VCR junction, if it was indeed a circularized cassette (Figure 4.13A). DNA sequence analysis confirmed the presence of circular cassettes, as the sequence starts from the middle of *vlpA* using the primer #2187 and reads downstream until the stop codon of *vlpA* (Figure 4.13B). This sequence is then continued across the full VCR and connects to the start of *vlpA* again. Additionally, the primer #2186 showed the same result except that the sequence goes in the opposite direction.

4.3 Discussion

Collis and Hall (1992a,b) demonstrated that an integrase can functions to catalyze the insertion or deletion of gene cassettes in integrons. Therefore, *vlpA*, which is one of the ORFs in the *V. cholerae* chromosomal integron, was used in this study to test for its ability to be mobilized in the presence of an integrase IntI1 or IntI4. Sequencing analysis in Chapter 3 and Figure 4.13 confirmed that *vlpA* is a gene cassette that is always associated

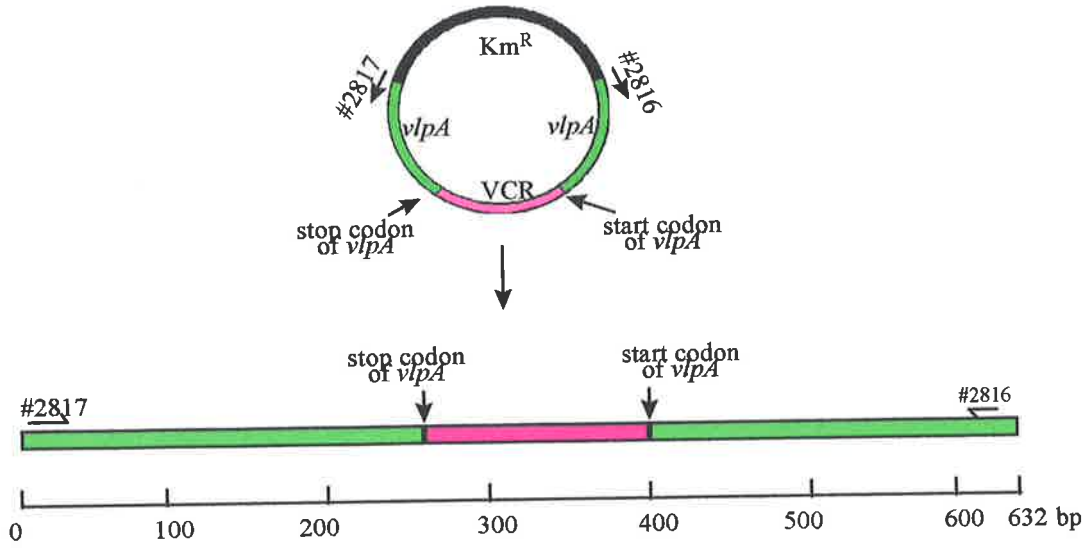
Figure 4.13 Sequence analysis of the putative *vlpA* circular cassette excised from *V. cholerae* chromosome.

(A) Schematic representation to show the oligonucleotide primers (#2816 and #2817) used for the PCR amplification and dye terminator sequencing, including the PCR product resulting from amplification of a circular cassette.

(B) Sequence of *vlpA* circular cassette using oligonucleotide primers #2816 and #2817 (see text in **Section 4.2.3.6** for details).

Figure 4.13

(A)



(B)

#2817

```
1   GAATGGCTCA ACAGATGGCT ATCTGAAGGT TTCATTTTTT GGCCCGTTTT ATGGCTCCTA
61  CGTAGTGTTT GAGTTAGACC GTGAAAACTA CAGTTATGCT TTTGTGTCAG GGCCGAATAC
121 AGAATATCTG TGGTTACTTT CAAGAACGCC GACTGTAGAA CGAGGCATTC TGGACAAGTT
181 CATAGAAATG TCGAAAGAGC GTGGTTTTGA TACAAATCGG CTTATTTACG TTCAGCAGCA
241 ATAAATACCT AACAAACGCC TCAAGAGGGA CTGTCAACGC GTGGCGTTTT CAGTCCCATT
301 GAGCCGAAGT GGTTACGGTT GTTGTGTTTG AGTTTAGTGT TATGCGTTGC CAGCCCCTTA
361 GGCGGGCGTT AGGTATTTGG AGACATCATG AGAGCTATCT TTTTGATTCT TTGCGCTGTT
421 TTATTGAATG GCTGCTTGGG CATGCCCGAA TCAGTAAAAC CGGTGTCGGA TTTTGAACTG
481 AACAACTATT TAGGCAAATG GTACGAGGTT GCTCGACTCG ATCACTCCTT TGAAAGAGGT
541 TTAAGTCAGG TTACTGCGGA ATACCGTGTT GGAAATGATG GTGGTATTTC GGTTCTTAAT
601 CGTGGTTATT CTGAAGAGAA AGGTGAGTGG AA
```

#2816

with a short imperfect inverted repeat sequence termed VCR, located downstream of the *vlpA* coding region. The *vlpA* gene forms a unique family of small mobile elements that include only a single *vlpA* coding region and a 59-be or VCR in this study. In addition, the *vlpA* cassette can exist in a circular form as shown in Figure 4.1B. Free circular *vlpA* cassettes can be generated by religation of both ends of the linear cassettes *in vitro*, thus each circular cassette will contain a unique *vlpA* sequence associated with VCR.

Testing for gene cassette movement was initially performed in *E. coli* using plasmids expressing the integrase IntI1 isolated from a class 1 integron in *P. aeruginosa*. IntI1 is known to function in class 1 and 2 integrons, while the newly identified integrase IntI4 from *V. cholerae* has not been studied to date. Either IntI1 or IntI4 was supplied in *trans* from a plasmid coexisting with a second plasmid carrying an *attI* site and a *vlpA*::Km^R cassette flanked by VCRs. Cassette movement was detected by screening for Km^S, looking for deletion of a *vlpA*::Km^R cassette in the presence of either IntI1 or IntI4. IntI4 has successfully been shown for the first time to be active in this study. However, the frequency of deletion catalyzed by IntI1 was higher than IntI4. It is possible that pSU2056, which carries the *intI*1 gene, expressed the integrase at higher levels than for of pPM5029 which carries the *intI*4 gene. pSU2056 contains *intI*1 from Tn*21* downstream of the β-galactosidase promoter of pUC9. This construction results in increasing the expression of IntI1 at least 1000-fold and a higher frequency of cointegrate formation (Martinez and de la Cruz, 1990). Conversely, pPM5029 contains *intI*4 from *V. cholerae* 569B cloned into pBAD24 carrying the P_{BAD} promoter from the arabinose operon and its regulatory gene, *araC* (Guzman *et al.*, 1995).

In general, genes cloned under the control of the *araC*-P_{BAD} promoter system are efficiently repressed. The concentration of arabinose used in inducing conditions, the

ability of the strain to degrade arabinose, the type of medium and the amount of other carbon sources present in the medium are all factors that contribute to an efficient repression of gene expression (Miyada *et al.*, 1984; Russell *et al.*, 1989; Guzman *et al.*, 1995). In this particular case, these factors may affect the expression of the *intI*4 gene by the *araC*-P_{BAD} promoter.

Cassette movement measured by excision or gene loss was observed only in the presence of the plasmid encoding integrase IntI1 (pSU2056) or IntI4 (pPM5029). Deletion occurred only when the *attI*1 site was present with integrase IntI1 or the *attI*4 site was present with integrase IntI4. It seems likely that the presence of the integrase and a recombination site, *attI,* from the same integron is important. IntI1 and the *attI*1 site are from a class 1 integron in *P. aeruginosa*, whereas IntI4 and *attI*4 are from the *V. cholerae* chromosomal class 4 integron. Thus, it is possible that IntI1 requires an initial interaction with *attI*1, and IntI4 with *attI*4.

Gravel and colleagues (1998) used purified maltose-binding protein fused with the integrase (MBP-IntI1) and native IntI1 protein, and gel retardation assays with fragments containing the complete and partial *attI*1 site for the detection of IntI1-DNA complexes. Chemical modification of specific nucleotides within the *attI*1 site was used to investigate their interference with binding of the integrase protein. IntI1 bound specifically to four regions in the *attI*1 site and a GTTA consensus sequence is found in three of the four regions, indicating that the integrase interacts with both sides of the DNA helix in *attI*1. In addition, Gravel *et al.* (1998) also showed that the interaction of IntI1 with the *attI*1 site results in the formation of four distinct complexes. They suggested that the GTTA residues are the most important in sequence recognition by the integrase protein. Hansson *et al.* (1997) reported that a few integrons carry the sequence GTTG, instead of the GTTA

usually found at the crossover site. However, they have been shown to be just as active as other integron *attI* sites in site-specific recombination. These observations could reflect the possibility that IntI1 is more promiscuous and is able to recognize all *attI* sequences. The effect of these binding sites on the recombination reaction have not yet been examined.

In the integron site-specific recombination systems examined to date, there are no reports as to whether cooperative binding occurs between an integrase and *attI* site from the same class of integron, and consequently whether integrases recognize alternative *attI* sites or 59-be. However, in this study both IntI1 and IntI4 recognize VCR which is a 59-be. There is only one difference between *attI*1 and *attI*4 which appears at the seven-base core site (GTTRRRY). This site in *attI*1 is GTTAGGT, and *attI*4 is GTTAGTT. The difference between purine G and pyrimidine T in *attI*1 and *attI*4 may affect the interaction or sequence recognition by integrase IntI1 to *attI*4 or IntI4 to *attI*1. Further study could be undertaken by base substitution of these residues.

Gravel *et al.* (1998) suggested that cassette integration usually involves the *attI* site, while cassette excision uses 59-bes. However, both VCR and the *attI* site have been shown to be used as the recombination sites in the deletion events in this study. Nucleotide sequence analysis (**Section 4.2.3.4**) is consistent with other data (Collis and Hall, 1992a) which have shown that the preferred deletion site is the cassette core site GTTRRRY, located at the 3′ end of *attI* and VCR. Recchia and Hall (1994) reported that the structure of *attI*1 is very different from that of a 59-be. Similarly, both the *attI*1 and *attI*4 sequences used in this study lack both an identifiable 7-bp inverse core site (AACYYYR) and the extensive inverted repeat associated with VCR. The only obvious feature that *attI*1 and *attI*4 share with VCR is the 7-bp core site (GTTRRRY).

Sequences of the VCR junctions were determined of the experimentally-derived deletion events present in pPM5022 and pPM5023. A deleted *vlpA*::Km^R gene cassette from pPM5022 showed the fusion of two VCRs forming a hybrid VCR at the GT residues of the core site, as well as the deletion of more than one gene cassette. In addition, two plasmids derived by cassette deletion from pPM5023 demonstrated the fusion between the *attI*1 and VCR sites (pPM5026), and between two VCRs (pPM5027). When a cassette excised, the deletion occurred at the 5′end between G and the first T of the cassette and the last base at the 3′end which is a G residue. These experimental data demonstrate directly that the gene cassette can be deleted at the core site GTT catalyzed by integrase. This result is also supported by the fact that deletions occurred only in the presence of integrase.

Having established that IntI1- and IntI4-mediated cassette deletion of *vlpA* can occur from plasmids constructed in *E. coli*, it was of interest to see whether the cassette movement could be made to occur in *V. cholerae*. The *vlpA*::Km^R cassette was used as a marker for this event and enhanced with a high gene copy number of integrase. The integrase was supplied in *trans* via pSU2056 (IntI1), or pPM5029 (IntI4) or pPM5034 (IntI4). However, the results from **Section 4.2.3.5** showed that only IntI1 was capable of catalyzing the deletion of gene cassettes in the *V. cholerae* chromosome. This may reflect a lower level of expression of IntI4 from the pBAD promoter, compared to that of IntI1 in pSU2056 and IntI4 in pPM5034. All deletions were screened for the loss of *vlpA*::Km^R or Tet^R, and PCR-amplification was performed to confirm the event had occurred. The lack of any PCR-amplified product in a number of the Km^S or Tet^S *V. cholerae* colonies suggested that deletions had occurred at other VCRs located either further upstream or downstream of the *vlpA* cassette. Further analysis is required to determine the extents of cassette deletion,

although at present no corresponding DNA sequence data are available for these flanking regions.

The boundary generated after the gene cassette excision in *V. cholerae* showed identity to the excision events in *E. coli* when VCRs and *vlpA* were present on a plasmid (**Sections 4.2.3.2 and 4.2.3.3**). The ability of integrase IntI1 to recognize VCRs has previously been shown using a conduction assay by Mazel *et al.* (1998). They found that a cassette associated with VCR can be directed to the insertion site of integrons when catalyzed by IntI1. However, the integrase IntI1 or more importantly IntI4 from *V. cholerae* has not previously demonstrated to function in *V. cholerae*. Therefore, this study was the first to demonstrate gene cassette deletion from the *V. cholerae* chromosome by overexpression of either integrase IntI1 or IntI4.

It was proposed that since cassettes were excised from the *V. cholerae* chromosome then the non-replicating circular gene cassettes should also be present in these strains. PCR-amplification was used to detect these circles using primers reading outwards from the center of the *vlpA* gene (Figure 4.13), thereby only generating a product if a circular form existed. Collis and Hall (1992a) attempted to obtain direct evidence for circular gene cassettes but were not successful because of the presence of plasmids containing two copies of the cassette gene insert in the DNA extract used. However, no plasmids containing the *vlpA* gene cassette were used in this study. As a result, there was no plasmid interference. The PCR product from the proposed circular cassettes were sequenced (Figure 4.13) and this data strongly support the presence of these circles e.g. VCRs are hybrids of the two flanking *vlpA* VCRs.

As mentioned in **Section 4.2.3.5.1**, more than one gene cassette also appeared to be deleted from the *V. cholerae* chromosome. A larger PCR product of approximately 1.3-kb

also appeared when primers #2816 and #2817 were used to detect the circles. It is possible that dimers between two *vlpA*::Km^R might occur, but in this case the PCR product from dimers should be approximately 2.3 kb (Figure 4.14). Therefore, the PCR-product may come from the adjacent cassette which excised as a circle together with *vlpA*::Km^R as shown in Figure 4.15.

The deletion of more than one gene from an integron has been reported (Wiedemann *et al.*, 1987; Schmidt *et al.*, 1989). In addition, Southern hybridization (Chapter 3) demonstrated that there is only one copy of *vlpA* in *V. cholerae* El Tor strain C31, whereas other strains of El Tor showed at least four copies of *vlpA*. It was reported that a large part of this region was deleted from the chromosome (van Dongen and De Graaf, 1986; Barker and Manning, 1997), but there is no report of the deleted fragment size and the mechanism by which this fragment was deleted. It is possible that the loss of this region may be due to a deletion event catalyzed by the integron integrase IntI4. Future studies could analyze this region to see whether this excised region is within the integron area, and if the deletion end points are associated with VCR sequences.

The results in this study demonstrate that VCR functions as a 59-be i.e. a specific recombination site that results in cassette deletion. Although in this case cassette insertion has not been observed, Mazel *et al.* (1998) have shown the integration of a single VCR containing cassette or the co-integration of a plasmid carrying the cassette into a target integron when catalyzed by IntI1. Initial insertion of cassette genes into an integron involves a circularized gene unit, followed by a single site-specific recombination event catalyzed by the integron integrase (Hall *et al*, 1991). Therefore, the excised gene cassettes found in this study could be reinserted at a new location in the *V. cholerae* chromosomal integron. However, Southern hybridization analysis of the *V. cholerae* antibiotic sensitive

Figure 4.14 Schematic representation to show formation of dimers or multimers of *vlpA*::Km^R cassette.

The *vlpA*::Km^R cassettes deleted from the chromosome can form dimers or multimers. PCR-amplification using primers #2186 and #2187 will give rise to different PCR products as shown by the horizontal line. Numbers refer to the expected fragment sizes in kb.

Figure 4.14

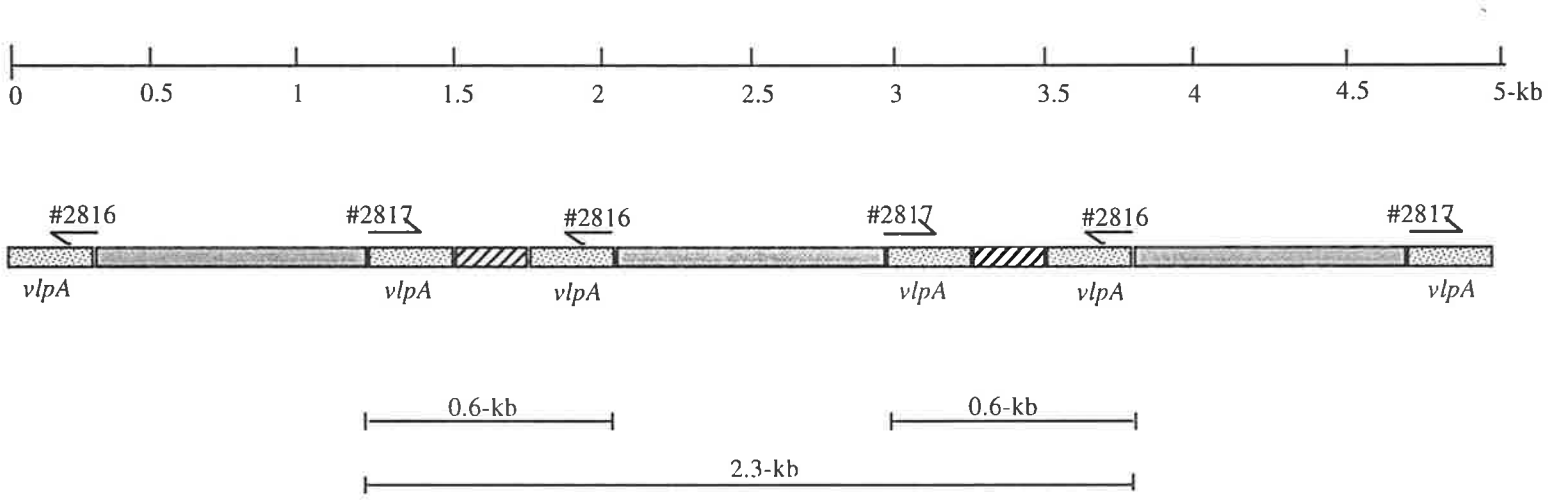
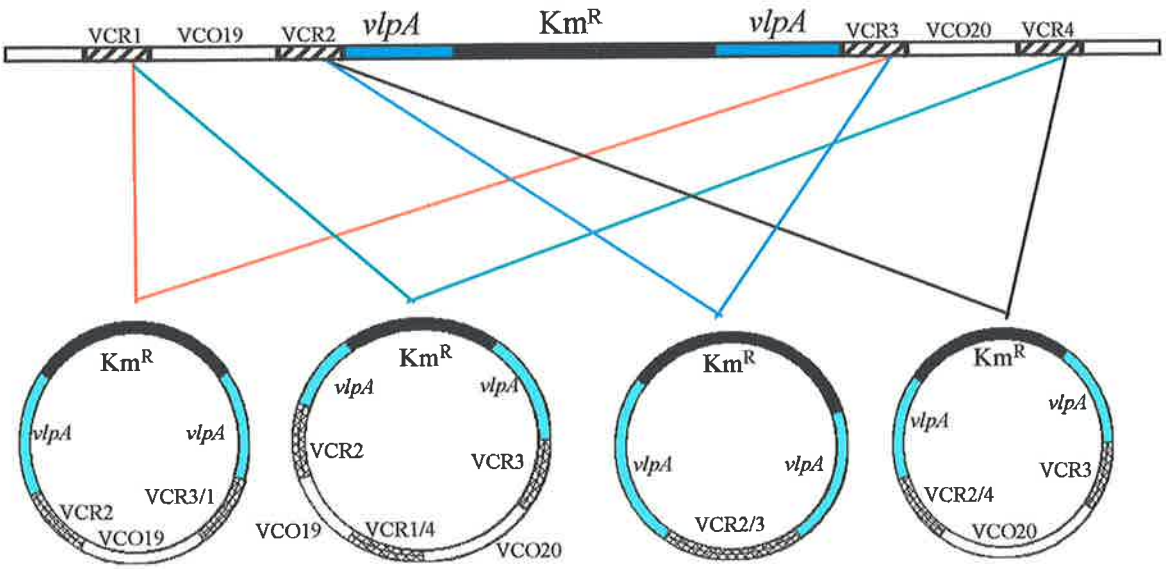

0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5-kb
#2816
#2817
#2816
#2817
#2816
#2817
vlpA
vlpA
vlpA
vlpA
vlpA
vlpA
0.6-kb
0.6-kb
2.3-kb

Figure 4.15 Model for the site-specific gene cassette deletion in *V. cholerae* chromosome.

The formation of circular cassettes resulting from integrase-mediated deletion from different VCRs. Site-specific recombination can take place at different VCRs present in the *V. cholerae* chromosomal integron, leading to the excision of different fragments within VCRs to form circular cassettes (not to scale).

Figure 4.15


VCR1
VCO19
VCR2
vlpA
KmR
vlpA
VCR3
VCO20
VCR4
KmR
vlpA
vlpA
VCR3/1
VCR2
VCO19
KmR
vlpA
vlpA
VCR2
VCR3
VCO19
VCR1/4
VCO20
KmR
vlpA
vlpA
VCR2/3
KmR
vlpA
vlpA
VCR3
VCR2/4
VCO20

strains tested in this study did not show any rearrangement of the *vlpA* cassettes in the chromosome. This suggests that the reinsertion of *vlpA* cassettes did not occur in these strains. However, future studies may need to analyze a number of the antibiotic resistance colonies which may have had rearrangement of the *vlpA*::Km^R or *vlpA*::Tet^R cassettes.

The lack of success in cassette insertion in this study could be explained in two ways; the existence and the amount of the circular cassettes. Free circular cassettes are not stable because they are not able to replicate by themselves, and the amount of the circular cassettes would be expected to be extremely low. They might be degraded during the experimental process before the occurring of cassette insertion event into the chromosome. In addition, the insertion of the circular unit requires only one site-specific recombination event for each gene insertion unit, whereas the deletion can occur at any recombination site, both *attI* and VCR sites. It is likely that deletion occurs more readily than insertion because of the proximity of VCR (or 59-be) sites within the integron. A free circular cassette needs to find itself in close proximity to a VCR to insert. This is supported by the finding of Collis and Hall (1992a) who demonstrated that the frequency of insertion was extremely low and deletion occurred ten times more frequently than cassette insertion.

The over-expression of integrase in *V. cholerae* may allow an increase in cassette reinsertion, that subsequently leads to the rearrangement of genes in the chromosomal integron. It is possible that the *vlpA* gene cassettes can undergo excision, with the excised cassettes integrating into different VCRs, or on the chromosomal integron of recipient *V. cholerae* cells or between *V. cholerae* strains. However, if the event occurred when the cell was dividing, the mobile *vlpA* gene cassettes can be incorporated into the integron more than once. Although the reinsertion of the *vlpA* gene cassettes to become associated with the *V. cholerae* integron area was not observed in this study, the deletion and

reinsertion of this gene is likely to explain the presence of *vlpA* as multiple copies in some strains of *V. cholerae* chromosome.

In conclusion, a number of gene cassettes including *vlpA* in this study, can be excised from *V. cholerae* integron catalyzed by integrase IntI1 or IntI4. A mobile gene cassette utilizes both the *attI* site and VCRs for movement through site-specific recombination. Hall (1998) indicated that integron and gene cassettes in many different Gram-negative bacteria have been very successful in achieving horizontal transfer. It seems reasonable to conclude that the excised circular gene model is one of a variety of mechanisms for the spread of mobile gene cassettes in bacterial populations.

Table 4.4 Bacterial strains used in this Chapter.

Bacterial strains	Characteristics	Source/Reference
E5191	DH5α with pSU2056	Dr. Christina Collis, CSIRO Division of Biomolecular Engineering, Sydney
E5227	DH5α with pPM5029	Mr. Christopher Clark, The University of Adelaide
E5242	DH5α with pSU2056 and pPM5021	This study (**Section 4.2.2.3**)
E5243	DH5α with pPM5022	This study (**Section 4.2.3.1.1**)
E5245	DH5α with pPM5029 and pPM5022	This study (**Section 4.2.3.3**)
E5246	DH5α with pSU2056 and pPM5022	This study (**Section 4.2.3.3**)
E5247	DH5α with pPM5023	This study (**Section 4.2.3.2**)
E5252	DH5α with pSU2056 and pPM5023	This study (**Section 4.2.3.3**)
E5260	DH5α with pPM5029 and pPM5023	This study (**Section 4.2.3.3**)
V911	O139 wildtype (AI-1837)	John Albert (I.C.D.D.R.B). Strain was from Bangladesh
V1269	V911 with *vlpA*1::Km^R	This study (Chapter 5)
V1271	V911 with *vlpA*4::Km^R	This study (Chapter 5)
V1279	V911 with *vlpA*1::Tet^R and *vlpA*4::Km^R	This study (Chapter 5)
V1285	V1271 with pSU2056	This study (**Section 4.2.3.5.1**)
V1286	V1271 with pPM5029	This study (**Section 4.2.3.5.2**)
V1287	V1285 with the deletion of *vlpA*4::Km^R mediated by IntI1	This study (**Section 4.2.3.5.1**)

Table 4.4 (continued)

Bacterial strains	Characteristics	Source/Reference
V1288	V1285 with the deletion of more than one cassette including *vlpA*4::Km^R mediated by IntI1	This study (**Section 4.2.3.5.1**)
V1289	V1285 with the deletion of more than one cassette including *vlpA*4::Km^R mediated by IntI1	This study (**Section 4.2.3.5.1**)
V1290	V1279 with pSU2056	This study (**Section 4.2.3.5.3**)
V1291-V1301	V1279 with the deletion of more than one cassette including *vlpA*4::Km^R mediated by IntI1	This study (**Section 4.2.3.5.3**)
V1305	V1279 with pPM5034	This study (**Section 4.2.3.5.4**)
V1308	V1305 with the deletion of more than one cassette including *vlpA*4::Km^R mediated by IntI4 from pPM5304	This study (**Section 4.2.3.5.4**)
V1309	V1305 with the deletion of *vlpA*4::Km^R mediated by IntI4 from pPM5304	This study (**Section 4.2.3.5.4**)

Table 4.5 Plasmids used in this Chapter

Plasmids	Characteristics	Source/Reference
pPM471	6.3-kb *Bam*HI fragment from *V. cholerae* 569B cloned into pBR322	Franzon and Manning (1986)
pPM5001	4.6-kb *Bam*HI fragment from *V. cholerae* 569B cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5002	1.0-kb *Eco*RI fragment from *V. cholerae* El Tor O17 cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5003	2.1-kb *Eco*RV fragment from *V. cholerae* O139 cloned into pBC-KS	This study (**Section 3.2.2**)
pPM5004	1.2-kb *Eco*RV fragment from *V. cholerae* O139 cloned into pBC-SK	This study (**Section 3.2.2**)
pPM5007	516-bp *vlpA* from PCR amplification using primers #1071 and #1072 cloned into pGEM-T.	This study (Chapter 3)
pPM5013	1.1-kb Km^R cartridge from pUWeKT inserted into *Hin*dIII site in the centre of *vlpA* in pPM5004.	This study (**Section 4.2.2.1**)
pPM5020	2.3-kb *Sph*I - *Sal*I fragment containing an N-terminal truncated *intI*4 with the adjacent *attI*4 site from *V. cholerae* 569B, and two copies of VCR in the plasmid vector pSU23	Mr. Christopher Clark (The University of Adelaide)
pPM5021	1.4-kb *Hind*III fragment from pPM5020 containing two VCRs and *attI*4 from *V. cholerae* 569B cloned into pSU23	This study (**Section 4.2.2.3**)
pPM5022	2.3-kb *Mlu*I fragment containing a full VCR, a truncated VCR and *vlpA*::Km^R from pPM5013 cloned into *Mlu*I site in pPM5020	This study (**Section 4.2.3.1**)

Table 4.5 (continued)

Plasmids	Characteristics	Source/Reference
pPM5023	2.2-kb *Eco*RV fragment containing *vlpA*::Km^R from pPM5013 cloned into *Bam*HI site in pRMH313	This study (**Section 4.2.3.2**)
pPM5026	pPM5023 with the deletion of two cassettes between *attI*1 and VCR mediated by IntI1	This study (**Section 4.2.3.4**)
pPM5027	pPM5023 with the deletion of *vlpA*::Km^R between two VCRs mediated by IntI1	This study (**Section 4.2.3.4**)
pPM5028	pPM5023 with the deletion of *vlpA*::Km^R between two VCRs mediated by IntI4	This study (**Section 4.2.3.4**)
pPM5029	963-bp fragment containing the *intI*4 gene from PCR amplification of *V. cholerae* 569B using primers #2596 and #2609 cloned into pBAD24	Mr. Christopher Clark (The University of Adelaide)
pPM5034	1.0-kb *intI*4 from PCR amplification of *V. cholerae* O139 chromosome using primers #3200 and #3201 and cloned into pGEM-T Easy.	This study (**Section 4.2.3.5.2**)
pBAD24	Ap^R, P_{BAD} promoter, cloning vector	Guzman *et al.*, 1995
pACYC184	Ap^R, *lacZ* promoter, cloning vector	Chang and Cohen (1972)
pRMH313	263-bp *Taq*I fragment included *attI*1 from Tn*21* cloned into pACYC184 at *Eco*RV and *Taq*I sites	Dr. Christina Collis (CSIRO Division of Biomolecular Engineering, Sydney)
pSU23	Cm^R *lacZ*, cloning vector	Bartolomé *et al.*, 1991
pSU2056	1.176-kb *Bam*HI-*Rsa*I fragment of In2 from Tn*21* inserted into pUC9 (Ap^R Int^+).	Martinez and de la Cruz (1990)
pUWEKT	Km^R cartridge originated from Tn903	Dr. Uwe Stroeher (The University of Adelaide).

Chapter 5

The role of VlpA in the pathogenesis of *Vibrio cholerae* O139

5.1 Introduction

The role of VlpA in the pathogenesis of *V. cholerae* is of considerable interest, as it represents the first bacterial lipocalin to be characterized. In earlier studies in this laboratory, Barker (1993) attempted to demonstrate a role for VlpA as a haemin binding protein by using the acid diazo Congo red, which correlates with haemin binding (Stugard *et al.*, 1989). By overproducing the protein in *E. coli* K-12 under the control of the T7 promoter/RNA polymerase system, it was shown that cell expressing VlpA can bind haemin as well as the related compounds hematoporphyrin IX and Congo red. While consistent with a role for VlpA in iron uptake, these data are not conclusive and do not reveal the function if any, of the VlpA protein in *V. cholerae* pathogenesis.

The TcpC lipoprotein has been postulated to provide a membrane anchor for the toxin-coregulated-pilus (TCP) (Parsot *et al.*, 1991; Ogierman and Manning, 1992; Iredell and Manning, 1994). For this reason, Barker (1993) speculated that VlpA might play a role as a membrane anchor for the MFRHA (Mannose-fucose resistant haemagglutinin) adhesin. A possible effect of VlpA on MFRHA activity was examined by comparing the haemagglutinating capacity of *E. coli* strains into which *mrhA,B* had been introduced in the presence or absence of *vlpA*. However, no differences in titre were observed, indicating that VlpA is not required as a membrane anchor in this system.

Therefore, further examination of the expression and function of *vlpA* was required to ascertain its role in *V. cholerae* infection. An earlier attempt to construct a *V. cholerae*

insertional mutant using a kanamycin resistance cartridge (*vlpA*::Km^R) in the O1 serogroup strain 569B was unsuccessful (Barker, 1993). In the present study, a pathogenic strain of the O139 serogroup (V911) was selected for mutational analysis, as this strain contains only two copies of *vlpA* (see Chapter 3). These two copies of *vlpA* were both inactivated individually and as a double mutant strain, both of which were then assessed for virulence in an infant mouse model of infection.

5.2 Results

5.2.1 Construction of suicide vectors carrying *vlpA*::Km^R and *vlpA*::Tet^R

SInce the V911 chromosome contains two copies of *vlpA* (Chapter 3), two antibiotic resistance cartridges (Km^R and Tet^R), were used sequentially to construct a *vlpA* double mutant. The approach used to construct the *vlpA* mutants was a modification of that described by Manning (1992) and is outlined in Figure 5.1. The 516-bp *vlpA* PCR product generated using the internal *vlpA* oligonucleotide primers (#1071 and #1072) was cloned into the pGEM-T (Promega) vector to produce pPM5007, and transformed into *E. coli* DH5α. This plasmid was subsequently digested with *Hin*dIII, which cleaved at approximately the middle of the *vlpA* ORF (nt 260), followed by insertion of either a Km^R or Tet^R cartridge at this site.

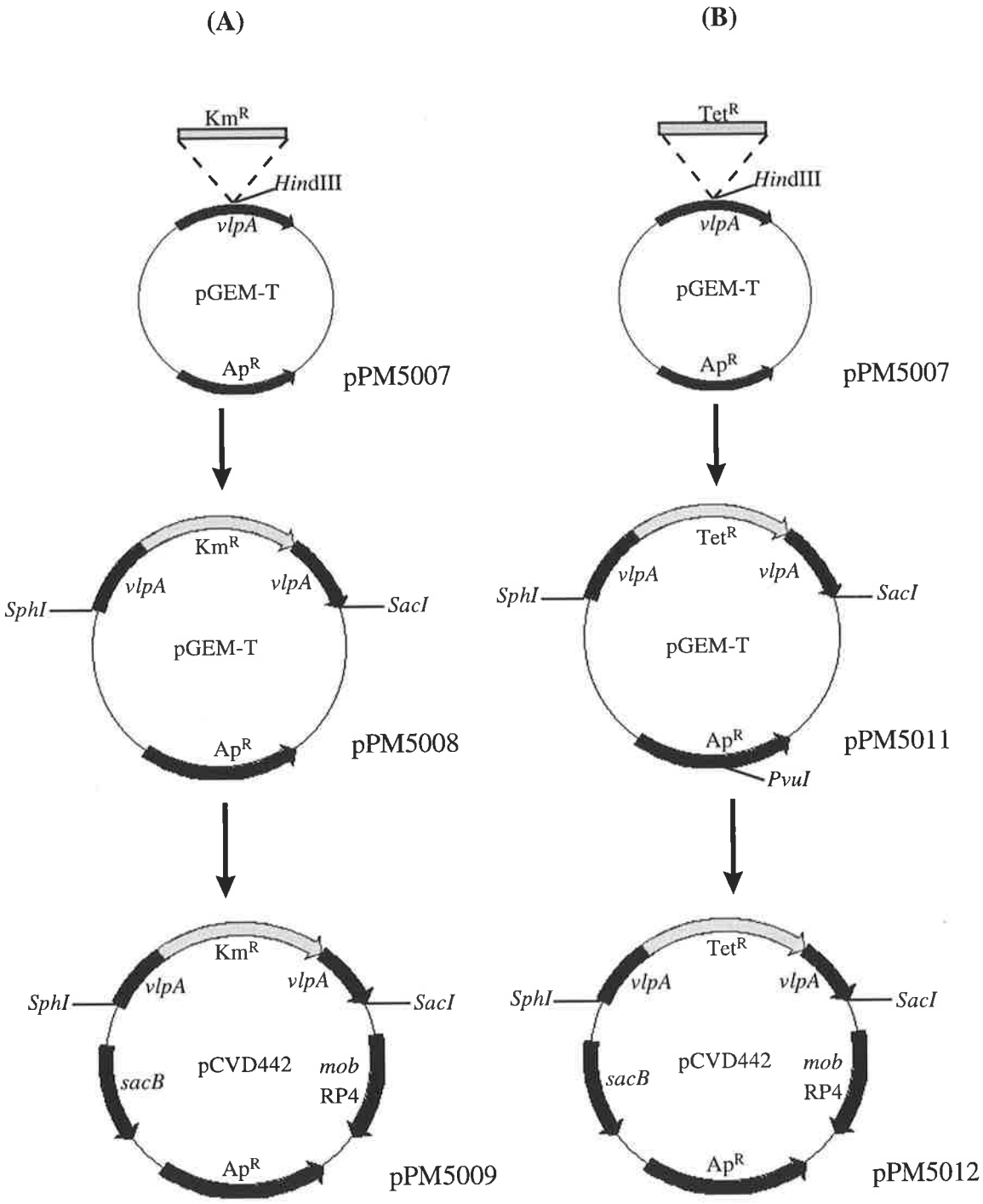
For the construction of *vlpA*::Km^R mutants, a 1.1-kb Km^R cartridge (originating from Tn*903*) was excised from the plasmid pUWEKT with *Hin*dIII, and inserted at the unique *Hin*dIII site in pPM5007 to create pPM5008 (Figure 5.1A). The ligation mixture was transformed into *E. coli* DH5α, with selection of Ap^R Km^R colonies. Restriction analysis and PCR amplification of plasmid isolated from these transformants allowed selection of plasmids in which the Km^R cartridge was inserted in the appropriate

Figure 5.1 Construction of pPM5009 and pPM5012 .

(A) A ca. 1.1 kb Km^R cartridge from pUWEKT was inserted at the unique *Hin*dIII site in pPM5007 to created pPM5008. This plasmid was digested with *Sph*I and *Sac*I, then a ca. 1.6 kb *Sph*I-*Sac*I fragment was inserted into pCVD442, resulting in pPM5009.

(B) A ca. 2.0 kb Tet^R cartridge from pBSA383 (Yansura and Henner, 1983) was inserted at the unique *Hin*dIII site of *vlpA* in pPM5007 to created pPM5011. This plasmid was digested with *Sph*I and *Sac*I, then a ca. 2.5 kb *Sph*I-*Sac*I fragment was inserted into pCVD442, resulting in pPM5012.

Figure 5.1


(A)
(B)
Km^R
Tet^R
HindIII
vlpA
pGEM-T
Ap^R
pPM5007
SphI
SacI
pPM5008
pPM5011
PvuI
pCVD442
sacB
mob
RP4
pPM5009
pPM5012

orientation, such that the transcription of the Km^R cartridge was in the same direction as the *vlpA* gene from the plasmid promoter.

The *vlpA*::Km^R gene was subcloned into the suicide plasmid vector, pCVD442 (Donnenberg and Kaper, 1991) for subsequent integration into the V911 genome via homologous recombination. pCVD442 contains the *mob*, *ori* and *bla* regions from pGP704 (Miller and Mekalanos, 1988). The R6K origin of replication (*oriR6K*) is dependent on the Pir (π) protein encoded by the *pir* gene (Miller and Mekalanos, 1988), and therefore maintenance of the suicide vector is dependent on the *pir* gene product, a copy of which is usually incorporated in the host chromosome to facilitate its maintenance in the donor strain. pCVD442 also contains the *sacB* gene from *Bacillus subtilis*. *sacB* expression in *E. coli* and other Gram-negative bacteria is lethal in the presence of sucrose at 30 °C. (The gene encodes levansucrase that converts sucrose to levan, which accumulates and is eventually lethal to the cell) (Gay *et al.*, 1985). Finally, pCVD442 can be efficiently mobilized into *V. cholerae* from an appropriate *E. coli* donor containing the necessary genes for conjugal transfer of the plasmid. These genes are usually supplied by a chromosomally integrated copy of the plasmid RP4 (Donnenberg and Kaper, 1991).

A 1.6-kb *Sph*I-*Sac*I fragment containing *vlpA*::Km^R from pPM5008 was transferred into pCVD442 at the *Sph*I and *Sac*I sites (Figure 5.1A). The resultant plasmid was transformed into *E. coli* S17-1λ*pir* (de Lorenzo *et al.*, 1990), selecting for Km^R colonies. The plasmid was confirmed to be carrying the *vlpA*::Km^R insert by restriction analysis and PCR amplification. This plasmid was designated pPM5009, and the *E. coli* S17-1λ*pir* carrying this plasmid was referred to as E2889.

A similar approach was used to construct a suicide vector carrying *vlpA*::Tet^R. A 2.0-kb Tet^R cartridge (originally from Tn*10*) was excised from pBSA383 (Yansura and

Henner, 1984) with *Hin*dIII and inserted into pPM5007 at the unique *Hin*dIII site to generate pPM5011 (Figure 5.1B). This plasmid was transformed into *E. coli* DH5α and Ap^R Tet^R colonies were obtained. The *vlpA*::Tet^R fragment was subcloned into pCVD442 by cutting pPM5011 with *Sph*I, *Sac*I and *Pvu*I. The last enzyme digests the residual vector to allow convenient isolation of the 2.5-kb *SphI*/*Sac*I fragment. This fragment was ligated into *SphI*/*Sac*I digested pCVD442 to produce pPM5012. This plasmid was confirmed by restriction analysis and PCR-amplification, then transformed into *E. coli* S17-1λ*pir*, yielding strain E5216.

5.2.2 Conjugation of pPM5009 and pPM5012 into *V. cholerae* O139.

5.2.2.1 Transfer of pPM5009 into *V. cholerae* O139 and selection for *vlpA*::Km^R mutants.

Plasmid pPM5009 was conjugated from E2889 into V911 using the methods described in **Section 2.10**. The bacterial mixture was plated onto minimal medium (to select against the donor) with kanamycin (to select against the recipient). Two exconjugants were selected and inoculated separately into LB without antibiotic selection, then incubated with shaking at 37 ° C overnight. These cultures were plated onto LA in the absence of NaCl and antibiotics, with 10% sucrose added. After overnight incubation at 30 ° C, one hundred and twenty-five colonies were patched onto nutrient agar plates with either kanamycin, or ampicillin added. Twenty-two of 125 colonies were Km^R Ap^S, and six of these were selected. The Km^R Ap^S phenotype was consistent with the absence of *sacB*, implying that these colonies might have arisen from double recombination events in the chromosome.

Chromosomal DNA was isolated and digested with *Eco*RV (this enzyme does not digest within *vlpA* or the Km^R cartridge), and analyzed by Southern hybridization. After transfer to nylon membrane, the digested DNA was incubated with a DIG-labeled *vlpA* probe (PCR product from oligonucleotides #1071 and #1072). Southern hybridization showed that allelic exchange had occurred between pPM5009 and one of the two chromosomal copies of *vlpA* (Figure 5.2A). These copies were referred to as *vlpA*1 (2.1-kb *Eco*RV band) and *vlpA*4 (1.2-kb *Eco*RV band) (Chapter 3 Figures 3.5C, D). The recombination event in one of the *vlpA* genes resulted in an increase in size of 1.1-kb to 3.2 and 2.3-kb, respectively (due to the 1.1-kb Km^R cartridge insertion).

DNA isolated from five of the six selected Km^RAp^S colonies showed a 3.2-kb fragment hybridizing with the *vlpA* probe. Therefore, recombination had occurred in the 2.1-kb *Eco*RV fragment resulting in a mutation in *vlpA*1 (lanes 1, 2, 4, 5 and 6 in Figure 5.2A). The sixth mutant carried the Km^R cartridge in *vlpA*4 (1.2-kb *Eco*RV fragment), with a 2.3-kb fragment identified by the *vlpA* probe (lane 3). Hybridization with a DIG end-labeled Km^R cartridge (*Eco*RI fragment) probe was used to confirm the location of the Km^R cartridge in the chromosome of these six mutants (Figure 5.2B). As expected, the hybridization showed a 2.3-kb band in lane 1 which is a strain with the *vlpA*4 mutation, and 3.2-kb bands in lanes 2-6 that are the strains with mutations in *vlpA*1. No hybridization was observed with wild-type DNA (lane 7).

PCR-amplification was performed for all 22 putative exconjugants (data not shown) using oligonucleotide primers generated from the *vlpA*4 flanking genes (#3141 and #3142) (Figure 5.2C). Only one in 22 colonies showed a PCR product, indicative of an insertionally-inactivated *vlpA*4. Therefore the oligonucleotide primers #1071 and #2905, generated from *vlpA*1 flanking genes (Figure 5.2C) were used to determine whether

Figure 5.2 Southern hybridization to confirm the insertion of Km^R and Tet^R into the *vlpA* genes in the *V. cholerae* O139 chromosome.

(A) Chromosomal DNA from *V. cholerae* O139 putative *vlpA* mutants (*vlpA*::Km^R) and V911 (wild-type) were digested with *Eco*RV, then probed with a DIG-labeled *vlpA*-specific PCR probe. Lanes **1-6** are putative *vlpA* mutants, lane **7** is V911 (control). Lanes **1, 2, 4, 5** and **6** show homologous recombination of the *vlpA*::Km^R mutant at the *vlpA*1 locus (V1269), while lane **3** shows the *vlpA*::Km^R mutation at the *vlpA*4 locus (V1271).

(B) Chromosomal DNA from V911 and putative *vlpA* mutants (*vlpA*::Km^R) was digested with *Eco*RV, then probed with a DIG end-labeled kanamycin probe. Lane **1** is a putative *vlpA*4 mutant, lanes **2-6** are putative *vlpA*1 mutants, lane **7** is the V911 wild-type control.

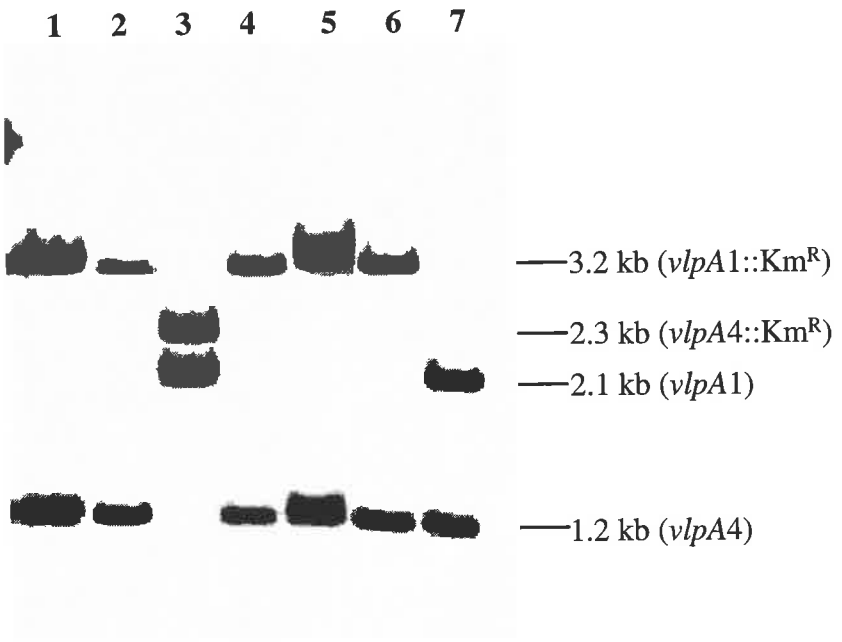
(C) Schematic to show restriction enzyme sites in *vlpA*1::Tet^R and *vlpA*4::Tet^R. (**B** = *Bam*HI, **C** = *Sac*I, **H** = *Hin*dIII, **M** = *Sma*I)

(D) Chromosomal DNA from putative single *vlpA* mutants (*vlpA*::Tet^R, *vlpA*::Km^R) and V911 were digested with *Eco*RV, then probed with DIG-labeled *vlpA*-specific PCR probe. Lanes **1-2** are putative mutants, which show replacement of the *vlpA*1 allele by *vlpA*::Tet^R. The doublet 2.1-kb bands are indicated by asterisks due to the presence of an *Eco*RV enzyme restriction site within the Tet^R sequence. One of these strains was retained and designated V1280. Lanes **3, 4** and **5** represent *vlpA*1::Km^R (V1269), *vlpA*4::Km^R (V1271) mutants, and V911 respectively.

The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.). The approximate sizes of fragments is indicated in kb.

Figure 5.2

(A)


1 2 3 4 5 6 7
—3.2 kb (vlpA1::KmR)
—2.3 kb (vlpA4::KmR)
—2.1 kb (vlpA1)
—1.2 kb (vlpA4)

(B)

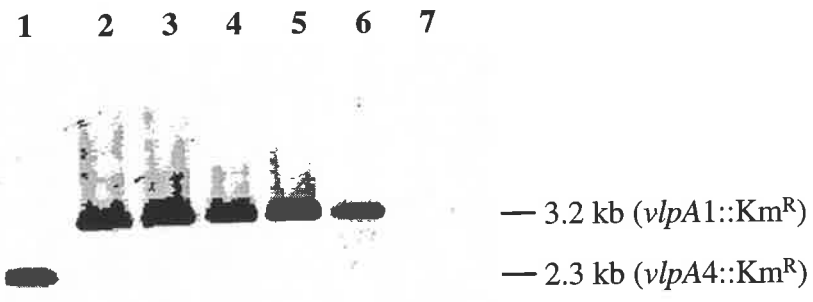
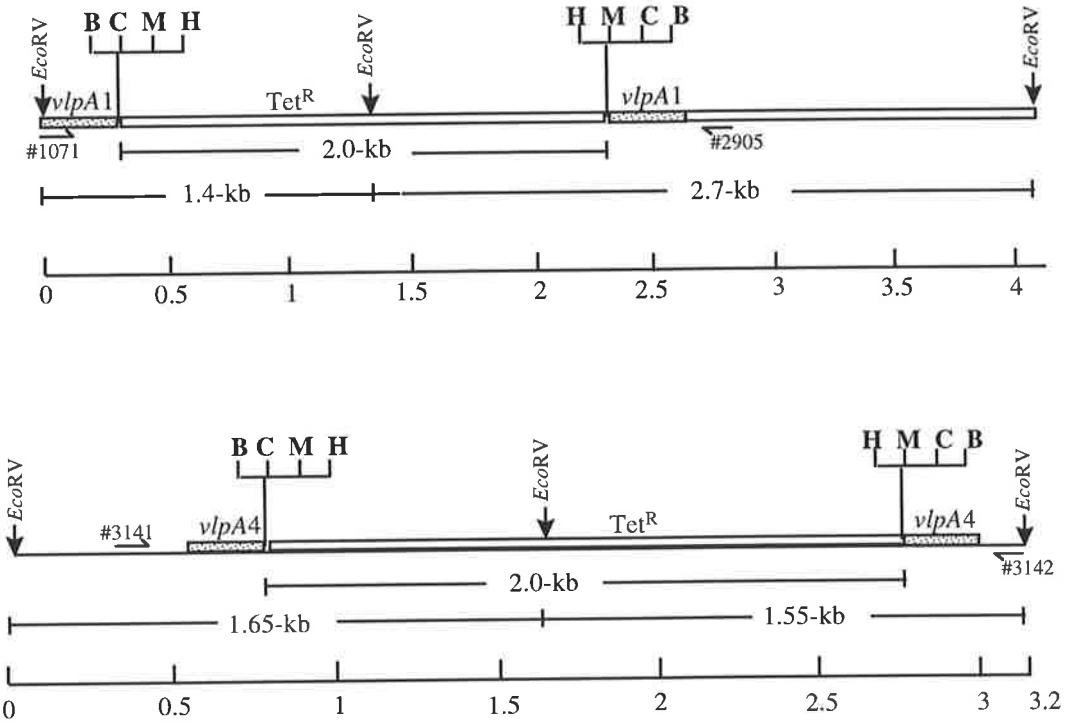
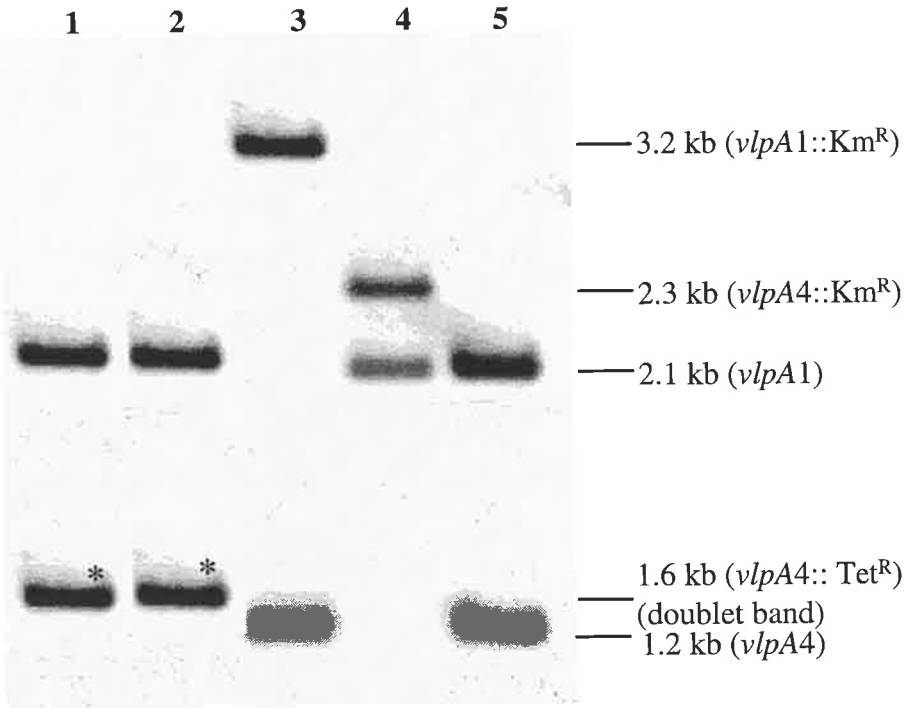

1 2 3 4 5 6 7
— 3.2 kb (vlpA1::KmR)
— 2.3 kb (vlpA4::KmR)

Figure 5.2

(C)


B C M H
EcoRV
vlpA1
Tet^R
EcoRV
H M C B
vlpA1
EcoRV
#1071
2.0-kb
#2905
1.4-kb
2.7-kb
0
0.5
1
1.5
2
2.5
3
3.5
4
B C M H
EcoRV
#3141
vlpA4
EcoRV
Tet^R
H M C B
vlpA4
EcoRV
#3142
2.0-kb
1.65-kb
1.55-kb
0
0.5
1
1.5
2
2.5
3
3.2

(D)


1
2
3
4
5
3.2 kb (vlpA1::Km^R)
2.3 kb (vlpA4::Km^R)
2.1 kb (vlpA1)
1.6 kb (vlpA4:: Tet^R)
(doublet band)
1.2 kb (vlpA4)

recombination with the insertion mutant copy of *vlpA* had occurred at *vlpA*1 in the remaining 21 exconjugants. The size of the PCR products obtained was consistent with insertion of the Km^R cartridge in *vlpA*1 in all of these strains (data not shown). One of the *V. cholerae* O139 strains with Km^R insertion mutations at either *vlpA*1 or *vlpA*4 were designated V1269 and V1271, respectively.

5.2.2.2 Transfer of pPM5012 into *V. cholerae* O139 and selection for *vlpA*::Tet^R mutants.

Conjugation between E5216 (pPM5012) and V911 was performed, and exconjugants were selected on minimal medium with tetracycline. A similar approach to **Section 5.2.2.1** was used for isolation and analysis of *vlpA*::Tet^R mutants. Two exconjugants were selected and grown separately in LB without antibiotic selection at 37°C overnight, then plated onto LA in the absence of NaCl and antibiotics, but in the presence of 10% sucrose. After overnight incubation at 30°C, seventy-two of the exconjugants obtained were patched onto NA plates with either tetracycline or ampicillin. Only two from 72 colonies screened were Tet^RAp^S, and chromosomal DNA was prepared from both.

Southern hybridization analysis was performed using the DIG-labeled *vlpA*-specific PCR probe. If homologous recombination occurs between the chromosomal *vlpA* copy and the insertional inactivation copy, the *Eco*RV fragment recognized by the probe should increase in size from 2.1 to 4.1-kb (*vlpA*1), or from 1.2 to 3.2-kb (*vlpA* 4) i.e. a 2.0-kb increase in size due to the Tet^R cartridge. However, neither band showed this expected size increase. This is because it was later realized that *Eco*RV cuts within the Tet^R cartridge (nt 942), and therefore replacement of the wild-type copy with the Tet^R inactivated copy would result in a 1.4 and 2.7-kb band (*vlpA*1::Tet^R) (see diagram in Figure 5.2C), as well as the

wild-type copy of the remaining gene of 1.2-kb (*vlpA*4). Alternatively, interruption of *vlpA*4 would result in 2.1-kb (*vlpA*1) band of the wild-type gene and a 1.6-kb doublet band (*vlpA*4::Tet^R). Southern hybridization (Figure 5.2D) showed that allelic exchange between pPM5012 and the *V. cholerae* O139 chromosome had occurred at *vlpA*4 in both cases. One colony was picked, and this strain was designated V1280. The enzymes *Bam*HI, *Eco*RI, *Hin*dIII were also used to digest chromosomal DNA from these colonies. Unfortunately, use of these enzymes make it more difficult to identify the hybridization bands (data not shown). The fragments resulting from *Bam*HI or *Eco*RI digestion of V911 DNA are larger than the markers used, and these enzymes also digest at the polylinker of the Tet^R cartridge. *Hin*dIII digests within *vlpA* and is therefore also unsuitable for identifying the hybridizing bands.

5.2.2.3 Transfer of pPM5012 into V1269 or V1271, pPM5009 into V1280, and selection for *vlpA* double mutants.

V1269 (*vlpA*1::Km^R) and V1271 (*vlpA*4::Km^R) were separately conjugated with E5216 (pPM5012) in an attempt to generate a strain with mutations in both copies of *vlpA*. Exconjugants were selected on minimal medium with kanamycin and tetracycline added. Colonies were grown separately at 37°C shaking overnight in Luria broth (LB), and plated onto Luria agar (LA) in the absence of NaCl and antibiotics, with 10% sucrose. After overnight incubation at 30°C, seventy-two exconjugants were obtained from conjugation between the donor E5216 and recipient V1271. Colonies were patched onto nutrient agar plates with either tetracycline, kanamycin, or ampicillin to screen for a Tet^{R}Km^{R}Ap^S phenotype. Only two of the 72 colonies screened had the desired phenotype. These presumably resulted from the introduction of *vlpA*::Tet^R into the chromosome via

homologous recombination, replacing the wild-type copy of *vlpA*1. No colonies with the desired phenotype were recovered from conjugation between E5216 and V1269. In addition, V1280 (*vlpA*1::Tet^R) was also used as a recipient for conjugation with E2889 (pPM5009) in the same manner as described above, and screened for Tet^{R}Km^{R}Ap^S colonies. No colonies of the desired phenotype were observed.

Southern hybridization was performed using the DIG-labeled *vlpA*-specific PCR probe for verification of these two colonies (exconjugants) containing mutations in both *vlpA* genes. Chromosomal DNA from these colonies, together with V1269, V1271 and V911 (wild-type) was digested with *Eco*RV. Southern hybridization showed that allelic exchange between pPM5012 and V1271 had indeed occurred in *vlpA*1 (2.1-kb *Eco*RV) (Figure 5.3A, lanes 1 and 2). Since *Eco*RV cuts within the Tet^R cartridge, two bands of 1.4 and 2.7-kb were observed as predicted (**Section 5.2.2.2**).

The PCR products amplified using oligonucleotide primers #1071 and #1072 (which bind at both ends of *vlpA* and read internally) confirmed that Tet^R and Km^R cartridges were recombined into *vlpA*1 and *vlpA*4 respectively (Figure 5.3B). The 2.5-kb (*vlpA*1::Tet^R) and 1.6-kb (*vlpA*4::Km^R) bands were detected in both putative *vlpA* double mutants (Lanes 1 and 2). A 516-bp band observed with the single *vlpA* mutants (Lanes 3, 4 and 5), corresponding to a wild-type copy of the gene was not apparent in the double mutants. One of the double mutants (i.e. No. 1) was designated V1279, and retained for later studies.

5.2.3 Characterization of *vlpA* single and double mutants

Before proceeding to a study of the role of VlpA in pathogenesis, it was necessary to characterize the *V. cholerae* O139 single and double *vlpA* mutants to ensure that the *vlpA*

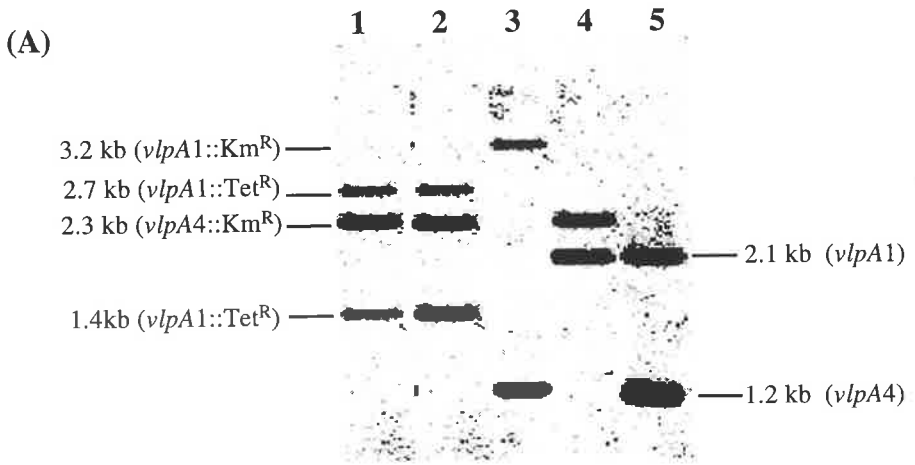
Figure 5.3 Southern hybridization and PCR amplification to confirm the double mutation of *vlpA*1 and *vlpA*4 in V911.

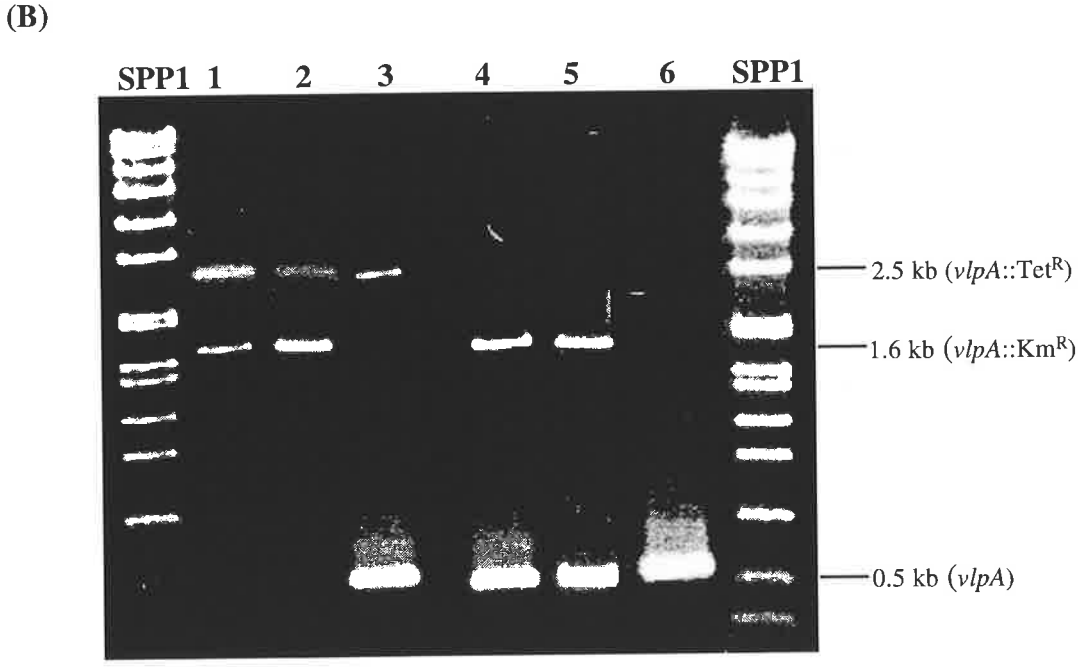
(A) Chromosomal DNA from the putative double *vlpA* mutants and the parent strain V911 was digested with *Eco*RV, then probed with a PCR DIG-labeled *vlpA* specific probe. Lanes contain: **1,** *vlpA*1::Tet^R and *vlpA*4::Km^R No.1; **2,** *vlpA*1::Tet^R and *vlpA*4::Km^R No.2; **3,** *vlpA*1:: Km^R (V1269); **4,** *vlpA*4::Km^R (V1271); and **5,** V911 (control).

(B) PCR amplification using primers #1071 and #1072 which bind at both ends of *vlpA*. Lanes contain : **1,** *vlpA*1::Tet^R and *vlpA*4::Km^R No.1; **2,** *vlpA*1::Tet^R and *vlpA*4::Km^R No.2; **3,** *vlpA*1::Tet^R (V1280); **4,** *vlpA*1::Km^R (V1269); **5,** *vlpA*4::Km^R (V1271); and **6,** V911 (control).

The molecular size marker (SPP1) is bacteriophage SPP1 DNA digested with *Eco*RI.). The approximate sizes of fragments is indicated in kb.

Figure 5.3


(A)
1
2
3
4
5
3.2 kb (vlpA1::KmR)
2.7 kb (vlpA1::TetR)
2.3 kb (vlpA4::KmR)
1.4kb (vlpA1::TetR)
2.1 kb (vlpA1)
1.2 kb (vlpA4)


(B)
SPP1
1
2
3
4
5
6
SPP1
2.5 kb (vlpA::TetR)
1.6 kb (vlpA::KmR)
0.5 kb (vlpA)

mutations did not affect growth rate or production of major virulence determinants. Growth rates were compared by determining generation time in nutrient broth (NB) with the appropriate antibiotic added. The strains V1269, V1271, V1279 and V911 were each grown in duplicate from single colonies (four hours at 37 ° C shaking), to a concentration of approximately 3 x 10^8 organisms per ml. At this stage all cultures were diluted 1 in 10^4 into pre-warmed NB and viable counts were performed. After a second incubation period of four hours at 37°C, viable counts were performed again and used to calculate generation times. The mean generation times of V911, V1269, V1271 and V1279 were not significantly different (Table 5.1).

Studies in the infant mouse cholera model have shown that TCP is a critical virulence determinant for both O1 and O139 *V. cholerae* (Taylor *et al.*, 1987; Attridge *et al.*, 1996; Voss *et al.*, 1996; Sharma *et al.*, 1997a). Therefore, it was necessary to check that expression of TCP was not affected in the *vlpA* single and double mutants. V911, V1269, V1271 and V1279 were grown in AKI medium for assessment of TCP expression (Voss and Attridge, 1993). Immunoblotting of whole cell lysates (rabbit polyclonal anti-TCP was provided by Dr. S. Attridge, The University of Adelaide, see Chapter 2 **Section 2.20.2**) confirmed that production of the TcpA pilin subunit was not decreased in any of the *vlpA* mutant strains (Figure 5.4).

5.2.4 *In vivo* behavior of *vlpA* mutants

5.2.4.1 Virulence studies

The infant mouse cholera model (Attridge and Rowley, 1983) was used to assess the virulence of the parent strain V911, and both single and double *vlpA* mutants. LD_{50} values were determined for each strain by feeding graded doses to groups of mice and

Table 5.1 Generation times of wild-type and mutants in NB.

Strains	Generation times (minutes)		
	set 1	set 2	mean
V911	23.9	23.9	23.9
V1269	24.7	24.4	24.6
V1271	24.6	23.9	23.8
V1279	22.4	24.2	23.3

Figure 5.4 Western immunoblot analysis of *V. cholerae* O139 and *vlpA* mutants using anti-TCP antibody.

Whole cell lysates (equivalent to 1 x 10^8 cells) of *V. cholerae* O139 wild-type and *vlpA* mutants were electrophoresed on a 15% SDS polyacrylamide gel and transferred to a nitrocellulose membrane. Proteins were detected using (absorbed) polyclonal anti-TCP (1 : 40,000 dilution). The protein band in each lane has a molecular mass of approximately 20 kDa. Lanes contain : **1**, V911; **2**, V1279; **3**, V1269; **4**, V1271; **5**, V912 *tcpA*::Km^R (negative control); and **6**, V912 wild-type (positive control).

Figure 5.4

following survival for 48 hours. No differences were detected in the survival rates of the mice; indeed, all mice died even at the lowest doses used (≈200 organisms per mouse). Therefore, the LD_{50} was <200 bacteria for both wild-type and mutants. A repeated experiment gave the same result confirming that VlpA did not significantly affect virulence, at least in this model.

A second experiment provided independent confirmation of this finding. V1306 and V1307 (see **Section 4.2.3.5.1** for construction of these mutants) are two strains derived from V911 both of which lack functional copies of *vlpA*. These strains carry large deletions spanning *vlpA* as opposed to the insertionally-inactivated copies. In addition, the closely linked *mrhA* and *mrhB* genes downstream of *vlpA*4 have also been deleted. When titrated for virulence in parallel with wild-type, neither V1306 nor V1307 showed any evidence of attenuation; the LD_{50} value for all three strains is <200 bacteria.

5.2.4.2 Competition experiments

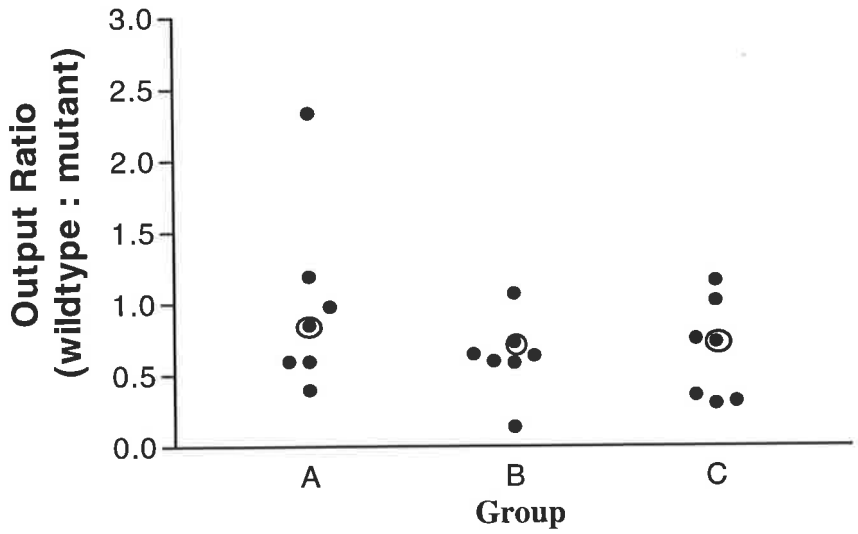
Competition experiments were also performed with the single and double *vlpA* mutants, to determine whether these could compete with wild-type strains *in vivo*. Three mixed bacterial suspensions were prepared, comprising wild-type and mutant bacteria (V1269, V1271 or V1279) in approximately equal concentrations. Each suspension was fed to a group of seven infant mice, each animal receiving ≈ 10^3 bacteria (≈ 10 LD_{50} doses) of each strain. Viable counts were performed retrospectively and confirmed the input ratios of Km^S:Km^R bacteria in these mixtures to be approximately 1 : 1 (see Figure 5.5). The mice were held at 25°C for 24 hours, at which time the animals were sacrificed and the small intestines excised, and homogenized in PBS. Dilutions of each gut homogenate were plated onto both NA and agar containing kanamycin. This allowed the separate enumeration of

Figure 5.5 Comparison of the *in vivo* persistence of the *vlpA* mutants and wildtype parents.

Data show the results of competition experiments performed between the V911 parent strain and each of three *vlpA* mutants (V1269, V1271, V1279). Each dot represents the ratio of parent : mutant bacteria recovered from an individual mouse 24 hours after feeding; circles show the median values. Input ratios and median output ratios were as follows:

	Input ratio	**Median output ratio**
Group A; wild-type : V1269	0.85 : 1	1.02 : 1
Group B; wild-type : V1271	0.64 : 1	0.78 : 1
Group C; wild-type : V1279	0.73 : 1	0.78 : 1

Figure 5.5



wild-type and mutant bacteria, so that an output ratio of the two strains could be determined for each animal. For each competition experiment the median output ratios were similar to the input ratios (Figure 5.5). In this model, none of the *vlpA* mutants have any defect in colonization or persistence.

5.3 Discussion

A previous study (**Section 3.1**) had shown that V911 has only two copies of *vlpA*, and this strain was therefore selected for the construction of a mutant(s) with which to evaluate the pathogenic significance of VlpA. It was decided to sequentially inactivate both copies of the gene by homologous recombination of insertionally-inactivated gene copies. Interestingly, allelic exchange between pPM5009 (*vlpA*::Km^R) and V911 chromosomal copies of this gene was observed at a higher frequency with *vlpA*1 than *vlpA*4. PCR analysis in **Section 5.2.2.1** showed that 21 of the 22 Km^R colonies were *vlpA*1::Km^R. A second plasmid pPM5012 carrying *vlpA*::Tet^R was also constructed and used for conjugation with V911, in order to select for recombination of an insertionally-inactivated copy of *vlpA* at the second allele. However, the mutant (Tet^R) copies of *vlpA* were only observed at *vlpA*1 (V1280). Although selection for double mutants was performed using V1269 (*vlpA*1::Km^R), V1271 (*vlpA*4::Km^R) and V1280 (*vlpA*1::Tet^R), the double mutant was only obtained from V1271. In contrast, cassette deletion of double mutants (V1279 with *vlpA*1::Tet^R, *vlpA*4::Km^R) in Chapter 4 (**Section 4.2.3.5.1**), showed that the event occurred in *vlpA*4::Km^R at a much higher frequency than *vlpA*1::Tet^R. Of the twelve colonies with a deleted copy of *vlpA* obtained, eleven were Km^S, and only one was Tet^S. It seems likely that homologous recombination occurred more readily in *vlpA*1 than *vlpA*4,

but cassette deletion, which involves site-specific recombination favoured the *vlpA*4 locus over that of *vlpA*1.

Before the effects of the *vlpA* mutations on *V. cholerae* virulence were examined, both single and double mutants (V1269, V1271, V1279) were compared with their parent strain (V911) for properties which could potentially obscure results obtained from the infant mouse model. Initially, the growth rates were characterized by determining the generation time. None of the mutants differed from the wild-type (Table 5.1). Since TCP is a critical virulence determinant for *V. cholerae* strains of both the O1 and O139 serogroup, the mutant strains were also compared with the wild-type for expression of these pili. When bacteria were grown under conditions shown to support TCP expression, it was confirmed that VlpA had no effect on the production of TCP in both wild-type and mutant strains.

The infant mouse model for accessing the virulence of *V. cholerae* strains revealed that *vlpA* was not a critical virulence factor. Moreover, the competition experiments confirmed that both single and double *vlpA* mutants have the ability to compete with wild-type *in vivo*. There was no defect in the colonization or persistence of the mutants compared to wild-type in the small intestine of mice. In addition, *V. cholerae* mutants generated from the deletion of both *vlpA* copies in the strains V1306 and V1307 (Chapter 4, **Section 4.2.3.5.1**) displayed identical properties. Additional genes flanking *vlpA* were also deleted in these strains, including the closely linked *mrhA, B* genes, which have previously been implicated in virulence (Franzon *et al.*, 1993). However, in our case, no reduction in virulence was detected in the infant mouse model. No differences were detected in the survival rates of the mice, and all of the mice died at the lowest doses used (≈200 organisms per mouse).

The effect of *mrhA* on *V. cholerae* virulence in this study is contradictory to the results of Franzon *et al.* (1993), where the *mrhA* mutant from *V. cholerae* O1 stain 569B was shown to be an important virulence factor in the same infant mouse model. The mutant strain showed a marked defect in its ability to persist in the infant mouse gut and was incapable of competing with the wild-type organism. However, these studies were performed in different *V. cholerae* strains with very different levels of virulence, and capacity to colonize. Franzon *et al.* (1993) used *V. cholerae* O1 classical strain 569B where the wild-type strain exhibited an LD_{50} of 9×10^5 organisms per ml after 48 hours, whereas the LD_{50} of V911 in this study is less than 2×10^2. It is possible that 569B is a weakly virulent strain, and loss of the *mrhA* gene may be sufficient to affect its virulence, while in V911 other virulence determinants may mask its effect. It would therefore be interesting to carry out further studies comparing the contribution of *vlpA* to virulence in strain 569B to that of *V. cholerae* O139 in this study.

The function of VlpA, a member of the lipocalin family, is still unknown. The lipocalins are a structurally and functionally diverse family of proteins, with several functions were reported in eukaryotes. Clinical studies have shown the practical importance of many lipocalins as biochemical markers in health and disease (Flower, 1996). The lipocalins are best known for their binding of a remarkable array of small hydrophobic ligands. Based on the homology between VlpA and human ApoD, which functions to bind a hydrophobic ligand, it seems likely that VlpA also functions as a hydrophobic ligand-binding protein.

Barker and Manning (1997) constructed *E. coli* strains capable of overproducing VlpA, and suggested that VlpA is able to bind haemin, although at low affinity. The activity of *vlpA* as a haemin-binding protein suggested that VlpA may be involved in iron

uptake. However, a 26 kDa cytoplasmic membrane protein (HutB) and a 77 kDa outer membrane protein (HutA) from *V. cholerae*, that confer the ability to utilize iron from haemin and hemoglobin, were shown to be clearly distinct from VlpA protein (Henderson and Payne, 1993). Conversely, it is possible that there may be more than one system for haemin uptake in *V. cholerae*.

Since VlpA is a bacterial lipocalin, it was predicted that VlpA, similar to *E. coli* bacterial lipocalin (Blc), might serve a function that contributes to the adaptation of cells to starvation conditions in bacteria (Bishop *et al.*, 1995). It is possible that environmental stress may induce the response of *vlpA*. Although Bishop *et al.* (1995) suggested that the bacterial lipocalins may function in adaptation to starvation conditions, it has not been demonstrated in this study due to the unclear nature of gene expression of *vlpA* in *V. cholerae* at present.

Collis and Hall (1995) revealed that the definitive feature of integrons was the site-specific recombination system which was able to capture a vast array of different resistance genes. Therefore, a comparison between the *vlpA* sequence and a number of antibiotic resistance gene sequences which are commonly found in Gram-negative bacteria was performed, i.e. GenBank accession number U73849 (Tetracycline), AF100174 (Kanamycin), AF100174 (Chloramphenicol), AF100177 (Gentamycin), AF202976 (Streptomycin), AF244574 (Spectinomycin), and AJ272109 (beta-lactam). However, no significant similarity was found between these sequences, suggesting that VlpA is not related to these antibiotic resistance genes. It is possible that VlpA may have no important function, like several other cassettes in the *V. cholerae* chromosomal integron (Mazel *et al.*, 1998; Rowe-Magnus *et al.*, 1999). Gene cassettes have been found to be capable of being integrated individually into integrons independent of a specific functional basis (Hall

and Collis, 1998). Although VlpA function is still undefined, the *vlpA* gene cassette has been successful in the integration of a varied number of copies into a host chromosomal integron. Further analysis may provide a major contributory function of VlpA to bacterial cells.

Table 5.2 Plasmids used in this chapter.

Plasmid	Characteristics	Source/Reference
pPM5007	516-bp *vlpA* from PCR amplification using primers #1071 and #1072 cloned into pGEM-T	This study
pPM5008	1.1-kb *Hin*dIII fragment containing Km^R cartridge from pUWEKT cloned into the unique *Hin*dIII site of pPM5007	This study
pPM5009	a 1.6-kb *Sph*I-*Sac*I from pPM5008 cloned into pCVD442	This study
pPM5011	pPM5007 with a 2.0-kb *Hin*dIII fragment containing Tet^R cartridge from pBSA383 inserted at the unique *Hin*dIII site	This study
pPM5012	2.5-kb *Sph*I and *Sac*I fragment containing *vlpA*::Tet^R from pPM5011 cloned into pCVD442	This study
pBSA383	cloning vector	Yansura and Henner, 1983
pCACTUS	cloning vector, *sacB*, Cm^R	Christopher Clark (The University of Adelaide)
pCVD442	Suicide vector with Pir-dependent RK6 replicon. Contain *sacB* gene for positive selection of recombinants (Ap^R).	Donnenberg and Kaper (1991)
pGEM-T	cloning vector, pGEM5Zf+ linearized with T- overhang for direct cloning of PCR products (Ap^R *lacZ*) .	Promega
pUWEKT	Km^R cartridge originated from Tn*903*	Dr Uwe Stroeher (The University of Adelaide).

Table 5.3 Bacterial strains used in this chapter.

Bacterial strains	Characteristics	Source/Reference
E2889	*E. coli* S17-1λ pir with pPM5009	This study (**Section 5.2.1**)
E5216	*E. coli* S17-1λ pir with pPM5012	This study (**Section 5.2.1**)
V911	*V. cholerae* O139 wildtype	Laboratory strain
V912	*V. cholerae* O139 wildtype	Laboratory strain
V1269	*V. cholerae* O139 with *vlpA*1::Km^R	This study (**Section 5.2.2.1**)
V1271	*V. cholerae* O139 with *vlpA*4::Km^R	This study (**Section 5.2.2.1**)
V1279	*V. cholerae* O139 double *vlpA* mutant with *vlpA*1::Tet^R and *vlpA*4::Km^R	This study (**Section 5.2.2.3**)
V1280	*V. cholerae* O139 with *vlpA*1::Tet^R	This study (**Section 5.2.2.2**)

Chapter 6

Transcriptional analysis of *vlpA*

6.1 Introduction

The *vlpA* gene has been over-expressed in *E. coli* by Barker and Manning (1997), who suggested that one of the functions of VlpA is to bind haemin, a human iron-binding protein (Otto *et al.*, 1992). Congo red binding of the *E. coli* strain expressing this gene was readily apparent, indicating that the induced protein may bind porphyrin or haemin. Gene expression studies were subsequently performed in *V. cholerae* 569B and O17 under iron-limiting conditions (using the iron chelator 2, 2′ dipyridyl) in order to detect *vlpA*-specific transcription. However, no *vlpA*-specific transcripts were detected under these conditions (Barker, 1993).

Although an analysis of the *vlpA* gene sequence shows the first in-frame initiation codon to lie within the cassette boundary, further analysis of the *vlpA* cassette has not confirmed any possible promoter sequences. Barker (1993) identified the promoter region of the *mrhA* and *mrhB* genes using primer extension assays and showed the length of the extension products corresponded to a promoter located immediately upstream of these genes. Furthermore, Northern blot analysis indicated that the promoters and terminators for *mrhA*, *mrhB* and the ORF upstream of these genes (VCO3.1 or ORF3.1), are within the VCRs, with no products further upstream of VCO3.1.

Barker *et al.* (1994) have suggested that VCRs might play a role in transcription termination in *V. cholerae*. RNase protection with VCR-specific probes of either orientation was demonstrated, and multiple bands consistent with the presence of full-length copies of VCR within *V. cholerae* total RNA were observed. Similarly, Recchia and

Hall (1995) have reported that the majority of integron transcriptional units are truncated at defined lengths of the full length transcripts corresponding to individual gene cassettes. Collis and Hall (1995) suggested that 59-be may not only function as recombination sites, but also as transcriptional terminators or processing signals for endonucleolytic cleavage of transcripts.

Many known gene cassettes do not include promoter signals (Recchia and Hall, 1995). Collis and Hall (1995) examined the expression of resistance genes encoded in integron-associated gene cassettes, and reported that all transcripts detected commenced at the common promoter P_{ant}, which is the only integron promoter region with strong expression. Alterations in the sequence of P_{ant} affect the level of resistance expressed by cassette genes. When the gene cassette is closest to P_{ant}, the level of expression is highest, but it is reduced if the cassette is situated downstream of one or more cassettes (Collis and Hall, 1995). Levesque *et al.* (1994) reported that a strong promoter found in *V. cholerae* is located upstream from the *attI* site, and directs the transcription of the cassettes. However, some cassettes carry their own promoter, and could allow the expression of other downstream cassettes (Hall and Collis, 1998). As *vlpA* is a cassette gene in the *V. cholerae* chromosomal integron, it was of interest to determine the location of the *vlpA* promoter if it exists, and how this gene is expressed. It may be possible that the expression of the *vlpA* gene is dependent on integration of the cassette in the correct orientation into the *V. cholerae* chromosomal integron, which is the case with other integrons.

6.2 Results

6.2.1 T7 RNA polymerase over-expression of *vlpA*.

Over-expression assays using the T7 RNA polymerase/promoter system were performed in an attempt to visualize the VlpA protein. The advantage of this system is the ability to make complete transcripts of almost any DNA that is placed under control of a T7 promoter (Studier *et al.*, 1990). In addition, it is possible to selectively inhibit the host RNA polymerase with rifampicin, permitting the exclusive expression of genes (Tabor and Richardson, 1985). The *vlpA* coding region was amplified by PCR from *V. cholerae* O139 (V911) chromosomal DNA, using oligonucleotide primers #1071 and #1072 which bind at both ends of *vlpA*. This PCR product was cloned into pGEM-T (Promega) in the T7 RNA polymerase promoter (P_{T7}) orientation, and designated pPM5007. This was confirmed by dye-terminator sequencing using the T7 promoter primer (#2926). pPM5007 was introduced into *E. coli* containing pGP1-2 (E2096) which contains the gene for T7 RNA polymerase under control of the λP_L promoter that is repressed by a temperature-sensitive repressor (*c*I857) (Ausubel *et al*, 1994). pPM5007 and pGP1-2 were maintained in the same cell by selection with kanamycin (for pGP1-2) and ampicillin (for pPM5007), and referred to as E5218.

T7 RNA polymerase over-expression was performed in E5218 along with E5217 (E2096 with pGEM5Zf+) as a negative control. Cells containing the two plasmids were grown overnight at 30°C, with induction at 42°C for 15 min, followed by incubation at 37°C for two hours. The amino acid sequence from the 516-bp *vlpA* gene encodes an approximately 19 kDa protein, and the protein size in this study (from E5218) is in agreement with the predicted size (lanes 5,6) which was not detected in the control E5217 (Figure 6.1A lanes 3, 4). Although rifampicin can be subsequently added to inhibit

Figure 6.1 Over-expression and detection of the VlpA proteins by SDS-PAGE (A), and cell fractionation (B).

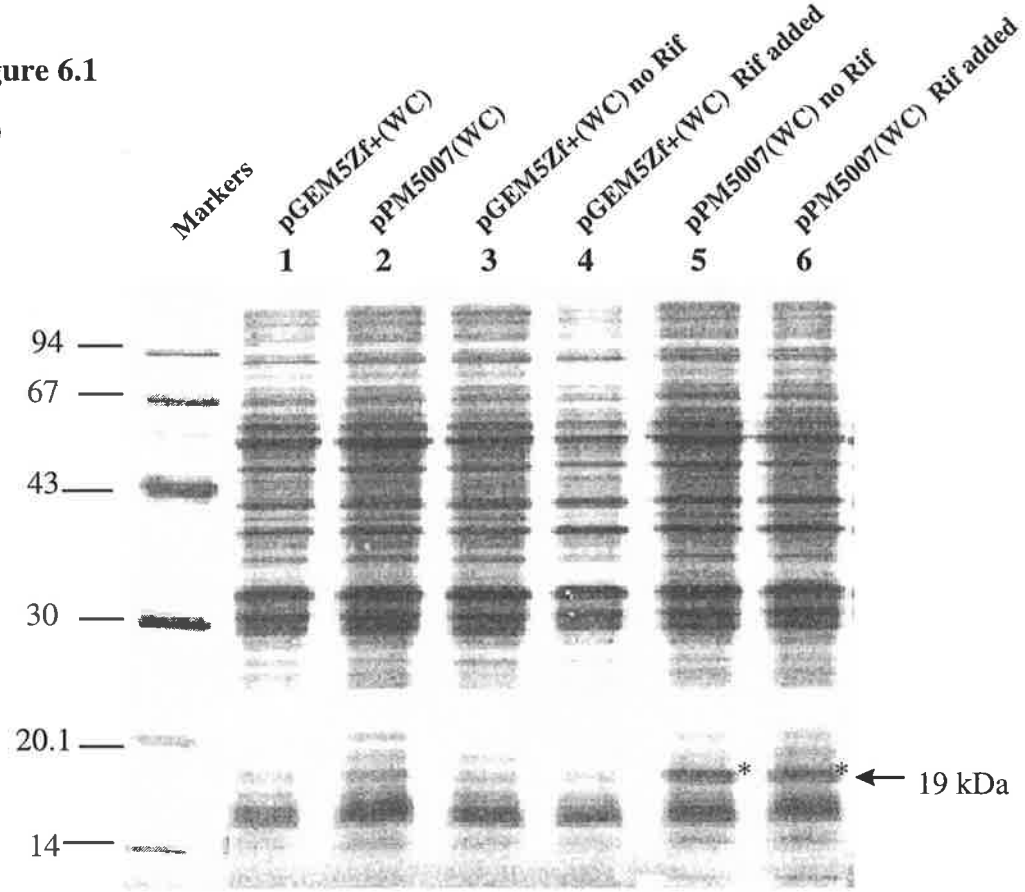
(A) The proteins were over-expressed using the T7 promoter / RNA polymerase system (Tabor and Richardson, 1985). Whole cells from E5218 (pPM5007) and E5217 (pGEM5Zf+), in the presence and the absence of rifampicin, were electrophoresed on a 15 % SDS polyacrylamide gel, then stained with Coomassie Brilliant Blue. Temperature shift (30°C to 42°C), was used to induce the λP_L promoter in lanes **3, 4, 5** and **6**.

(B) Cell-fractionation and Coomassie Blue stained on a 15% SDS-PAGE of E5217 and E5218 fractions.

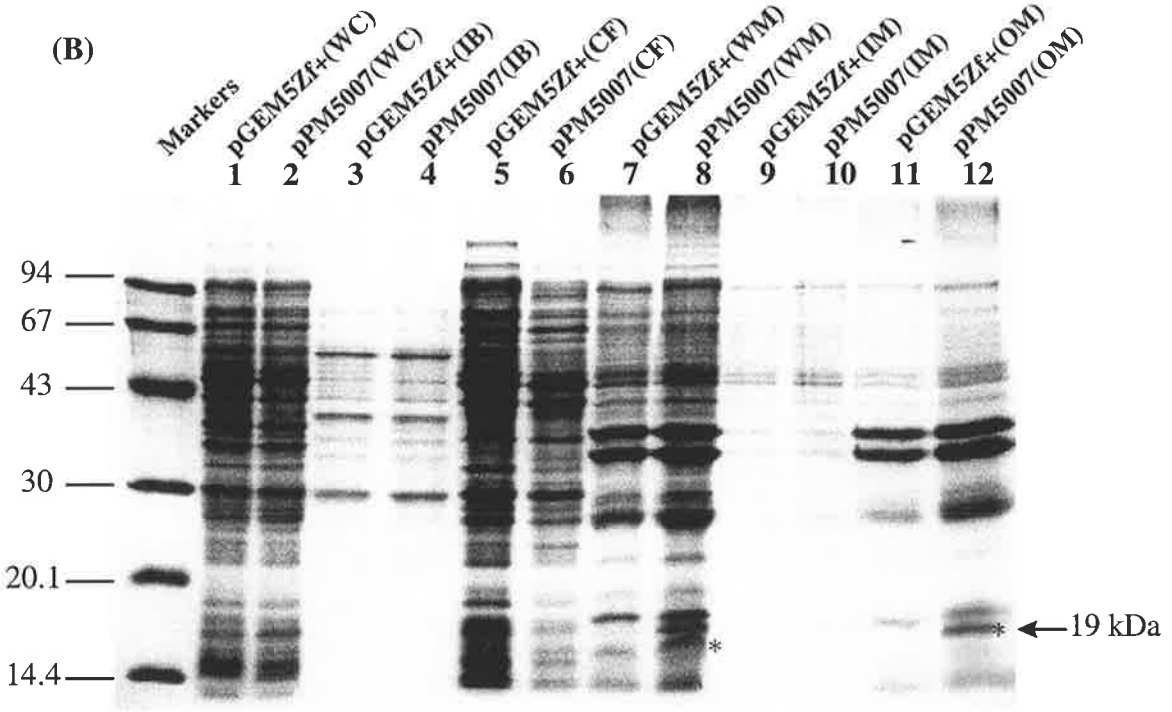
Plasmids harbored in the strains are indicated at the top of the figure, as are the fractionated cell components : whole cells (WC), inclusion bodies (IB), whole membranes (WM), cytoplasmic fraction (CF), inner membranes (IM) and outer membranes (OM). The ~19 kDa VlpA protein is shown by the asterisk. Migration positions of the molecular mass standards in kDa (Pharmacia) : alpha-lactalbumin (14.4), soybean trypsin inhibitor (20.1), carbonic anhydrase (30), ovalbumin (43), bovine serum albumin (67) and phosphorylase b (94).

Figure 6.1

(A)


Markers
pGEM5Zf+(WC)
pPM5007(WC)
pGEM5Zf+(WC) no Rif
pGEM5Zf+(WC) Rif added
pPM5007(WC) no Rif
pPM5007(WC) Rif added
1
2
3
4
5
6
94
67
43
30
20.1
14
*
*
19 kDa

(B)


Markers
pGEM5Zf+(WC)
pPM5007(WC)
pGEM5Zf+(IB)
pPM5007(IB)
pGEM5Zf+(CF)
pPM5007(CF)
pGEM5Zf+(WM)
pPM5007(WM)
pGEM5Zf+(IM)
pPM5007(IM)
pGEM5Zf+(OM)
pPM5007(OM)
1
2
3
4
5
6
7
8
9
10
11
12
94
67
43
30
20.1
14.4
*
*
19 kDa

transcription by *E. coli* RNA polymerase, in this case it was not necessary (Figure 6.1A lanes 5, 6).

6.2.2 Over-expression and cell-fractionation using pPM5007.

Cell-fractionation was performed in order to determine the cellular location of VlpA, and to enrich for protein to allow purification and antiserum production. A 500 ml culture of E5218 was induced at 42°C for 15 min to produce the ~19 kDa VlpA protein. E5217 and E5218 were fractionated into the following components : inclusion bodies (IB), whole membrane (WM), cytoplasmic fraction (CF), inner membrane (IM) and outer membrane (OM). The ~19 kDa band corresponding to the VlpA protein appeared exclusively in the whole cell membrane fraction (Figure 6.1B, lane 8), and this was further localized to the outer membrane fraction (lane 12), confirming its predicted cell surface location.

6.2.3 Development of anti-VlpA serum and VlpA detection.

The outer membrane fraction containing the VlpA protein was used to produce a specific antiserum. Proteins were electrophoresed on a 15% SDS polyacrylamide gel and the ~19 kDa protein band was excised from the gel (Figure 6.1A), then homogenized in incomplete Freund's adjuvant. The judicious use of adjuvants is essential to induce a strong antibody response to soluble antigens, although they may not always be required for particulate or whole-cell antigens (Harlow and Lane, 1988). The gel-purified VlpA was used to develop a rabbit polyclonal anti-VlpA serum by subcutaneously injecting a rabbit with 360 µg of purified soluble outer membrane protein. This crude antiserum has been used primarily in Western blot analysis with whole cell lysates of *E. coli* overproducing

VlpA (E5218). A protein of approximately 19 kDa was detected as the major component capable of reacting with the antiserum (Figure 6.2B), although other proteins were observed which reacted less intensely. The crude serum developed by this method was found by Western immunoblotting to contain one other band (~35 kDa) detected by non-specific cross-reactive antibodies in *E. coli* (Figure 6.2B). Non-specific antibodies from this antiserum were then absorbed out using partially purified outer membrane fractions from E5217.

Whole cell lysates of E5218 (equivalent to 1 x 10^8 cells) along with E5217 as a negative control were electrophoresed on a 15% SDS polyacrylamide gel. Western immunoblotting using absorbed anti-VlpA serum as a probe demonstrated that cross-reacting *E. coli* antibodies were no longer present (Figure 6.2C). This absorbed anti-VlpA serum was then used for further Western blot analysis with whole cell lysates and whole membrane fractions of *V. cholerae*.

6.2.4 Detection of VlpA in *V. cholerae*.

To see whether anti-VlpA serum could detect VlpA in *V. cholerae*, whole cell lysates from V911 (wild-type) and V911 derivatives containing *vlpA* mutants (V1269, V1271, V1279) (see **Section 5.2.2.1** and Table 6.1) were used. Western immunoblotting showed that the absorbed anti-VlpA serum cross-reacted with a larger protein of approximately 40 kDa found in whole cell lysates of all the *V. cholerae* strains tested (Figure 6.3A, lanes 1-4), but not in their outer membrane fractions of E5218 (lane 5). Conversely, this anti-VlpA serum did not give a detectable specific reaction of any proteins at the ~19 kDa region to the whole cell lysates of *V. cholerae* strains used in this study (lanes 1-4). Whole membrane components were then used to enrich for VlpA, but still no

Figure 6.2 Outer membranes from the over-expression and cell fractionation (A), Western Immunoblot of E5218 producing VlpA using crude antibodies (B), and absorbed antibodies (C).

(A) Coomassie Blue stained SDS-PAGE of E5217 and E5218 outer membrane (OM) fractions.

(B) Outer membranes from E5218 and E5217 were electrophoresed on a 15% SDS polyacrylamide gel and transferred to nitrocellulose. Anti-VlpA (crude antibodies) was used to detect proteins.

(C) Proteins were electrophoresed on a 15% SDS polyacrylamide gel and transferred to nitrocellulose membrane. Absorbed anti-VlpA antibodies were used to detect proteins. Temperature shift (30°C to 42°C), was used to induced the λP_L promoter in lanes **2, 4,** and **6.**

Plasmids harbored in the strains are indicated at the top of the figure. The fractions are: whole cells (WC) and outer membranes (OM). The 19 kDa VlpA outer membrane protein produced is indicated by an asterisk. Migration positons of the molecular mass standard (Pharmacia) are indicated on the left side (in kDa) : alpha-lactalbumin (14.4), soybean trypsin inhibitor (20.1), carbonic anhydrase (30), ovalbumin (43), bovine serum albumin (67) and phosphorylase b (94).

Figure 6.2

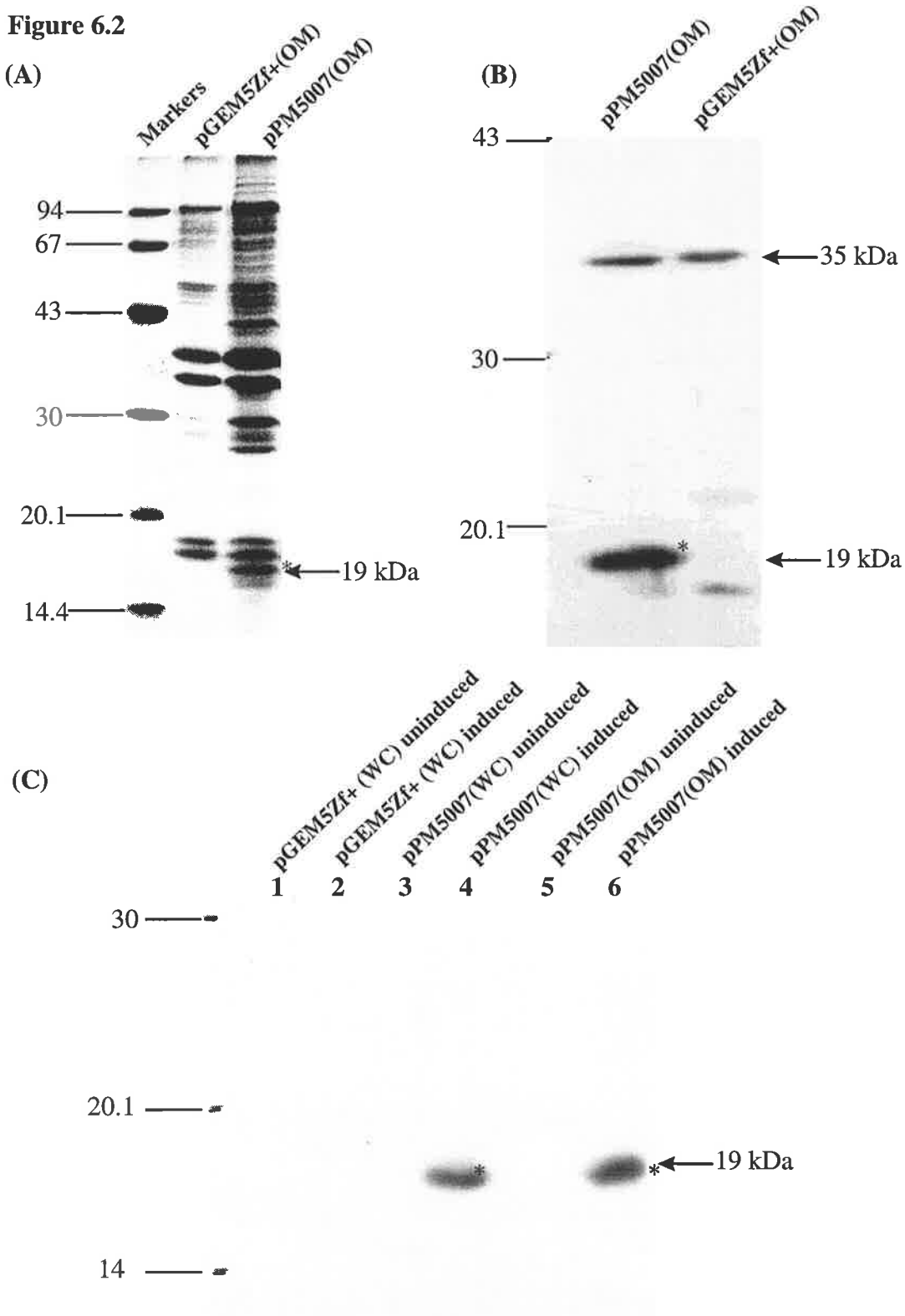

(A)
Markers
pGEM5Zf+(OM)
pPM5007(OM)
94
67
43
30
20.1
14.4
19 kDa
(B)
pPM5007(OM)
pGEM5Zf+(OM)
43
30
20.1
35 kDa
19 kDa
(C)
pGEM5Zf+ (WC) uninduced
pGEM5Zf+ (WC) induced
pPM5007(WC) uninduced
pPM5007(WC) induced
pPM5007(OM) uninduced
pPM5007(OM) induced
1 2 3 4 5 6
30
20.1
14
19 kDa

Figure 6.3 Western Immunoblot of *V. cholerae* producing VlpA.

Proteins from whole cell lysates (equivalent to 1 x 10^8) were electrophoresed on a 15% SDS polyacrylamide gel and transferred to nitrocellulose membrane. Absorbed anti-VlpA antibodies were used to detect the proteins. The VlpA protein is indicated with an arrow and asterisk on the right side of the figure.

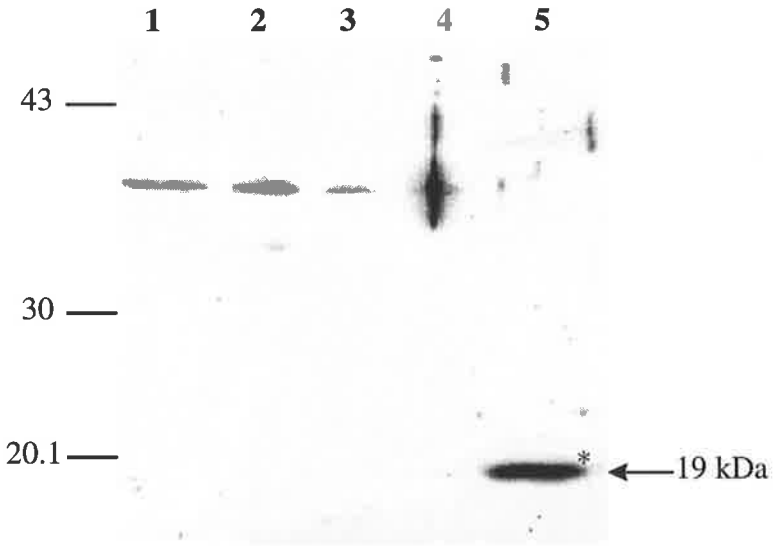
(A) Lanes contain the whole cells from: **1,** V911; **2,** V1269; **3,** V1271; **4,** V1279; and **5,** outer membrane fraction of VlpA expressed in *E. coli*.

(B) Lanes contain whole membranes from: **1,** V911; **2,** V1269; **3,** V1271; **4,** V1279; **5,** E5218; **6,** E5217; and **7,** outer membrane fraction of VlpA expressed in *E. coli*.

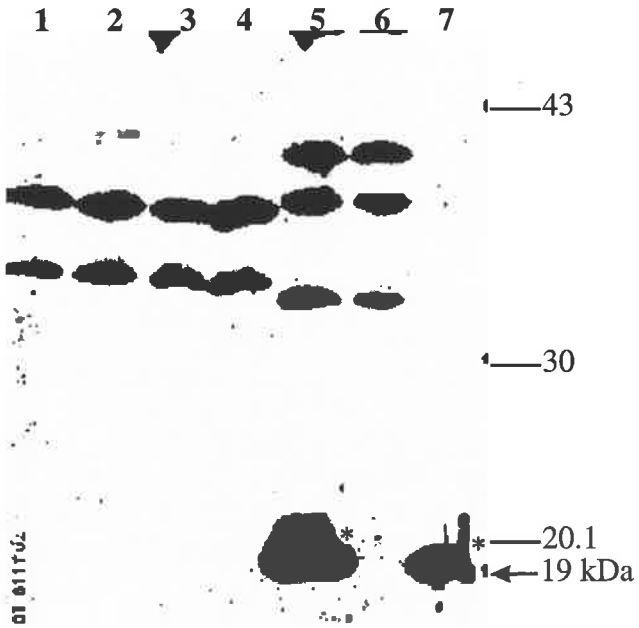
Migration positions of the molecular mass standards (Pharmacia) are indicated on the left side (in kDa) : alpha-lactalbumin (14.4), soybean trypsin inhibitor (20.1), carbonic anhydrase (30), ovalbumin (43), bovine serum albumin (67) and phosphorylase b (94).

Figure 6.3

(A)



(B)



specific bands at 19 kDa were detected in *V. cholerae* (Figure 6.3B). However, a number of cross-reacting bands were present at approximately 33 and 38 kDa (Figure 6.3B, lanes 1,2,3,4), whereas E5218 and E5217 (*E. coli* strains with a plasmid-encoded *vlpA*) also showed another band of 40 kDa (lanes 5,6). As the 33 kDa protein bands were present in all *V. cholerae* strains including *vlpA* mutants, these are presumably due to cross-reacting antibodies or antibodies specific for other outer membrane proteins still present in the absorbed antiserum. In addition, the whole cell components from *V. cholerae* O1 classical strain 569B and El Tor strain O17 (which contain 4 copies of *vlpA*) were also used to detect VlpA. However, Western immunoblotting showed the same results as mentioned above (data not shown). It is possible that other outer membrane proteins react with the anti-VlpA serum. Therefore, either the expression of the chromosomal copy of *vlpA* may be very low and insufficient to produce a detectable reaction, conditions favoring its expression were not used, or *vlpA* just is not expressed in this study and other strains of *V. cholerae*.

Gene cassettes do not generally include a promoter to signal initiation of transcription (Hall and Collis, 1998), and expression usually relies on a promoter P_{ant} upstream of the *attI* site in the integron. Therefore, expression of *vlpA* was determined in a number of different plasmids containing *vlpA* with or without an external promoter, i.e. the promoter supplied by the plasmid vectors. Western immunoblotting was performed using outer membrane purified VlpA from *E. coli* and whole membranes of E5218 as positive controls, whereas whole membranes from the following strains were used as negative controls: *V. cholerae* V1279 (*vlpA* double mutant), and *E. coli* DH5α harboring pPM5007 (E5208), pPM5010 (E5212), pPM5013 (E5214), pPM5024 (E5251), pGEM5Zf+ (E5265) and pGP1-2 (E2096), respectively (for details of these strains see Table 6.1). The strains which were used to determine the expression of *vlpA* from different promoters were E2096

harboring pPM5010 (E5250), *E. coli* harboring pPM5001 (E2889), pPM5016 (E5220), and *V. cholerae* V1279 carrying pPM5016 (V1281). The details of each plasmid are presented as follows and also listed in Table 6.2.

pPM5007 is the 516 bp *vlpA* open reading frame, minus VCRs, cloned into pGEM–T in the T7 RNA polymerase (P_{T7}) promoter orientation. This was used as a positive control. pPM5010 is pPM5007 digested with *Hin*dIII (present in the middle of the *vlpA* ORF at nt 260), end-filled using Klenow fragment of *E. coli* DNA polymerase I, and re-ligated. This causes a frame-shift mutation that results in a truncated VlpA. Similarly, plasmid pPM5013 also originated from pPM5007, but with a Km^R cartridge inserted at the unique *Hin*dIII site to interrupt the expression of *vlpA*. pPM5024 contains *vlpA* cloned in the reverse orientation in pGEM–T Easy, and therefore immediately downstream of the *lac* promoter.

Other plasmids used to detect the expression of *vlpA* were pPM5001 and pPM5016. pPM5001 is a 4.8-kb *Bam*HI fragment from *V. cholerae* O1 strain 569B containing *vlpA* and flanking genes cloned into pBluescript II KS+ as described in Chapter 3 **Section 3.2.2**. It was proposed that VlpA may be expressed from this plasmid if the *vlpA* promoter is located upstream of the *vlpA* coding region, i.e. within the VCR region. Plasmid pPM5016 contains the 516-bp fragment of the *vlpA* coding region cloned into the broad host-range vector pBBR1MCS-4 (Kovach *et al.*, 1995) immediately downstream of the *lac* promoter. pBBR1MCS-4 contains the T7 and *lac* promoter from nt 381 to nt 1031 of pBluescript II KS- (Stratagene) (Kovach *et al.*, 1994), and is stably retained in many bacterial genera *in vitro* (>10 days) and *in vivo* (>4 weeks in BALB/C mice) without antibiotic selection (Elzer *et al.*, 1995). This plasmid has been tested and found to replicate in *V. cholerae* (Antoine and Locht, 1992). pPM5016 was introduced into V1279 (the *vlpA* double mutant)

to see if VlpA would be expressed from this plasmid. The *lac* promoter (which is known to function in *V. cholerae*) in pPM5016 was induced using IPTG at a final concentration 0.5mM.

Whole membranes from *E. coli* and *V. cholerae* strains carrying the plasmids mentioned above, were analyzed by Western immunoblotting using absorbed anti-VlpA serum as a probe (Figure 6.4A). However, VlpA was observed only in E5218 (pGP1-2, pPM5007) and outer membrane-purified VlpA which were used as positive controls. The truncated VlpA that resulted from the frame-shift mutation of *vlpA* in pPM5010 (E5250) was not detected, even though it was also under control of the T7 promoter. This was probably due to increased proteolytic degradation of the truncated form. In addition, pPM5016 in which *vlpA* was induced from the *lac* promoter, did not express any detectable levels of *vlpA*. In E2889 (pPM5001) containing *vlpA* and the flanking genes, it was hypothesized that a promoter sequence could be located either within the VCRs or further upstream of the gene, however these was no detectable expression of *vlpA*. These results demonstrated that the expression of *vlpA* was observed only after over-expression of *vlpA* via a T7 promoter. As VlpA was not detected in *V. cholerae* strains, or under the control of other *E. coli* promoters, it is possible that the chromosomal copy of *vlpA* is not expressed, at least in the *V. cholerae* strains tested, or that this gene may only be expressed under specific conditions, some of which will be tested.

6.2.5 The expression of *vlpA* under iron starvation conditions

The lack of expression of *vlpA* in *V. cholerae* under normal growth conditions lead to further investigation as to, if and when this gene could be expressed. It was suggested by Barker and Manning (1997) that VlpA functions as a haemin binding protein, indicating

Figure 6.4 Western Immunoblot of *V. cholerae* O139 and *E. coli* producing VlpA.

Proteins from whole membranes (A) and whole cell lysates (B, C) were electrophoresed on a 15% SDS polyacrylamide gel and transferred onto nitrocellulose membrane. Absorbed anti-VlpA antibodies were used to detect proteins.

(A) The expression of *vlpA* with different promoters. Temperature shift (30°C to 42°C), was used to induced the λP_L promoter in lanes **3, 6,** and **7**. IPTG (50 mM for the final concentration) was used to induce *lac* promoter in lanes **4, 5, 8, 10** and **11**. Lanes contain: **1**, E5208 (pPM5007); **2**, E5212 (pPM5010); **3**, E5213 (pPM5010, pGP1-2); **4**, E5214 (pPM5013); **5**, E5254 (pPM5024); **6**, E5217 (pGEM5Zf+, pGP1-2); **7**, E5218 (pPM5007, pGP1-2); **8**, E5220 (pPM5016); **9**, V1279, (10) V1281 (pPM5016); **11**, E2889 (pPM5001); **12**, E2096 (pGP1-2); and **13**, VlpA (OM).

(B, C) The expression of *vlpA* under iron-limiting condition.

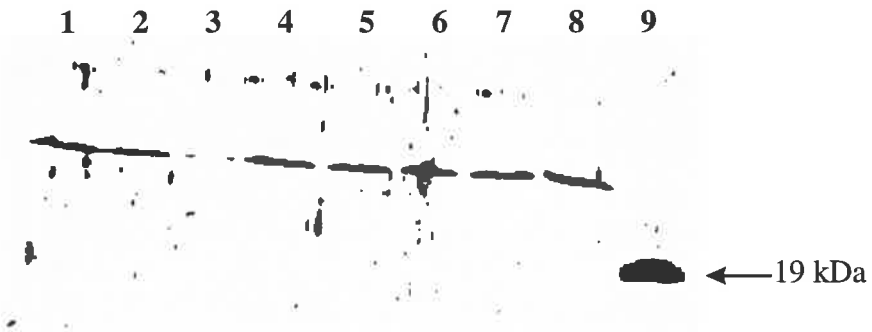
Strains in lanes **1-4, 5-8, 10-13** were grown in minimal medium, NB with 2,2′-dipyridyl added, and NB respectively. Temperature shift (30°C to 42°C), was used to induce the λP_L promoter in lanes **16, 17.** Lanes contain: **1**, V911; **2**, V1269; **3**, V1271; **4**, V1279; **5**, V911; **6**, V1269; **7**, V1271; **8**, V1279; **9**, VlpA (OM); **10**, V911; **11**, V1269; **12**, V1271; **13**, V1279; **14**, E5217 in NB; **15**, E5217 in NB with 2,2′-dipyridyl added; **16**, E5218 in NB; **17**, E5218 in NB with 2,2′-dipyridyl added; **18**, VlpA (OM).

Migration positions of the molecular mass standards (Pharmacia) are indicated on the left side (in kDa) : soybean trypsin inhibitor (20.1) and carbonic anhydrase (30).

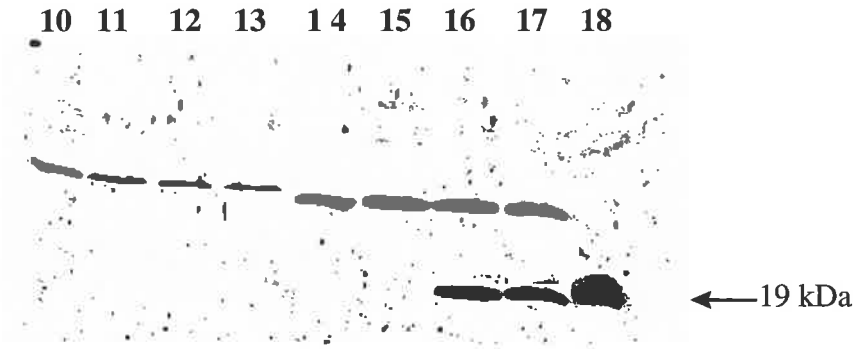
Figure 6.4

(A)

(B)



(C)



the *vlpA* gene may be expressed under conditions of iron starvation. Stressing for free iron was performed by adding the iron chelator 2,2′-dipyridyl into NB at a final concentration of 0.2 mM. This was compared with *V. cholerae* strains grown in minimal medium and NB. Single colonies of *V. cholerae* strains (V911, V1269, V1271 and V1279) were grown in NB, NB with 2,2′-dipyridyl added, and minimal medium at 37°C with shaking for five hours to an A_{600} = 0.5, whereas *E. coli* strains (E5127 and E5128) were grown at 30°C in the same media and induced at 42°C for 15 min before growing at 37°C for 2 hours.

Western immunoblotting using absorbed anti-VlpA serum as a probe was performed from the whole cell lysates after electrophoresis on a 15% SDS polyacrylamide gel. No positive bands at the 19 kDa region were observed in any of the different growth conditions used (Figure 6.4B, C). These results are consistent with Barker and Manning (1997) who could not detect *vlpA* mRNA transcripts in *V. cholerae* O1 strains 569B and O17 under similar conditions.

6.2.6 Analysis of transcription in the *vlpA* region

6.2.6.1 Northern blot analysis

Northern blots were performed to detect specific *vlpA* mRNA. Total RNA (30 μg) extracted from V911, V1269, V1271 and V1279 was separated by agarose gel electrophoresis. RNA was then transferred overnight onto nylon membrane (Amersham) at room temperature using 10X SSPE pH7.4 as the transfer buffer, as described by Sambrook *et al.* (1989). A radiolabeled PCR-generated *vlpA*-specific probe (oligonucleotides #1071 and #1072) to the RNA was then hybridized, followed by autoradiography. However, no hybridization was observed in all the *V. cholerae* strains tested. This indicated that the *vlpA* gene was not transcribed. Although it appears that no promoter is present within the gene

cassette or the VCR for *vlpA*, at least under the conditions used, gene activity may require additional as yet undefined factors.

6.2.6.2 Primer extension

Although Northern blot analysis did not detect any message, primer extension assays were also performed to identify the 5′ end start point of any transcripts of the message, if it existed. Primer extension is a sensitive assay and may have been successful in detecting a transcript which was not seen with Northern blots. The oligonucleotide primer #3092 which read out at position 12-bp from the start of *vlpA* was labeled at the 5′end using polynucleotide kinase and γ (^{32}P) ATP. RNA from V911, V1269 and V1279 in amounts of either 10ng or 20ng were used to compare the suitable amount for annealing and extension reactions. The dried extension products were separated on a conventional DNA sequencing gel at a constant voltage setting of 1200V (~25 mA). Although the gel was placed overnight in contact with a phosphorimage screen (Bio-Rad Molecular Imager® FX), and also various exposure times, no extension products were observed in this study (data not shown).

6.2.6.3 Reverse transcription-polymerase chain reaction (RT-PCR)

RT-PCR was performed as a final method to detect any possible message from the *vlpA* gene. This procedure is also extremely sensitive and reliable, in which RNA is reverse transcribed to produce complementary DNA (cDNA) templates for use with PCR, and is suitable for the detection of gene expression (Aatsinki, 1997). Specifically, deoxyoligonucleotides are used as primers for extension on mRNA templates.

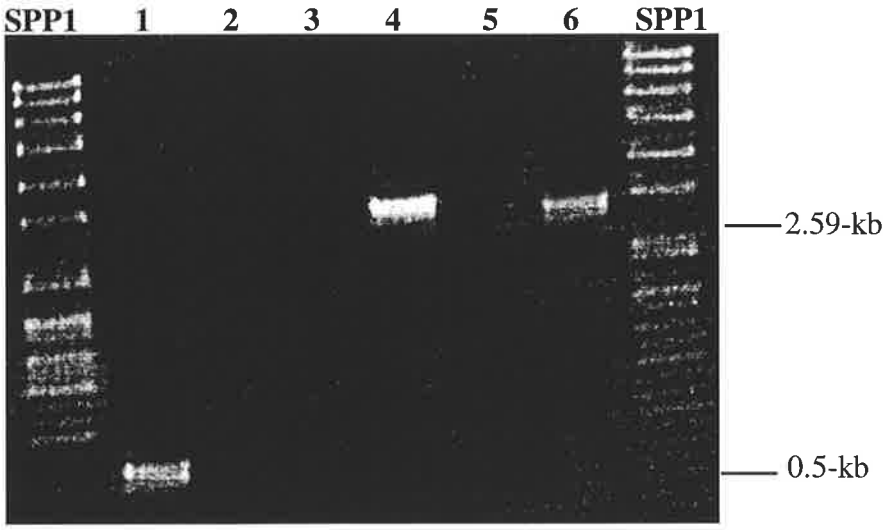
The procedures for RT-PCR in this study were described by Aasinki (1997) using Moloney murine leukemia virus (MMLV) from SuperScript^{TM} II (Life Technologies) as the choice of reverse transcriptase enzyme. Total RNA (50 ng) from *V. cholerae* O139 strains, i.e. V911, V1269, V1271 and V1279 was used as templates in the first strand DNA synthesis. The primer for cDNA synthesis was oligonuclotide #3170 which binds near the 3′ end (nt 486) and reads through *vlpA*. The cDNA was then used as a template for amplification in PCR. The primers for PCR were oligonucleotides #3170 and #3171 which bind at both ends (nt 38 and nt 486) of the *vlpA* gene. Ten percent of the first strand cDNA was used in the reaction with the initially denaturation at 95°C, 5 min. This was followed by 25 cycles of 30 sec denaturation at 95°C, 1 min annealing at 55°C, and 1 min extension at 72°C using a DNA thermal cycler (Perkin Elmer). No PCR products were observed, supporting the data from Northern blot and primer extension analyses which were also negative.

As a positive control for the RT-PCR, the oligonucleotides #1009 and #1019 which are specific for *hly*A (a known expressed gene in *V. cholerae*) (Williams and Manning, 1991), were used to confirm the procedure and the RNA quality. PCR amplification of the *V. cholerae* O139 (V911) chromosomal DNA spanning the *hlyA* region using oligonucleotides #1009 and #1019 generated a 2.5-kb fragment. Therefore, these primers were used in the RT-PCR for amplification cDNA from V911 as a positive control. Reverse transcription was observed when oligonucleotides #1009 and #1019 were used to detect *hly*A (Figure 6.5). This result suggests that the chromosomal *vlpA* gene is not expressed or expressed below detectable limits, under the conditions used here and specifically for *V. cholerae* O139 (**Sections 6.2.6.1 and 6.2.6.2**).

Figure 6.5 RT-PCR to detect *vlpA* mRNA.

Total RNA from *V. cholerae* O139 (V911) was used to produce cDNA which was then used as a template in the PCR-amplification. Oligonucleotides for *vlpA* (#3170 , #3171) were used in lanes **1-3**, and for *hlyA* (#1009, #1019) in lanes **4-6**. PCR products in each lane were from templates as follow: **1**, chromosomal DNA (positive control); **2**, cDNA with RT; **3**, cDNA with no RT; **4**, chromosomal DNA (positive control); **5**, cDNA with no RT; **6**, cDNA with RT.

Figure 6.5



6.3 Discussion

Expression of *vlpA* was detected in *E. coli*, with a predicted protein size of ~19 kDa. Interestingly, this protein was detected by Coomassie blue staining and later by Western blot analysis only when the powerful T7 promoter was placed upstream of the *vlpA* ORF. Cell fractionation confirmed that VlpA is localized in the outer membrane. The purified VlpA from the cell fractionation was used to develop a rabbit polyclonal anti-VlpA serum, and was used in Western blot analysis. This anti-VlpA serum reacted strongly against the outer membrane fraction of VlpA, and also whole cell lysates of *E. coli* (E5218) over-expressing, the protein using the T7 phage expression system (pGP1-2, pPM5007). To detect VlpA in *V. cholerae*, whole cell lysates from *V. cholerae* O139 and its mutated *vlpA* derivatives (i.e. V911, V1269, V1271, and V1279) were used. The anti-VlpA serum did not give a detectable specific reaction at 19 kDa to any of the *V. cholerae* O139 strains used in this study. In addition, whole cell lysates from *V. cholerae* O1 classical strain 569B and El Tor strain O17 (which have up to 4 copies of *vlpA*) were also used in Western blot analysis, and no specific reactions were detected (data not shown).

It was considered that the protein concentration from whole cell lysates may have been insufficient to detect very low levels of expressed VlpA using anti-VlpA serum. Therefore, to enrich for the protein whole membranes were used. However, there were no specific reactions to show that VlpA was expressed in *V. cholerae*. Interestingly, although different promoters from plasmid vectors were used in different strains of *E. coli* and *V. cholerae*, *vlpA* was expressed when under the control of the strong T7 promoter. Limiting iron availability by adding iron chelator 2,2′ dipyridyl, also did not result in *vlpA* expression in *V. cholerae* O139 strains in the stationary phase as determined by Western blot analysis.

Other procedures were also utilized to determine if transcription of *vlpA* occurred. Northern blot analysis using a radiolabeled ^{32}P-PCR-generated *vlpA*-specific probe did not show any hybridizing bands. Similarly, no signal was observed for primer extension. RT-PCR was also performed as this procedure is extremely sensitive and suitable for the detection of gene expression. Again, no specific products were observed.

It is possible that this gene does not have its own promoter as is the case with gene cassettes in other integrons. They lack their own promoter and do not generally include promoter-like signals (Stokes *et al.*, 1989; Lévesque *et al.*, 1994; Recchia and Hall, 1995). In addition, VCR might play a role in transcription termination in *V. cholerae* (Barker *et al.*, 1994) as with other 59-bes (Collis and Hall, 1995). Transcription has been shown to terminate within at least one of the repeats, but it seems unlikely that all copies of VCR act as transcription terminators (Barker *et al.*, 1994).

The expression of *vlpA* may require specific conditions, as is the case with the *blc* gene in *E. coli*. Bishop *et al.* (1995) reported that *blc* is normally expressed only weakly. This expression is optimized in stationary phase under the control of the RpoS σ-factor global regulator. A starvation-inducible lipoprotein (Slp) was shown to be expressed in stationary phase cultures independently of *rpoS* (Bishop *et al.*, 1995). However, stressing for iron using 2,2′ dipyridyl and stationary phase cultures in this study did not detect any VlpA.

To date, nothing is known concerning the expression of the cassette encoded genes in the *V. cholerae* chromosomal integron. It is expected that a promoter such as P_{ant} may exist upstream of the *attI*4 site (Rowe-Magnus *et al.*, 1999), although this has not yet been shown. This is consistent with Lévisque *et al.* (1994) who reported that a strong promoter is located upstream from the *attI* site, directing the transcription of the cassettes. The

activity of the P_{ant} promoter in class 1 integrons has been demonstrated experimentally (Stokes *et al.*, 1989; Collis and Hall, 1995), with the level of expression of gene cassettes highest when located closest to the P_{ant} promoter.

Data from the TIGR database show that the first *vlpA* is located at ~14.5-kb upstream of the *intI*4 gene in the *V. cholerae* El Tor strain N16961 chromosomal integron. There are 25 ORFs (VCA0292 - VCA0316) between these genes (Heidelberg *et al.*, 2000) (Chapter 3, Figure 3.14). Collis and Hall (1995) demonstrated that the level of expression of each gene cassette is influenced by its position in the integron and the number of upstream cassettes. In addition, some cassettes may carry their own promoter, and transcriptional read-through could allow the expression of cassettes located downstream of them (Collis and Hall, 1995; Hall, 1998; Rowe-Magnus *et al.*, 1999). Therefore, the lack of expression may be due to the fact that *vlpA* does not has its own promoter. In addition, the number of cassettes between *vlpA* and *intI*4 (25 ORFs) may too distant for the *vlpA* gene to be influenced by the P_{ant}. However, it is interesting to note that *vlpA* when placed under control of a strong *E. coli* promoter such as the *lac* promoter, does not produce a detectable protein band, even with Western blot analysis. This may indicate that *vlpA* is expressed at very low levels, with the block in protein expression at the translational level.

In conclusion, *vlpA* encodes an approximately 19 kDa outer membrane protein, but its expression from the *V. cholerae* chromosome is still unclear. Further analysis should be performed to determine if a P_{ant} like promoter is present within the *intI*4 gene and flanking DNA, and if in fact it is responsible for gene-cassette expression for genes such as *vlpA*.

Table 6.1 Bacterial strains used in this Chapter.

Bacterial strains	Characteristics	Source/Reference
E5208	DH5α with pPM5007	This study (**Section 6.2.4**)
E5212	DH5α with pPM5010	This study (**Section 6.2.4**)
E5213	DH5α with pPM5010 and pGP1-2	This study (**Section 6.2.4**)
E5214	DH5α with pPM5013	This study (**Section 6.2.4**)
E5217	DH5α with pPM5GEM5Zf+ and pGP1-2	This study (**Section 6.2.3**)
E5218	DH5α with pPM5007 and pGP1-2	This study (**Section 6.2.3**)
E5220	DH5α with pPM5016	This study (**Section 6.2.4**)
E5248	DH5α with pPM5001 and pGP1-2	This study (**Section 6.2.3**)
E5249	DH5α with pPM5016 and pGP1-2	This study (**Section 6.2.4**)
E5250	DH5α with pPM5010 and pGP1-2	This study (**Section 6.2.4**)
E5251	DH5α with pPM5024	This study (**Section 6.2.4**)
E5265	DH5α with pGEM5Zf+	This study (**Section 6.2.4**)
V911	O139 wild-type (AI-1837)	John Albert (I.C.D.D.R.). Strain was from Bangladesh
V1269	V911 with *vlpA*1::Km^R	This study (Chapter 5)
V1271	V911 with *vlpA*4::Km^R	This study (Chapter 5)
V1279	V911 with *vlpA*1::Tet^R and *vlpA*4:: Km^R	This study (Chapter 5)

Table 6.2 Plasmids used in this chapter.

Plasmid	Characteristics	Source/Reference
pPM5001	4.6-kb *Bam*HI fragment from *V. cholerae* 569B cloned into pBC-KS	This study (Chapter 3, **Section 3.2.2**)
pPM5007	516-bp *vlpA* from PCR amplification using primers #1071 and #1072 cloned into pGEM-T.	This study (Chapter 3)
pPM5010	pPM5007 digested with *Hin*dIII located at the center (nt260) of *vlpA*, then end-filled using Klenow fragment of *E. coli* DNA polymerase I, and re-ligated again	This study (**Section 6.2.4**)
pPM5013	1.1-kb Km^R cartridge from pUWeKT inserted into *Hin*dIII site in the center of *vlpA* in pPM5004.	This study (Chapter 4, **Section 4.2.2.1**)
pPM5016	516-bp *vlpA* cloned into pBBR1MCS4 at *Apa*I and *Sac*I sites in the *lacZ* promoter	This study (**Section 6.2.4**)
pPM5024	516-bp *vlpA* cloned into pGEM-T Easy in the *lacZ* promoter	This study (**Section 6.2.4**)
pBBR1MCS4	cloning vector Ap^R, *lacZ*	Kovach *et al.*, 1995
pGEM-T Easy	cloning vector Ap^R, *lacZ*	Promega
pBBR1MCS4	cloning vector Ap^R, *lacZ*	Kovach *et al.*, 1995

Chapter 7

Discussion

7.1 Introduction

Recently, the complete genome sequence of *V. cholerae* El Tor strain N16961 has been described by Heidelberg *et al.* (2000), representing the DNA sequence of both *V. cholerae* chromosomes. The small chromosome (chromosome 2) is 1.07 megabases, and carries the large integron island (125.3-kb) (Rowe Magnus *et al.*, 1999; Heidelberg *et al.*, 2000). There are four copies of *vlpA* in N16961 located at nt 325,225; nt 343,997; nt 378,083; and nt 398,645 in the integron area, and have been designated VCA0317, VCA0350, VCA0406, and VCA0443 respectively (http://www.tigr.org).

The *vlpA* gene encodes ~19 kDa outer membrane lipoprotein which belongs to a group of lipocalins, or more specifically, the α_2-microglobulin superfamily. Members of this superfamily are generally found in eukaryotic organisms, and only a very few prokaryotic lipocalins have been reported. Very little is known about the function of bacterial lipocalins. *V. cholerae* is unique in that it contains a number of copies of this gene, in some strains up to four copies. Therefore, the characterization of a number of aspects of this novel prokaryotic gene including distribution, sequence and copy number, expression and function in *V. cholerae* O139 were undertaken in this study.

7.2 Distribution of the *vlpA* gene in the chromosome of *Vibrio spp.*.

The *vlpA* gene is present as multiple copies in all strains of *V. cholerae* O1, O139 and some of the non-O1 serogroups. However, this gene was not detected in other *Vibrio* species, such as *V. mimicus*, *V. fluvialis*, and *V. parahaemolyticus*, the only exception being

V. anguillarum, the most closely-related of the species. DNA sequence analysis of *vlpA* (**Section 3.2.3.2**) showed sequence identity from 97-99% within *V. cholerae* strains, and 95% between *V. cholerae* and *V. anguillarum*, indicating that this gene is highly conserved and spread within these *Vibrio* species. It is interesting to note that in *V. cholerae*, the *vlpA* gene is linked to the integron. This leads to the question of how the genes were generated.

The *vlpA* gene encodes a lipocalin which until recently have only found in eukaryotic organisms, mostly vertebrates. To date, three bacterial lipocalins (Blc) have been reported, i.e. CfBlc (from *Citrobacter freundii*, GenBank accession No. U21727), EcBlc (from *E. coli*, GenBank accession No. P39281), and VlpA (from *V. cholerae*, GenBank accession No. AF025662, and AF025663) in this study. It is possible that the bacterial lipocalins may have spread among the bacterial population or they may been acquired individually by horizontal transmission from a eukaryotic organism. Flower (1995) argued that the existence of one prokaryotic lipocalin can be explained by the mechanism of horizontal gene transfer, but with several examples from different species this seems less likely. The probability that a specific gene will be successfully transferred to a new host depends on specific mechanisms such as transformation, transduction and conjugation (Jain *et al.*, 1999), and even the presence of integrases in organisms (Hall, 1998). It is now known that a number of the virulence-associated genes can be transferred horizontally between *V. cholerae* strains by phages or other mobile elements (Taylor, 1999; Waldor and RayChaudhuri, 2000).

7.3 VCRs are involved in gene capture in the *V. cholerae* chromosomal integrons.

DNA sequence analysis in **Section 3.2.3** has demonstrated that *vlpA* is a mobile gene cassette consisting of a single ORF, associated with a VCR located downstream of the gene. Gene cassettes are mobilized by site-specific recombination at the 59-be seven-base

core site, and VCRs contain all the features of such an element (Hall, 1998). This study has shown that the *vlpA* gene varies significantly in copy number and chromosomal location in different strains of *V. cholerae* and other *Vibrio* species. In all cases, the presence of *vlpA* was associated with a VCR. The presence of the VCR element was suggestive of such site-specific exchange, and the VCR element was demonstrated in **Section 4.2.3** to function as a specific recombination site for gene capture in *V. cholerae*.

It is now accepted that DNA encoding a diverse range of genes in prokaryotes are acquired by horizontal transfer, both inter- and intra-species as well as different genera. This can lead to the formation of new pathogenic strains, e.g. the presence of "pathogenicity islands". It is reasonable to assume that the *vlpA* gene would be exchanged between *Vibrio* species via the gene capture system of integrons. The association between this gene and VCR, and the presence as multiple copies of *vlpA* in different *Vibrio* strains would therefore be a consequence of insertion at different sites.

There are at least 155 copies of VCR, and 216 ORFs (VCAs) located on the *V. cholerae* El Tor (strain N16961) integron (Heidelberg *et al.*, 2000). Most of these VCAs are also associated with VCRs (Mazel *et al.*, 1998; Rowe Magnus *et al.*, 1999; http://www.tigr.org). Some VCAs are present in more than one copy in the integron area, for example, *mccF* (VCA0337, VCA0439) homologous to the immunity protein for microcin C7 (MccC7) (Gonzalez-Pastor, 1995), and *relB* (VCA0349, VCA0504) (homologous to a translational inhibitor that accumulates during amino-acid starvation in *E. coli*) (Beach *et al.*, 1985), which are present as two identical copies. Despite the high variability in 59-be sequences, it has been established that antibiotic resistance cassettes can be substrates for the integron-encoded integrases (Mazel *et al.*, 1998), thus gene acquisition seems to be the easiest way to develop antibiotic resistance effects. Rowe-Magnus and Mazel (1999) suggested that the combination of the mobile cassettes in

integrons and the selection pressure exerted by antibiotics used, may have driven the specific capture of resistance cassettes. VCRs might contribute substantially to the process of gene acquisition, by providing a mechanism for gene capture.

7.4 Chromosomal integrons capture mobile gene cassettes.

Karaolis *et al.* (1995) reported that genes for major virulence factors such as TCP, can be transferred horizontally, and antigenic conversion can be achieved by the acquisition and loss of these genes. There is the formal possibility that any non-virulent *V. cholerae* strain could be transformed into a virulent strain, even an epidemic one (Karaolis *et al.*, 1995; Mekalanos *et al.*, 1997).

Mobile genetic elements may facilitate transfer of numerous pathogenic virulence genes between bacterial species (Finlay and Falkow, 1989). Based on polymorphism within the *V. cholerae* housekeeping gene encoding the aspartate-semialdehyde dehydrogenase (*asd*), Karaolis *et al.* (1995) have inferred that horizontal transfer of the *V. cholerae* O-antigen biosynthesis genes may occur, suggesting that *V. cholerae* could show a high level of intraspecies genetic exchange. The ability of *vlpA* cassettes to spread from one *Vibrio* species to another can be facilitated by the fact that this gene is located in an integron.

Gene capture is a significant pathway in the acquisition of antibiotic resistance gene (Hall, 1998; Rowe-Magnus and Mazel, 1999), although conjugation is the main process by which these genes disseminate among bacterial populations (Row-Magnus and Mazel, 1999). Moreover, antibiotic resistance genes are mobile units which can be inserted into several independent locations in different plasmids and transposons (Hall and Vockler, 1987; Stokes and Hall, 1989; Collis and Hall, 1992a). While integrons were previously thought to capture mobile gene cassettes on a strictly functional basis (Hall and Collis, 1995), this is the case with the *vlpA* gene cassette (which is present in multiple copies) in

the *Vibrio* chromosome. No virulence is associated with *vlpA* as the insertional inactivation, or even deletion of this gene (and surrounding genes) had no effect in the infant mouse model as demonstrated in **Section 5.2.4**. This suggests that even though this gene cassette is not a virulence gene or does not confer any known advantage to *V. cholerae*, it still can be captured or transferred within *Vibrio* species mediated by the integron system.

7.5 Integrase IntI4 mediates cassette rearrangement in the *Vibrio spp.* chromosome.

Until now, only integrase IntI1 has been demonstrated to be required for the movement of gene cassettes (Martinez and de la Cruz, 1990; Hall *et al.*, 1991; Collis and Hall, 1992a, 1992b; Collis *et al*, 1993; Recchia *et al.*, 1994). However, integrase IntI4 has been shown for the first time in this study to catalyze cassette deletion in a plasmid clone containing *vlpA*/VCR, and more specifically in *V. cholerae* O139. Although IntI1 is more efficient than IntI4 (**Section 4.2.3.5**), the deletion mediated by both integrases still occurred at the VCR seven-base core site. IntI1 can recognize VCRs in this study in *E. coli* and *V. cholerae*. Although IntI1 was shown to recognize VCRs in *E. coli* by Mazel *et al.* (1998), this was cassette integration, not deletion as in this study. The *vlpA* cassette deletion was also predicted to generate a covalently closed-circular gene cassette, and this was confirmed by PCR analysis (**Section 4.2.3.6**). These circularized cassettes can therefore presumably be re-inserted into the chromosome at different VCR sites. Circular cassettes resulting from cassette deletion may re-insert into another chromosome especially during cell division, leading to multiple copies of the gene cassette in a new cell.

Interestingly, a larger size non-replicating circular cassette was also detected by PCR, implying that more than one gene cassette was excised during the deletion event. This is known to occur with IntI1 (Collis and Hall, 1992b). Whether this was as a result of

over-expression of IntI1 and / or IntI4 (i.e. under control of the *lac* promoter), or if this is the case with normal levels of expression under control of its own promoter is not clear, but it allows the movement of more than a single gene in the one circular cassette. In fact, when comparing other flanking genes of *vlpA*, some conservation is seen between different strains. Therefore, these cassettes may move as either a single cassette unit or as a larger gene cassette. To test this hypothesis, it was expected that larger deletions around the *vlpA* region could be detected. In fact, this was seen to be the case with a single gene deletion, where the recombination crossover occurred at either side of VCRs flanking *vlpA*. Larger deletions also occurred often to an unknown extent, as flanking genes further upstream or downstream of *vlpA* were unknown and therefore could not be mapped. Due to variation of the integron region in *V. cholerae* strains, it has not possible to use the information from the sequenced strain N16961 to map this deleted region further.

It is interesting to note that no background deletions (monitored by loss of antibiotic resistance, in our case Km^R or Tet^R inserted into *vlpA*) was observed when no plasmid-expressed copy of *intI*4 was present. Whether this represents a low level of *intI*4 expression, and therefore not enough to detect movement with the numbers screened in our system, or if *intI*4 is even expressed, is not clear. In any case, the amount of the cassette movement would be expected to be much lower than other plasmid-based integrons as it potentially involves movement of large regions of chromosomal DNA, and therefore constraints may be placed on the degree of chromosomal shuffling the cell can tolerate. If this is the case, low regulation of *intI*4 expression may have occurred over time evolved. The requirement for this is to know if in fact IntI4 is expressed from the chromosome. Future work, could focus on the expression of *intI*4; however, it should be noted that specific conditions may be required as well.

Evidence for cassette movement is two-fold, firstly the variation in integron gene cassette composition and distribution within and between *Vibrio* species. For example, the *vlpA* cassettes in different strains of *V. cholerae* are not located at the same position as shown in *V. cholerae* O1 classical strain 569B, El Tor strains O17 and N16961 and the O139 serogroup. Secondly, a PCR-generated copy of *intI4* is active, even though the *lac* promoter was used and a higher copy number of the gene is present due to its plasmids based location. A possible interesting application stemming from this work, is the ability to delete more than one gene cassette. This approach could be used to delete larger regions of the *V. cholerae* chromosomal integron, thereby making it feasible to investigate the possible role of the integron-located genes in processes such as virulence. Deleting larger regions of DNA containing at least ten years is far easier than the step-wise deletion of over 150 individual cassettes.The rearrangement of the gene cassettes is likely to be important in the dissemination and evolution of plasmids and bacterial genomes. Recchia and Hall (1995) reported that integration of a cassette at a secondary site leads to the permanent acquisition of a new gene.

7.6 The origin of gene cassettes associated with VCR.

The genome sequence of *V. cholerae* El Tor strain N16961 has revealed a large chromosomally-located integron with an array of gene cassettes, representing over 11% of the smaller chromosome 2 (Heidelberg *et al.*, 2000). These gene cassettes contain two functional components, a gene coding region and a VCR (commonly referred as a 59-be). Among the cassettes in strain N16961 integron, some genes appear more than once, for example *mccF* and *vlpA*, which are present in two and four copies, respectively. The *mccF* gene in *E. coli* encodes the resistance or immunity gene to microcin, a bacteriocin

(Gonzalez-Pastor *et al.*, 1995), whereas *vlpA* still has an unknown function, like the majority of the gene cassettes found in this integron.

Bacterial lipocalins have been found in a restricted number of species. They have homology to the mammalian apolipoprotein D (ApoD), and to Lazarillo found in the nervous system of the American grasshopper, *Schistocera americana*. This raises the possibility that bacterial lipocalins originated or were acquired by horizontal transmission from a eukaryotic source (Bishop and Weiner, 1998; Ganfornina *et al.*, 2000). However, the other bacterial lipocalins are not associated with a VCR or 59-be.

It is still unclear about the origins of the VCR or 59-be. Hall (1991) proposed a model for the formation of cassettes through a reverse transcription mechanism, suggesting that the VCRs were added to the ORFs to act as a primer for reverse transcriptase. The original gene cassettes may be formed by the reverse transcription of mRNA molecules (Hall, 1991). The cassettes may have many different origins, and the VCRs may be either part of the original transcript or added later to the ORFs. However, several *Vibrio* cassettes have their own promoter and / or are in the opposite orientation compared to the VCR (Rowe-Magnus and Mazel, 1999). These characteristics conflict with a model described by Hall (1991). Whether the same constructs are placed on this larger chromosomally-located integron or if this reflects a new model remains to be established.

7.7 The expression of *vlpA*.

The *vlpA* gene product has only been observed when expressed under the control of the T7 promoter (**Section 6.2.1**). Furthermore, *in situ* expression was not observed from the *V. cholerae* chromosome even under a variety of growth conditions. Northern blot, primer extension and RT-PCR analysis could not detect any *vlpA* message. It is possible that the *vlpA* gene cassette does not contain its own promoter, like many known gene cassettes

which do not generally include promoter signals. Expression of cassette-associated gene is usually dependent on integration of the cassette in the correct orientation into an integron that supplies an upstream promoter (Recchia and Hall, 1997).

Mazel *et al.* (1998) demonstrated that a cluster of ribosomal protein genes in *V. cholerae* is located immediately downstream from the *intI*4 gene. Specific signals were detected in both *V. mimicus* and *V. metschnikovii* by Southern hybridization with an *intI*4 probe. If all transcripts detected commenced at the integron promoter (P_{ant}), which by analogy to class1-3 integrons is located within *intI*4, the level of expression in gene cassettes would be affected by the distance between the promoter and the cassettes. Furthermore, the expression of gene cassettes would not only be affected by the position of cassettes in the integron, but the presence of additional upstream cassettes could also reduce the expression of downstream gene due to VCRs acting as transcriptional terminators. Levesque *et al.* (1994) reported that a strong promoter in a class 1 integron is located upstream from the *attI* site, and directs the transcription of the cassettes. Studies in P_{ant} have been performed (Collis and Hall, 1995; Recchia and Hall, 1995), which showed that all cassette-encoded genes are dependent on P_{ant} for their expression. Longer transcripts exist that cover several cassettes, but the majority of transcripts are shorter corresponding to single genes.

A second strong promoter, P_2 (Levesque *et al.*, 1994) can be created by the insertion of three bases between otherwise poorly spaced -35 and -10 signals. When the weakest version of P_{ant} is present, the level of cassette expression from a second promoter is increased. The genomic sequence of *V. cholerae* El Tor strain N16961 shows the first *vlpA* is located at 14.5-kb from *intI*4, with 25 VCAs in between, that may reduce the influence of the integron promoter. If the expression of *vlpA* relies on an integron promoter,

the distance between *vlpA* and *intI*4 may be too far for expression from P_{ant}. Further studies should be undertaken to determine if *intI*4 has a P_{ant}-like promoter.

Although a promoter for *vlpA* is not apparent in the upstream sequence, and no message specific for this gene could not be detected in this study, a number of integron gene cassettes are known to carry their own promoters. Stokes and Hall (1991) reported the *cmlA* gene cassette found in the integron In4 of Tn*1696* has a promoter required for the expression of the *cmlA* resistance gene. Barker (1996) also identified the *mrhA* promoter lying immediately 5′ to *mrhA*. The expression of *mrhA* is independent of the ToxRST and HlyU regulons (two regulators of virulence gene expression) (DiRita, 1992; William *et al.*, 1993). Attempts to identify the putative promoters for ORFs adjacent to the *mrhA* gene using a primer extension assay (Barker, 1993) were not successful. Since a VCR element is an imperfect inverted repeat, which can form a stable stem-loop structure (Barker *et al.*, 1994), it is possible that VCRs play a role in transcriptional termination. However, the VCR located upstream of *mrhA* is unlikely to act as a terminator because of the polarity of the flanking ORFs (Barker *et al.*, 1994). Further studies are required to address this issue.

7.8 Function of VlpA.

Insertional mutation in both chromosomal copies of *vlpA* did not reduce the virulence of *V. cholerae* O139 in the infant mouse model. The function of VlpA is still unclear, although it has been demonstrated to bind haemin, suggesting that it has a role in iron uptake (Barker *et al.*, 1997). Bishop *et al.* (1995) proposed that bacterial lipocalins may contribute to the adaptation of cells to starvation conditions. Growth of *V. cholerae* under iron limiting conditions using 2,2′ dipyridyl as an iron chelator, did not result in any detection of *vlpA* (**Section 6.2.5**). The bacterial lipocalins are membrane lipoproteins which are closely related to eukaryotic lipocalins. In general, the members of this family have

been classified as extracellular transport proteins (Flower, 1996). Future studies of VlpA would aim to define the substrates that may be transported by VlpA. These studies at present would be limited to *E. coli* where VlpA can only be expressed when placed under T7 promoter control.

7.9 The relationships between the integron and the *V. cholerae* smaller chromosome.

The presence of a large integron island on the *V. cholerae* on smaller chromosome (chromosome 2) raises the question as to how this chromosome originated. It may have arisen by the excision of a large segment of DNA from the ancestral genome catalyzed by the integron, or conversely, it was originally a megaplasmid captured by an ancestral *Vibrio* species (Heidelberg *et al.*, 2000). The origin of chromosome 2 can be explained in both ways using the gene capture mechanism of the *V. cholerae* integron.

If chromosome 2 was a megaplasmid before it was captured by an ancestral *Vibrio* species, either integron and/or phage integration could have had an affect on its increasing size, and the appearance of new genes in the megaplamid genome. The megaplasmid presumably acquired genes from diverse bacterial species before its capture by the ancestral *Vibrio* species (Heidelberg *et al.*, 2000). Many genes can integrate into the *V. cholerae* genome via phage vectors. These genes may be acquired inter- or intra-species. For example, the cholera toxin phage (CTXϕ) is the genome of a virus that integrates into the *V. cholerae* chromosome (Waldor and Mekalanos, 1996), containing the cholera toxin (*ctx*) gene. The *V. cholerae* classical biotype contains two copies of the *ctx* gene, and other *V. cholerae* strains can carry several copies of these elements (Trucksis *et al.*, 1998; Heidelberg *et al.*, 2000).

Conversely, the smaller chromosome may have been derived from chromosome 1 by cassette excision. Cassette deletion demonstrated in Section 4.2.3.5 revealed that gene

cassette deletion of *vlpA* from chromosome 2, mediated by either integrase IntI1 or IntI4, was not just limited to a single gene cassette, but resulted in the excision of many gene cassettes. It is possible that a large region of chromosome 1 was excised and re-circularized. If this fragment containing the integron accumulated several essential genes over time, including those encoding the ribosomal proteins L20 and L35, and the ability to replicate independently, this cassette could have evolved to formed a new chromosome (Waldor and RayChaudhuri, 2000).

If chromosome 2 was derived from cassette excision, the possibility for integration back into chromosome 1 also exists. However, the probability of this re-integration event is presumably less likely. Chromosome 2 carries the integron region, whereas it does not appear in chromosome 1. Therefore, gene cassettes have more chance of inserting into chromosome 2 at the integron region catalyzed by integrase.

7.10 Non-toxigenic *V. cholerae* have the capacity to be virulent via gene capture or phage integration.

Until recently, only *V. cholerae* strains of the O1 serotype were known to be the causative agents of cholera epidemics (Barua, 1992), whereas the non-O1 strains were usually associated with only sporadic cases of cholera (Bik *et al.*, 1995; 1997). Therefore, it was totally unexpected that the cholera epidemic which began in India late in 1992 was caused by a non-O1 strain with the novel serogroup O139 (Albert *et al.*, 1993; Albert, 1994). Colwell *et al.* (1995) suggested that *V. cholerae* non-O1 strains may convert to the *V. cholerae* O1 serotypes and vice versa under suitable conditions, a possible strategy for survival in the environment.

Toxigenic *V. cholerae* strains may have arisen from a strain containing a "*V. cholerae* pathogenicity island" (VPI), that are always present in epidemic and pandemic

strains, but absent from nonpathogenic strains (Karaolis *et al.*, 1998). VPI contains the gene cluster encoding TCP, an accessory colonization factor (ACF), and the virulence gene regulator ToxR which are located near each other (Brown and Taylor, 1995). The *tcpA* gene encodes an important colonization factor and the receptor for the CTXϕ. The *toxT*, *tcpP*, and *tcpH* genes encode regulators of virulence genes. VPI is proposed to be an originally functional phage that was acquired by *V. cholerae*, as it was found in two clinical non-O1/ non-O139 cholera toxin-positive strains (Karaolis *et al.*, 1998). It was suggested that VPI can be transferred within *V. cholerae*, then modified making it defective and unable to excise (Karaolis *et al.*, 1998). If the VPIϕ and CTXϕ have been introduced into *V. cholerae* by phage, the nonpathogenic strains may become toxigenic through the acquisition of these virulence determinants.

Interestingly, an integrase gene (*int*) that is closely-related to the *E. coli* cryptic prophage (CP4-57) integrase protein, is located adjacent to the TCP gene cluster (Kovach *et al.*, 1996), and VPI is flanked by the *attI* sites which function as specific attachment sites. Genomic analysis of *V. cholerae* O1 and non-O1 strains showed that only pathogenic strains possess the TCP-ACF in association with the integrase (Karaolis *et al.*, 1998). Although this *int* gene is different from the *intI*4 in the *V. cholerae* integron, the site-specific recombination at the *attI* sites, that is catalyzed by these integrase genes, has been implicated in the acquisition of virulence determinants (Kovach *et al.*, 1996; 1998; Mazel *et al.*, 1998). Future studies may attempt to observe the frequency of recombination between *attI* sites located in the integron and in the VPI. In addition, the introduction of virulence genes into VPI by phage should be also investigated.

7.11 Future directions

V. cholerae O1 has been highly successful in causing the epidemic cholera, and this has stimulated the study of virulence gene acquisition. At presence, there is little understanding of how non-O1 strains acquire the potential to cause epidemic cholera. It has been proposed that many genes within eukaryotes and prokaryotes have been acquired by horizontal transfer, utilizing a variety of genetic mechanisms, including gene capture by integrons, phage integration, transposition, conjugation, and uptake of naked DNA. This project has examined the *vlpA* gene which is located within the *V. cholerae* chromosomal integron area. The *vlpA* gene cassette has been shown to delete from integron when catalyzed by either integrase IntI1 from *Pseudomonas aeruginosa* or IntI4 from *V. cholerae*. The recombination event occurs at the VCR element seven-base core site. Although it has not been shown in this study that *vlpA* is a virulence gene, cassette deletion is still linked to other genes involved in virulence. The study of the diverse functions of eukaryotic lipocalins may lead to new insights into the role of VlpA in the life cycle of *V. cholerae*.

Bibilography

Aatsinki, J. T. 1997. Coupled one-step reverse transcription and polymerase chain reaction procedure for cloning large cDNA fragments. Vol. 67. In, White, B.A. (ed.). *Methods in Molecular Biology, PCR cloning protocols : From Molecular cloning to Genetic Engineering.* Humana Press, Totowa, New Jersey, USA.

Abremski, K. E. and Hoess, R. H. 1992. Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Protein Eng.* **5:**87-91.

Aiba, H., Adhya, S. and de Crombrugghe, B. 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J. Bio. Chem.* **256:**11905-11910.

Albert, M. J., Siddique, A. K., Islam, M. S., Faruque, A. S., Ansaruzzaman, M., Faruque, S. M. and Sack, R. B. 1993. Large outbreak of clinical cholera due to *Vibrio cholerae* non-O1 in Bangladesh. *Lancet.* **341:**704.

Albert, M. J. 1994. *Vibrio cholerae* O139 Bengal. *J. Clin. Microbiol.* **32:**2345-2349.

Albert, M. J. 1996. Epidemiology and molecular biology of *Vibrio cholerae* O139 Bengal. *Indian J.Med. Res.* **104:**14-27.

Alm, R. A., Stroeher, U. H. and Manning, P. A. 1988. Extracellular proteins of *Vibrio cholerae* : nucleotide sequence of the structural gene (*hlyA*) for the haemolysin of the haemolytic El Tor strain O17 and characterization of the *hlyA* mutation in the non-haemolytic classical strain 569B. *Mol. Microbiol.* **2:**481-488.

Alm, R. A. and Manning, P. A. 1990. Characterization of the *hlyB* gene and its role in the production of the El Tor haemolysin of *Vibrio cholerae* O1. *Mol. Microbiol.* **4:**413-425.

Alm, R. A., Mayrhofer, G., Kotlarski, I. and Manning, P. A. 1991. The amino-terminal domain of the El Tor haemolysin of *Vibrio cholerae* O1 is expressed in classical strains and is cytotoxic. *Vaccine.* **9:**588-594.

Altschul, S. F., Gish, W., Myers , E. W. and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215:**403-410.

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

Antoine, R. and Locht, C. 1992. Isolation and molecular characterization of a novel broad-host-range plasmid from *Bordetella bronchiseptica* with sequence similarities to plasmids from gram-positive organisms. *Mol. Microbiol.* **6:**1785-1799.

Aragawa, Y., Murakami, M. and Suzuki, K. 1995. A novel integron-like element carrying the metallo-β-lactamase gene bla_{IMP}. *Antimicrob. Agents. Chemother.* **39:**1612-1615.

Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalinoes, B., Narayana, S. V. L., Pierson, L. S. III., Sternberg, N. and Leong, J. M. 1986. The integrase family of site-specific recombinase : regional similarities and global diversity. *EMBO J.* **5:**433-440.

Arita, M., Honda, T., Miwatani, T., Ohmori, K., Takao, T. and Shimonishi, Y. 1991. Purification and characterization of a new heat-stable enterotoxin produced by *Vibrio cholerae* non-O1 serogroup Hakata. *Infect. Immun.* **59:**2186-2188.

Arruda, L. K., Vailes, L. D., Hayden, M. L., Benjamin, D. and Chapman, M. D. 1995. Cloning of cockroach allergen, Bla g4, identifies ligand binding proteins (or calycins) as a cause of IgE antibody responses. *J. Biol. Chem.* **270:**31196-31201.

Attridge, S. R. 1979. Infection and Immunity to *Vibrio cholerae*. Ph.D. thesis, Department of Microbiology and Immunology, The University of Adelaide.

Attridge, S. R. and Rowley, D. 1983. The role of the flagellum in the adherence of *Vibrio cholerae*. *Infect. Dis.* **147:**864-872.

Attridge, S. R. and Rowley, D. 1993. The role of the toxin-coregulated pili in the pathogenesis of *Vibrio cholerae* El Tor. *Micro. Pathog.* **15:**421-431.

Attridge, S. R., Manning, P. A., Holmgren, J., and Jonson, G. 1996. Relative significance of mannose-sensitive haemagglutinin and toxin-coregulated pili in colonization of infant mice by *Vibrio cholerae* El Tor. *Infect. Immun.* **64:**3369-3373.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Albright, L. M., Coen, D. M. and Varki, A. 1994. Current protocols in Molecular Biology. Vol. 1-3. In, Janssen (ed.). John Wiley & Sons, U.S.A.

Balbin, M., Freije, J. M., Fueyo, A., Sanchez, L. M. and Lopez-Otin, C. 1990. Apolipoprotein D is the major protein component in cyst fluid from women with human breast gross cystic disease. *Biochem. J.* **271:**.

Band, L., Yansura, D. G. and Henner, D. J. 1983. Construction of a vector for cloning promoters in *Bacillus subtilis*. *Gene.* **26:**313-315.

Barker, A. 1993. Haemagglutinins of *Vibrio cholerae* O1 : studies on the organization of the genes encoding the mannose-fucose-resistant haemagglutinin (MFRHA). Ph. D. Thesis. Department of Microbiology and Immunology, The University of Adelaide.

Barker, A., Clark, C. A. and Manning, P. A. 1994. Identification of VCR, a repeated sequence associated with a locus encoding a haemagglutinin in *Vibrio cholerae* O1. *J. Bacteriol.* **176:**5450-5458.

Barker, A. and Manning, P. A. 1997. VlpA of *Vibrio cholerae* O1 : the first bacterial member of the α_2-microglobulin lipocalin superfamily. *Microbiol.* **143:**1805-1813.

Barker, A. and Manning, P. A. 1998. Origin of *Vibrio cholerae* lipocalin : which gene came first? *Microbiol.* **144:**816.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Natl. Acad. Sci. USA.* **91:**2216-2220.

Barrette, T. J. and Blake, P. A. 1981. Epidemiological usefulness of changes in hemolytic activity of *Vibrio cholerae* biotype El Tor during the seventh pandemic. *J. Clin. Microbiol.* **13:**126-129.

Barry, A. L. and Thornsberry, C. 1991. Susceptibility tests : diffusion test procedures, p 1117-1125. In, Balows, A., Hausler, Jr., W. J., Herrmann, K. L., Isenberf, H. D. and Shadomy, H. J. (eds.). *Mannual of clinical Microbiology.* 5th ed. American Society for Microbiology, Washington D.C.

Bartolomé, B., Jubete, Y., Martinez, E. and de la Cruz., F. 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene.* **102:**75-78.

Barua, D. 1992. History of cholera. In, Barua, D. and Greenough III, W. B.. (eds.). *History of cholera.* Plenum Medical Book, New York, USA. pp.1-36.

Baumann, P., Furniss, A. L. and Lee, J. V. 1984. Genus I *Vibrio* Pacini 1854, 411AL pp. 518-528. In, Krieg, N. R. and. Holt, J. G. (eds.). *Bergey's Mannual of Systematic Bacteriology.* Vol 1. Williams & Wilkins, Baltimore, MD, USA.

Bauuman, P. and Schubert, R. H. W. 1984. Family II *Vibrionaceae* Veron 1965, 5245AL. pp.516-517. In, Krieg, N. R. and Holt, J. G. (eds.). *Bergey's Mannual of Systematic Bacteriology.* Vol 1. Williams & Wilkins, Baltimore, MD, USA.

Beckmann, M. P., Betsholtz, C., Heldin, C. H., Westermark, B., Dimarco, E., Difiore, P. P., Robbins, K. C. and Aaronson, S. A. 1988. Comparison of biological properties and transforming potential of human PDGF-A and PDGF-B chains. *Science.* **77:**34-38.

Beneson, A. 1991. Cholera. In, Evans, A. S. and P. S. B. (eds.). *Bacterial infectious of humans.* Plenum Medical Book, New York.

Benitez, J. A., Spelbrink, R. G., Silva, A., Phillips, T. E., Stanley, C., Boesman-Finkelstein, M. and Finkelstein, R. A. 1997. Adherence of *Vibrio cholerae* to cultured differentiated human intestinal cells : an *in vitro* colonization model. *Infect. Immun.* **65:**3474-3477.

Bhaskaran, K. and Gorrill, R. H. 1957. A study antigen variation in *Vibrio cholerae. J. Gen. Microbiol.* **16:**721-729.

Bhattacharya, M. K., Bhattacharya, S. K., Garg, S., Saha, P. K., Dutta, D., Nair, G. B., Deb, B. C. and Das, K. P. 1993. Outbreak of *Vibrio cholerae* non-O1 in India and Bangladesh. *Lancet.* **341:**1346-1347.

Bik, E. M., Bunschoten, A. E., Gouw, R. D. and F. R. Mooi. 1995. Genesis of the novel epidemic *Vibrio cholerae* O139 strain : evidence for horizontal transfer of genes involved in polysaccharide synthesis. *EMBO J.* **14:**209-216.

Bik, E. M., Gouw, R. D. and Mooi, F. R. 1996. DNA fingerprinting of *Vibrio cholerae* strains with a novel insertion sequence element : a tool to identify epidemic strains. *J. Clin. Microbiol.* **34:**1453-1461.

Birch, A., Hausler, A., Ruttener, C. and Huttner, R. 1991. Chromosomal deletion and rearrangements in *Streptomyces glaucescens. J. Bacteriol.* **173:**3531-3538.

Bishop, R. E. and Weiner, J. H. 1993. Overproduction, solubilization, purification and DNA-binding properties of Amp^R from *Citrobacter freundii. Eur. J. Biochem.* **213:**405-412.

Bishop, R. E., Penfold, S. S., Frost, L. S. and Holtje, J. V. 1995. Stationary phase expression of a novel *Escherichia coli* outer membrane lipoprotein and its relationship with mammalian apoliprotein D. *J. Biol. Chem.* **270:**23097-23013.

Bishop, R. E. and Weiner, J. H. 1996. 'Outlier' lipocalins more than peripherol. *Trends Biochem. Sci.* **21:**127.

Bishop, R. E. and Weiner, J. H. 1998. Enterobacterial lipocalins precede *Vibrio* homologue. *Microbiol.* **144:**815-816.

Bissonnette, L., Champetier, S., Buisson, J.-P. and Roy, P. H. 1991. Characterization of the nonenzymatic chloramphenicol resistance (*cml*) gene of the In4 integron of Tn*1696* : similarity of the product to transmembrane transport proteins. *J. Bacteriol.* **173:**4493-4502.

Bissonnette, L. and Roy, P. H. 1992. Characterization of *Pseudomonas aeruginosa* plasmid pVSI, an ancestor of integrons of multiresistance plasmids and transposon of Gram-negative bacteria. *J. Bacteriol.* **174:**1248-1257.

Blake, P. A. 1994. Endemic cholera in Australia and the United States. In, Wachsmuth, I. K., Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives*. pp. 309-319. American Society for Microbiology, Washington, DC.

Blomfield, I. C., Vaughn, V., Rest, R. F. and. Eisenstein, B. I. 1991. Allelic exchange in *Escherichia coli* using the *Bacillus subtilis sacB* gene and a temperature-sensitive pSC101 replicon. *Mol. Microbiol.* **5:**1447-1457.

Booth, I. W., Levine, M. M. and Harries, J. Y. 1984. Oral rehydration therapy in acute diarrhea in childhood. *J. Pediatr. Gastroenterol. Nutr.* **3:**491-499.

Booth, B. A., Sciortino, C. V. and Finkelstein, R. A. 1986. Adhesins of *Vibrio cholerae*. In, Mirelman, D.(ed.). *Microbial lectins and agglutinins*. John Wiley and Sons, New York.

Boyles, J. K., Notterpek, L. M. and Anderson, L. J. 1990. Accumulation of apolipoproteins in the regenerating and remyelinating mammalian peripheral nerve : identification. *J. Biol. Chem.* **265:**17805--17815.

Brass, J. M. 1986. The cell envelope of Gram-negative bacteria : New aspects of its function in transport and chemotaxis. Vol. 129. In, Clarke, A. *et al.* (eds.). *Current topics in Microbiology and Immunity*. Springer - Verlag, New York, USA.

Braun, V., Schonherr, R. and Hobbie, S. 1993. Enterobacterial hemolysins: activation, secretion and pore formation. *Trends Microbiol.* **1:**211-216.

Brock, T. D., Madigan, M. T., Martinko, J. M. and Parker, J. 1994. Biology of microorganisms. 7th ed. Prentice - Hall, Englewood Cliffs, New Jersy USA.

Brown, M. C. M., Weston, A., Saunders, J. R. and Humphreys, G. O. 1979. Transformation of *Escherichia coli* C600 by plasmid DNA at different phases of growth. *FEMS Microbiol. Lett.* **5:**219-222.

Brown, C. R. and Taylor, R. K. 1995. Organization of *tcp*, *acf*, and *toxT* genes within a ToxT-dependent operon. *Infect. Immun.* **16:**425-439.

Bunny, K. L., Hall, R. M. and Stokes, H. W. 1995. New mobile gene cassettes containing an aminoglycoside resistance gene, *aac* A7, and a chloramphenicol resistance genes, *catB*3, in an integron in pBWH301. *Antimicrob. Agents. Chemother.* **39:**686-693.

Burnet, F. M. and Stone, J. D. 1947. Desquamation of intestinal epithelium by *Vibrio cholerae* filtrates. Characterization of mucinase and tissue disintegrating enzymes. *Aust. J. Exp. Biol. Med. Sci.* **25:**219-226.

Burnet, F. M. 1949. Ovomucin as a substrate for the mucinolytic enzymes of *Vibrio cholerae* filtrates. *Aust. J. Exp. Biol. Med. Sci.* **27:**245-252.

Burrows, W., Mather, A. N., McGann, V. G. and Wagner, S. M. 1946. Studies on immunity to Asiatic cholera. Part II The O and H antigenic structure of the cholera and related vibrios. *J. Infect. Dis.* **79:**168-179.

Cameron, F. H., Groot Obbink, D. J., Ackerman, V. P. and Hall, R. M. 1986. Nucleotide sequence of the AAD (2′) aminoglycoside adenylatetransferase determinant *aadB*. Evolutionary relationship of this region with those surrounding *aadA* in R538-I and *dhfrII* in R388. *Nucleic. Acid. Res.* **14:**8625-8635.

Campbell, K. S., Fuchs, B. A. and Munson, A. E. 1992. The T-lymphocyte is the primary cellular target for potentiation of the *in vitro* T-dependent IgM antibody response by the B subunit of cholera toxin. *Int. J. Immunopharmacol.* **14:**111-120.

Carroll, P. A., Tashima, K. T., Rogers, M. B., DiRita, V. J. and Calderwood, S. B. 1997. Phase variation in *tcpH* modulates expression of the ToxR regulon in *Vibrio cholerae*. *Mol. Microbiol.* **25:**1099-1111.

Champion, G. A., Neely, M. N., Brennan, M. A and DiRita, V. J. 1997. A branch in the ToxR regulatory cascade of *Vibrio cholerae* revealed by characterization of *toxT* mutant strains. *Mol. Microbiol.* **23:**323-331.

Chang, A. C. Y. and Cohen, S. N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. *J. Bacteriol.* **134:**1141-1156.

Chao, L., Vargas, C., Spear, B. B. and Cox, E. C. 1983. Transposable elements as mutator genes in evolution. *Nature*. **303:**633-635.

Clark, C. A., Purins, L., Kaewrakon, P. and Manning, P. A. 1997. VCR repetitive sequence elements in the *Vibrio cholerae* chromosome constitute a mega-integron. *Mol. Microbiol.* **26:**1137-1138.

Collis, C. M. and Hall, R. M. 1992a. Site-specific deletion and rearrangement of integron insert genes catalysed by the integron DNA integrase. *J. Bacteriol.* **174:**1574-1585.

Collis, C. M. and Hall, R. M. 1992b. Gene cassettes from the insert region of integrons are excised as covalently closed circles. *Mol. Microbiol.* **6:**2875-2885.

Collis, C. M., Grammaticopoulos, G., Briton, J., Stokes, H. W. and Hall, R. M. 1993. Site-specific insertion of gene cassettes into integrons. *Mol. Microbiol.* **9:**41-52.

Collis, C. M. and Hall, R. M. 1995. Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob. Agents. Chemother.* **39:**155-162.

Colwell, R. R. 1970. Polyphasic taxonomy of the genus vibrio : numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. *J. Bacteriol.* **104:** 410-433.

Colwell, R. R., Huq, A., Chowdhury, M. A., Brayton, P. R. and Xu, B. 1995. Serogroup conversion of *Vibrio cholerae*. *Can. J. Microbiol.* **41:**946-950.

Condon, C. and Weiner, J. H. 1988. Fumarate reductase of *Escherichia coli*: an investigation of function and assembly using *in vivo* complementation. *Mol. Microbiol.* **2:**43-52.

Connell, T., Metzger, D. J., Lynch, J. and Folster, J. 1998. Endochitinase is transported to the extracellular milieu by the esp-encoded general secretory pathway of *Vibrio cholerae*. *J. Bacteriol.* **180:**5591-5600.

Crennell, S., Garman, E., Laver, G., Vimr, E. and Taylor, G. 1994. Crystal structure of *Vibrio cholerae* neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. *Structure*. **2:**535-544.

Cromie, G., Collins, J. and Leach, D. 1997. Sequence interruptions in enterobacterial repeated elements retain their ability to encode well-folded RNA secondary structures. *Mol. Microbiol.* **24:**1311-1314.

Dalsgaard, A., Skov, M. N., Serichantalerges, O., Echeverria, P., Meza, R. and. Taylor, D. N. 1997. Molecular evolution of *Vibrio cholerae* O1 strain isolated in Lima, Peru, from 1991 to 1995. *J. Clin. Microbiol.* **35:**1151-1156.

Davis, B. R., Fannning, G. R., Madden, J. M. and *et. al.* 1981. Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species *Vibrio mimicus*. *J. Clin. Microbiol.* **14:**631-639.

de Lorenzo, Herrero, M., Jakubzik, U. and Timmis, K. 1990. Mini-Tn*5* transposon derivatives for insertional mutagenesis, promoter probing and chromosomal insertion of cloned DNA in Gram-negative DNA. *J. Bacteriol.* **172:**6568-6572.

de Moor, C. E. 1949. Paracholera (El Tor), enteritis choleriformis El Tor, Van Loghem. Bull. WHO. **2:**5-19.

Delidow, B. C. 1993. Molecular cloning of polymerase chain reaction fragment with cohesive ends. pp. 217-228. In, White, B. A. (ed.). *Methods in molecular biology.* Vol 15. *PCR protocols : Current methods and applications.* Totowa, N.J., U.S.A.

Deneer, H. G., Healey, V. and Boychuk, I. 1995. Reduction of exogenous ferric iron by a surface-associated ferric reductase of *Listeria spp. Microbiol.* **141:**1985-1992.

DePaola, A. 1981. *Vibrio cholerae* in marine foods and environmental waters : literature review. *J. Food Sci.* **46:**66-70.

DePaola, A., Kaysner, C. A. and McPhearson, R. M. 1987. Elevated temperature method for recovery of *Vibrio cholerae* from oysters (*Crassostrea gigas*). *Appl. Environ. Microbiol.* **53:**1181-1182.

Desmarchelier, P. M. and Reichelt, J. L. 1981. Phenotypic characteristic of clinical and environmental isolate of *Vibrio cholerae* from Australia. *Curr. Microbiol.* **5:**123-127.

DiRita, V. J., Parsot, C., Jander, G. and Mekalanos, J. J. 1991. Regulatory cascade controls virulence in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA.* **88:**5403-5407.

DiRita, V. J. and Mekalanos, J. J. 1991. Periplasmic interaction between two membrane regulatory proteins, ToxR and ToxS, results in signal transduction and transcriptional activation. *Cell.* **64:**29-37.

DiRita, V. J. 1992. Coordinate expression of virulence genes by ToxR in *V. cholerae*. *Mol. Microbiol.* **6:**451-458.

DiRita, V. J. 1994. Multiple regulatory systems in *Vibrio cholerae* pathogenesis. *Trends Microbiol.* **2:**37-38.

DiRita, V. J., Neely, M., Taylor, R. K. and Bruss, P. M. 1996. Differential expression of the ToxR regulon in classical and El Tor biotypes of *Vibrio cholerae* is due to biotype-specific control over ToxT expression. *Proc. Natl. Acad. Sci. USA.* **93:**7991-7995.

Donnenberg, M. S. and Kaper, J. 1991. Construction of an *eae* deletion mutant of Enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* **59:**4310-4317.

Douet, J. P., Castroviejo, M., Dodin, A. and Bebear, C. 1996. Study of the haemolytic process and receptors of thermostable dirct haemolysin from *Vibrio parahaemolyticus*. *Res. Microbiol.* **147:**687-696.

Eisenach, K. D., Cave, M. D., Bates, J. H. and Crawford, J. T. 1990. Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J. Infect. Dis.* **161:**977-981.

Espeli, O. and Boccard, F. 1997. *In vivo* cleavage of *Escherichia coli* BIME-2 repeats by DNA gyrase : genetic characterization of the target and identification of the cut site. *Mol. Microbiol.* **26:**767-777.

Esposito, D. and Scocca, J. J. 1997. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. *Nucleic Acids Res.* **25:**3605-3614.

Everiss, K. D., Hughes, K. J. and Peterson, K. M. 1994a. The accessory colonization factor and toxin-coregulated pilus gene clusters are physically linked on the *Vibrio cholerae* O395 chromosome. *DNA Seq.* **5:**51-55.

Everiss, K. D., Hughes, K. J., Kovach, M. E. and Peterson, K. M. 1994b. The *Vibrio cholerae acfB* colonization of determinant encodes an inner membrane protein that is related to a family of signal-transducing proteins. *Infect. Immun.* **62:**3289-3298.

Faruque, S. M., Abdul-Alim, A. R., Rahman, M. M., Siddique, A. K., Sack, R. B. and Albert, M. J. 1993. Clonal relationships among classical *Vibrio cholerae* O1 strains isolated between 1961 and 1992 in Bangladesh. *J. Clin. Microbiol.* **31:** 2513-2516.

Faruque, S. M., Asadulghami, M. N. Saha, A. R. Alim, M. J. Albert, K. M. Islam and J. J. Mekalanos. 1998a. Analysis of clinical and environmental strains of nontoxigenic *Vibrio cholerae* for susceptibility to CTXϕ : molecular basis for origination of new strains with epidemic potential. *Infect. Immun.* **66:**5819-5825.

Faruque, S. M., Albert, M. J. and Mekalanos, J. J. 1998b. Epidemiology, Genetics, and Ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.* **62:**1301-1314.

Fasano, A., Baudry, B., Pumplin, D. W., Wasserman, S. S., Tall, B. D., Ketley J. M. and Kaper, J. B. 1991. *Vibrio cholerae* produces a second enterotoxin which affects intestinal tight junctions. *Proc. Natl. Acad. Sci. USA.* **88:** 5242-5246.

Fasano, A. 1998. Novel approaches for oral delivery of macromolecules. *J. Pharm. Sci.* **87:**1351-1356.

Finkelstein, R. A. and Mukerjee, S. 1963. Hemagglutination : a rapid method for differntiating *Vibrio cholerae* and El Tor vibrios. *Proc. Soc. Exp. Biol. Med.* **112:**355-359.

Finkelstein, R. A. 1975. Immunology of cholera. *Curr. Top. Microbiol. Immunol.* **69:**138-196.

Finkelstein, R. A. and Hanne, L. F. 1982. Purification and characterization of the soluble hemagglutinin (cholera lectin) produced by *Vibrio cholerae*. *Infect. Immun.* **36:**1199-1208.

Finkelstein, R. A., Boesman-FInkelstein, M., and Holt, P. 1983. *Vibrio cholerae* hemagglutinin/lecitin/protease hydrolyzes fibronectin and ovomucin : F.M. Burnet revisited. *Proc. Natl. Acad. Sci. USA.* **80:**1092-1095.

Finkelstein, R. A., Burks, M. F., Rieke, L. C., McDonald, R. J., Browne, S. K. and Dallas, W. S. 1985. Application of monoclonal antibodies and genetically-engineered hybrid B-subunit proteins to the analysis of the cholera/coli enterotoxin family. *Dev. Biol. Stand.* **59:**51-62.

Finkelstein, R. A. and Dorner, F. 1985. Cholera enterotoxin (choleragen). *Pharmacol. Ther.* **27:**37-47.

Finkelstein, T. A. 1988. Cholera-the great teacher : 100 years since Koch's postulates. In, Takeda, Y. and Kuwahara, S. (eds.). *Vibrio cholerae and cholera.* KTK Scientific Publishers, Tokyo.

Finkelstein, R. A. 1992. Cholera enterotoxin (choleragen)-a historical persfective. pp.155-187. In, Barua D. and Greenough III, W.B. (eds.). *Cholera*. Plenum Medical Book, New York.

Finkelstein, R. A., Boesman-Finkelstein, M., Chang, Y. and Hase, C. C. 1992. *Vibrio cholerae* hemagglutinin/protease, colonial variation, virulence, and detachment. *Infect. Immun.* **60:**472-478.

Finlay, B. B. and Falkow, S. 1989. Common themes in microbial pathogenicity. *Microbiol. Rev.* **53:**210-230.

Finn, T. M., Reiser, J., Germanier, R. and Cryz, S. J. 1987. Cell associated hemagglutinin deficient mutant of *Vibrio cholerae*. *Infect. Immun.* **55:**942-946.

Flower, D. R., North, A. C. T. and Attwood, T. K. 1991. Mouse oncogene protein 24p3 is a member of the lipocalin protein family. *Biochem. Biophys. Res. Commun.* **180:**69-74.

Flower, D. R. 1993a. Structural relationship of streptavidin to the calycin protein superfamily. *FEBS Lett.* **333:**99-102.

Flower, D. R., North, A. C. T. and Attwood, T. K. 1993b. Structure and sequence relationship in the lipocalins and related proteins. *Protein Sci.* **2:**753-761.

Flower, D. R. 1994. The lipocalin protein family : a role in cell regulation. *FEBS Lett.* **354:**7-11.

Flower, D. R., Sansom, C. E., Beck, M. E. and Attwood, T. K. 1995a. The first prokaryotic lipocalins. *Trends Biochem. Sci.* **20:**498-499.

Flower, D. R. 1995b. Multiple molecular recognition properties of the lipocalin protein family. *J. Mol. Recogn.* **8:**185-195.

Flower, D. R. 1996. The lipocalin protein family : structure and function. *Biochem. J.* **318:** 1-14.

Focareta, T. and Manning, P. A. 1991. Distinguishing between the extracellular DNases of *Vibrio cholerae* and development of a plasmid transformation system. *Mol. Microbiol.* **5:**2547-2555.

Francia, M. V., dela Cruz, F. and Garcia Lobo, J. M. 1993. Secondary sites for integration mediated by the Tn*21* integrase. *Mol. Microbiol.* **10:**823-828.

Francia, M. V., Zabala, J. C., de la Cruz, F. and Garcia-Lobo, J. M. 1999. The IntI1 integron integrase preferentially binds single-stranded DNA of the *attC* site. *J. Bacteriol.* **181:**6844-6849.

Franzon, V. L. and Manning, P. A. 1986. Molecular cloning and expression in *Escherichia coli* K-12 of the gene for a haemagglutinin from *Vibrio cholerae*. *Infect. Immun.* **52:**279-284.

Franzon, V. L. 1988. Haemagglutinins of *Vibrio cholerae* : molecular characterization of the mannose-fucose resistant haemagglutinin (MFRHA). Ph. D. Thesis. Department of Microbiology and Immunology, The University of Adelaide.

Franzon, V. L., Barker, A. and Manning, P. A. 1993. Nucleotide sequence encoding the mannose-fucose-resistant haemagglutinin of *Vibrio cholerae* O1 and construction of a mutant. *Infect. Immun.* **61:**3032-3037.

Freter, R. P., O' Brien, C. M. and Macsal, M. S. 1981. Role of chemotaxis in the association of motile bacteria with intestinal mucosa : *in vivo* studies. *Infect. Immun.* **34:**234-240.

Galen, J. E., Ketley, J. H., Fasano, A., Richardson, S. H., Wasserman, S. S. and Kaper, J. B. 1992. Role of the *Vibrio cholerae* neuraminidase in the function of cholera toxin. *Infect. Immun.* **60:**406-415.

Ganfornina, M. D., Sanchez, D. and Bastiani, M. J. 1995. Lazarillo, a new GPI-linked surface lipocalin is restricted to a subset of neurons in the grasshopper embryo. *Development.* **121:**123-134.

Ganfornina, M. D., Gutierrez, G., Bastiani, M. and Sanchez, D. 2000. A phylogenetic analysis of the lipocalin protein family. *Mol. Biol. Evol.* **17:**114-126.

Gangarosa, E. J., Sanati, A., Saghare, H. and Feely, J. C. 1967. Multiple serotypes of *Vibrio cholerae* isolated from a case of cholera : evidence suggesting *in vivo* mutation. *Lancet.* **25:**646-648.

Garay, E., Arnau, A. and Amaro, C. 1985. Incidence of *Vibrio cholerae* and related vibrios in a coastal lagoon and seawater influenced by lake discharges along an annual cycle. *Appl. Environ. Microbiol.* **50:**426-430.

Garger, S. J., Griffith, O. M. and Gill, L. K. 1983. Rapid purification of plasmid DNA by a single centrifugation in a two step caesium chloride-ethidium bromide gradient. *Biochem. Biophys. Res. Commun.* **117:**835-842.

Gaskell, A., Crennell, S. and Taylor, G. 1995. The three domains of a bacterial sialidase : a β-propeller, an immunoglobulin model and a galactose-binding jelly-roll. *Stucture.* **3:**1197-1205.

Gay, P., Le Coq, D., Steinmetz, M., Ferrari, E. and Hoch, J. A. 1983. Cloning structural gene *sacB*, which encodes for exoenzyme levansucrase of *Bacillus subtilis* : expression of the gene in *Escherichia coli. J. Bacteriol.* **153:**1424-1431.

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T. and Kado, C. I. 1985. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *J. Bacteriol.* **164:**918-921.

Gennaro, M. L. and Greenaway, P. J. 1983. Nucleotide sequence within the cholera toxin operon. *Nucleic Acids Res.* **12:**3855-3861.

Geschwend, G., Feist, H. and Erler, W. 1997. Investigation of outer membrane proteins of *Pasteurella*. 2 : iron-regulated outer membrane proteins of *Pasteurella multocida* and *Pasteurella haemolytica. Berl. Tierarztl. Wochenschr.* **110:**386-390.

Gilson, E., Clement, J. M., Brutlag, D. and Hofnung, M. 1984. A family of dispersed repetitive extragenic palindromic DNA sequences in *Escherichia coli. EMBO J.* **3:**1417-1421.

Gilson, E., Bachellier, S., Perrin, S., Perrin, D., Grimont, P. A. D., Grimont, F. and Hofnung, M. 1990. Palindromic unit highly repetitive DNA sequences exhibit species specificity within *Enterobacteriaceae. Res. Microbiol.* **141:**1103-1116.

Gilson, E., Perrin, D. and Hofnung, M. 1990. DNA polymerase I and a protein complex bind specifically to *Escherichia coli* palindromic unit highly repetitive DNA:implications for bacterial chromosome organization. *Nucleic Acid Res.* **18:**3941-3952.

Gilson, E., Saurin, W., Perrin, D., Bachellier, S. and Hofnung, M. 1991a. The BIME family of bacterial highly repetitive sequences. *Res. Microbiol.* **142:**217-222.

Gilson, E., Saurin, W., Perrin, D., Bachellier, S. and Hofnung, M. 1991b. Palindromic units are part of a new bacterial interspersed mosaic element (BIME). *Nucleic Acids Res.* **19:**1375-1383.

Giron, J. A., Ho, A. S. and Schoolnik, G. K. 1991. An inducible bundle-forming pilus of enteropathogenic *Escherichia coli. Science.* **254:**710-713.

Giron, J. A., Levine, M. M. and Kaper, J. B. 1994. Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli. Mol. Microbiol.* **12:**71-82.

Godovac-Zimmermann, J. 1988. The structural motif of β-lacto-globulin and retinol binding protein : a basic framework for binding and transport of small hydrophobic molecules? *Trends Biochem. Sci.* **13** :64-66.

Goldberg, I. and Mekalanos, J. J. 1986. Effect of a *recA* mutation on cholera toxin gene amplification and deletion events. *J. Bacteriol.* **165:**723-731.

Goldberg, M. B., DiRita, V. J. and Calderwood, S. B. 1990. Identification of an iron-regulated virulence determinant in *Vibrio cholerae*, using Tn*phoA* mutagenesis. *Infect. Immun.* **58:**55-60.

Goldhar, J., Perry, R. and Ofek, I. 1984. Extraction and properties of non fimbrial mannose-resistant hemagglutinin from a urinary isolate of *Escherichia coli. Curr. Microbiol.* **11:**49-54.

Goldhar, J., Perry, R., Golecki, J. R., Hoschuzky, H., Jann, B. and Jann, K. 1987. Nonfimbrial, mannose-resistant ashesins from uropathogenic *Escherichia coli* 083 : K1 : H4 and 014 : K? : H11. *Infect. Immun.* **55:**1837-1842.

Gonzalez-Pastor, J. E., San-Millan, J. L., Castilla, M. A. and Moreno, F. 1995. Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7. *J. Bacteriol.* **177:**7131-7140.

Gravel, A., Fournier, B. and Roy, P. H. 1998a. DNA complexes obtained with the integron integrase IntI1 at the *attI*1 site. *Nucleic Acids Res.* **26:**4347-4355.

Gravel, A., Messier, N. and Roy, P. H. 1998b. Point mutation in the integron integrase IntI1 that affect recombination and/or substrate recognition. *J. Bacteriol.* **180:**5437-5442.

Grunstein, M. and Hogness, D. 1975. Colony hybridization : a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA.* **72:**3961-3965.

Guentzel, M. N. and Berry, L. J. 1975. Motility as a virulence determinant for *Vibrio cholerae*. *Infect. Immun.* **11:**890-897.

Guzman, M. G., Resik, S., Sariol, C., Rodriguez, R., Alvarez, M., Morier. L, Castillo, A., Kouri, G. and Mas., P. 1995. The physicochemical characterization of the strains isolated during the outbreak of epidemic neuropathy in Cuba. II. *Rev. Cubana. Med. Trop.* **47:**36-40.

Hacker, J., Blum-Oehler, G., Muhldorfer, I. and Tschape, H. 1997. Pathogenicity islands of virulent bacteria : structure, function and impact on microbial evolution. *Mol. Microbiol.* **23:**1089-1097.

Hall, R. M. and Vockler, C. 1987. The region of the IncN plasmid R46 coding for resistance to β-lactam antibiotics, streptomycin/spectinomycin and sulphonamides is closely related to antibiotic resistance segments found in IncW plasmids and in Tn*21*-like transposons. *Nucleic Acid Res.* **15:**7491-7501.

Hall, R. M., Brookes, D. E. and Stokes, H. W. 1991. Site-specific insertion of genes into integrons : role of the 59-base element and determination of the recombination cross-over point. *Mol. Microbiol.* **5:**1941-1959.

Hall, R. M., Brown, H. J., Brookes, D. E. and Stokes, H. W. 1994. Integrons found in different locations have identical 5′ ends but variable 3′ ends. *J. Bacteriol.* **176:**6286-6294.

Hall, R. M. and Collis, C. M. 1995. Mobile gene cassettes and integrons : capture and spread of genes by site-specific recombination. *Mol. Microbiol.* **15:**593-600.

Hall, R. M. 1995. Integrons : the movers and shakers of antibiotic resistance. *Today's Life Science.* 38-44.

Hall, R. M. 1998. The role of gene cassettes and integrons in the horizontal transfer of genes in Gram-negative bacteria. Pp.53-62. In, Syvanen M. and Kado, C. (eds.). *Horizontal gene transfer*. Chapman and Hall, London.

Hall, R. M. and Collis, C. M. 1998. Antibiotic resistance in Gram-negative bacteria : the role of gene cassettes and integrons. *Drug Resistance Updates*. **1:**109-119.

Hanne, L. F. and Finkelstein, R. A. 1982. Characterization and distribution of the hemagglutinins produced by *Vibrio cholerae*. *Infect. Immun.* **36:**209-214.

Hansson, K., Skold, O. and Sundstrom, L. 1997. Non-palindromic *attl* sites of integrons are capable of site-specific recombination with one another and with secondary targets. *Mol. Microbiol.* **26:**441-453.

Harlow, E. and Lane. D. 1988. Antibodies : A laboratory manual. Cold Spring Harbor Laboratory, New York, USA.

Hase, C. C. and Finkelstein, R. A. 1991. Cloning and nucleotide sequence of the *Vibrio cholerae* hemagglutinin/protease (HA/protease) gene and construction of an HA/ protease-negative strain. *J. Bacteriol.* **173:**3311-3317.

Hase, C. C. and Mekalanos, J. J. 1998. TcpP protein is a positive regulator of virulence gene expression in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. USA*. **95:**730-734.

Hayashi, M., Fujii, J. and Unemoto, T. 1997. Genetic defect of the sodium pump-defective mutant Nap-1 from the marine *Vibrio alginolyticus*. *Biochem. Mol. Biol. Int.* **41:**41-47.

Heildelberg, J. F., Eisen, J.A., Nelson, W. C., Clayton, R.A, Gwinn, M. L., Dodson, R.J, Haft, D. H., Hickey, E.K., Peterson, J. D., Umayam, L., Gill, S. R., Nelson, K.E, Read, T. D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Haiying, Q., Dragol, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nireman, W. C., White, O., Salzberg, S. L., Smith, H.O., Colwell, R. R., Mekalanos, J.J., Venter, J. C. and Fraser, C. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*. **406:**477-484.

Henderson, D. P. and Payne, S. M. 1993. Cloning and characterization of the *Vibrio cholerae* genes encoding the utilization of iron from haemin and haemoglobin. *Mol. Microbiol.* **7:**461-469.

Henderson, D. P. and Payne, S. M. 1994. *Vibrio cholerae* iron transport systems : role of heme and siderophore iron transport in virulence and identification of a gene associated with multiple iron transport system. *Infect. Immun.* **62:**5120-5125.

Herrington, D. A., Hall, R. H., Losonsky, G. A., Mekalanos, J. J., Taylor, R. K. and Levine, M. M. 1988. Toxin, toxin-coregulated pili, and the *toxR* regulon are essential for *Vibrio cholerae* pathogenesis in humans. *J. Exp. Med.* **168:**1487-1492.

Higgins, D. G. and Sharp, P. M. 1988. CLUSTAL : a package for performing multiple sequence alignment on a microcomputer. *Gene*. **73:**237-244.

Higgins, C. F., Mclaren, R. S. and Newbury, S. F. 1988. Repetitive extragenic palindromic sequences, mRNA stability and gene expression : evolution by gene conversion? - a review. *Gene*. **72:**3-14.

Higgins, D. G. and Sharp, P. M. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* **5:**151-153.

Higgins, D. E. and DiRita, V. J. 1996. Genetic analysis of the interaction between *Vibrio cholerae* transcription activator ToxR and ToxT promoter DNA. *J. Bacteriol.* **178:**1080-1087.

Hirono, I., Masuda, T. and Aoki, T. 1996. Cloning and detection of the hemolysin gene of *Vibrio anguillarum*. *Microb. Pathog*. **21:**173-182.

Holden, H. M., Rypniewski, W. R., Law, J. H. and Rayment, I. 1987. The molecular structure of insecticyanin from the tobacco hornworm *Manduca sexta* at 2.6 A° resolution. *EMBO J.* **6:** 1565-1570.

Holmgren, J., Lonnroth, I., Mansson, J. and Svennerholm, L. 1975. Interaction of cholera toxin and membrane GM1 ganglioside of small intestine. *Proc. Natl. Acad. Sci. USA*. **72:**2520-2524.

Holmgren, J. 1991. Mucosal immunity and vaccination. *FEMS Microbiol. Immunol.* **89:**1-10.

Holmgren, J. 1992. Pathogenesis. p. 199-208. In, Barua, D. and Greenough III, W.B. (eds.). *Cholera*. Plenum Medical Book, New York.

Holt, R. G., Abiko, Y., Saito, S., Smorawinska, M., Hansen, J. B. and Curtiss III., R. 1982. *Streptococcus* mutans genes that code for extracellular proteins in *Escherichia coli* K-12. *Infect. Immun*. **38:**147-156.

Holzfeind, P., Merschak, P., Rogatsch, H., Culig, Z., Feichtinger, H., Klocker, H. and Redl, B. 1996. Expression of the gene for tear lipocalin/von Ebner's gland protein in human prostate. *FEBS Lett*. **396:**95-98.

Honda, T. and Finkelstein, R. A. 1979. Purification and characterization of a haemolysin produced by *Vibrio cholerae* biotype El Tor : another toxic substance produced by cholera vibrios. *Infect. Immun*. **26:**1020-1027.

Honda, T., Nishibuchi, M., Miwatani, T. and Kaper, J. B. 1986. Demonstration of a plasmid-borne gene encoding a thermostable direct hemolysin in *Vibrio cholerae* non-O1 strains. *Appl. Environ. Microbiol*. **52:**1218-1220.

Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Flak, H. and Kayser, H. 1987. Molecular structure of the bilin binding protein (BBP) from *Pieris brassica* after refinement at 2.0 A° resolution. *J. Mol. Biol.* **198:**499-513.

Hulton, C. S. J., Higgins, C. F. and Sharp. P. M. 1991. ERIC sequences : a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. *Mol. Microbiol.* **5:**825-834.

Ikigai, H., Akatsuka, A., Tsujiyama, H., Nakae, T. and Shimamura, T. 1996. Mechanism of membrane damage by El Tor hemolysin of *Vibrio cholerae* O1. *Infect. Immun.* **64:**2968-2973.

Ikigai, H., Ono, T., Iwata, M., Nakae, T. and Shimamura, T. 1997. El Tor hemolysin of *Vibrio cholerae* O1 forms channels in planar lipid bilayer membranes. *FEMS Microbiol. Lett.* **150:**249-254.

Ikigai, H., Ono, T., Nakae, T., Otsuru, H. and Shimamura, T. 1999. Two forms of *Vibrio cholerae* O1 El Tor hemolysin derived from identical precursor protein. *Biochim. Biophys. Acta.* **1415:**297-305.

Inouye, M. and Inouye, S. 1991. msDNA and bacterial reverse transcriptase. *Annu. Rev. Microbiol.* **45:**163-186.

Iredell, J. R. and Manning, P. A. 1994. The toxin-co-regulated pilus of *Vibrio cholerae* O1: a model for type 4 pilus biogenesis? *Trends Microbiol.* **2:**187-192.

Jain, R., Rivera, M. C. and Lake, J. 1999. Horizontal gene transfer among genomes : the complexity hypothesis. *Proc. Natl. Acad. Sci. USA.* **96:**3801-3806.

Jenkin, C. R. and Rowley, D. 1959. Possible factors in the pathogenesis of cholera. *Br. J. Exp. Pathol.* **40:**474-482.

Johnson, J. A., Morris, J. G. and Kaper, J. B. 1993. Gene encoding zonula occludens toxin (*zot*) does not occur independently from cholera enterotoxin genes (*ctx*) in *Vibrio cholerae*. *J. Clin. Microbiol.* **31:**732-733.

Jones, A. H. 1997. Literature and medicine: Garcia Marquez' Love in the time of cholera. *Lancet.* **350:**1169-1172.

Jonson, G., Holmgren, J. and Svennerholm, A. M. 1991a. Epitope differences in toxin-coregulated pili produced by classical and El Tor *Vibrio cholerae* O1. *Microb. Pathog.* **11:**179-188.

Jonson, G., Holmgren, J. and Svennerholm, A. M. 1991b. Identification of a mannose-binding pilus on *Vibrio cholerae* El Tor. *Microb. Pathog.* **11:**433-441.

Jonson, G., Holmgren, J. and Svennerholm, A. M. 1992. Analysis of expression of toxin-coregulated pili in classical and El Tor *Vibrio cholerae* O1 *in vitro* and *in vivo*. *Infect. Immun.* **60:**4278-4284.

Jouravlena, E. A., McDonald, G. A., Marsh, J. W., Taylor, R. K., Boesman-Finkelstein, M. and Finkelstein, R. A. 1998. The *Vibrio cholerae* mannose-sensitive hemagglutinin is the receptor for a filamentous bacteriophage from *Vibrio cholerae* O139. *Infect. Immun.* **66:**2535-2539.

Kaasen, I., Falkenberg, P., Styrvold, O. B. and Strom, A. R. 1992. Molecular cloning and physical mapping of the *otsBA* genes, which encode the osmoregulatory trehalose pathway of *Escherichia coli*: evidence that transcription is activated by *katF* (AppR). *J. Bacteriol.* **174:**889-898.

Kaniga, K., Delor, I. and Cornelis, G. R. 1991. A wide-host-range vector for inproving reverse genetics in Gram-negative bacteria : inactivation of the *blaA* gene of *Yersinia enterocolitica*. *Gene*. **109:**137-141.

Kaper, J. B., Fasano, A. and Trucksis, M. 1994. Toxins of *Vibrio cholerae*. pp. 145-176. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.) *Vibrio cholerae and cholera : Molecular to global perspectives*. ASM Press, Washington DC.

Kaper, J. B., Morris, J. G. and Levine, M. M. 1995. Cholera. *Clin. Microbiol. Rev.* **8:**48-86.

Karaolis, D. K. R., Lan, R. and Reeves, P. 1995. The sixth and seven cholera pandemics are due to independent clones separately derived from environmental nontoxigenic, non-O1 *Vibrio cholerae*. *J. Bacteriol.* **177:**3191-3198.

Karaolis, D. K. R., Johnson, J. A., Bailey, C. C., Boedeker, E. C. and Kaper, J. B. 1998. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc. Natl. Acad. Sci. USA*. **95:**3134-3139.

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

Karasawa, T., Mihara, T., Kurazono, H., Nair, G. B., Garg, S., Ramamurthy, T. and Takeda, Y. 1993. Distribution of the *zot* (zonula occludens toxin) gene among strains of *Vibrio cholerae* O1 and non-O1. *FEMS Microbiol. Lett.* **106:**143-145.

Kaufman, M. R. and Taylor, R. K. 1994. The toxin-coregulated pilus : Biogenesis and function. pp. 182-202. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives*. ASM Press, Washington DC.

Kay, B. A., Bopp, C. A. and Wells, J. G. 1994. Isolation and identification of *Vibrio cholerae* O1 from fecal specimens. pp. 3-25. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø.

(eds.). *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Kay, D., Fleisher, J. M., Salmon, R. L., Jones, F., Wyer, M. D., Godfree, A. F., Zelenauch-Jacuotte, Z. and Shore, R. 1994. Predicting likelihood of gastrenteritis from sea bathing : results from randomised exposure. *Lancet.* **344:**905-909.

Kaysner, C. A. and W. E. Hill. 1994. Toxigeinic *Vibrio cholerae* O1 in food and water. pp. 27-39. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Kazama, H., Hamashima, H., Sasatsu, M. and Arai, T. 1998. Distribution of the antiseptic-resitance gene *qacE* delta 1 in Gram-positive bacteria. *FEMS Microbiol. Lett.* **165:**295-299.

Keen, J. N., Caceres, I., Eliopoulos, E. E., Zagalsky, P. F. and Findley, J. B. C. 1991. Complete sequence and model for the A2 subunit of the carotenoid pigment complex, crustacyanin. *Eur J. Biochem.* **197:**407-417.

Kelly, J. T. and Parker, C. D. 1981. Identification and preliminary characterization of *Vibrio cholerae* outer membrane proteins. *J. Bacteriol.* **145:**1018-1024.

Kimsey, H. H. and Waldor, M. K. 1998. *Vibrio cholerae* haemagglutinin/protease inactivates CTXϕ. *Infect. Immun.* **66:**4025-4029.

King, C. A. and van Heyningen, W. E. 1973. Deactivation of cholera toxin by a sialidase-resistant monosialosylganglioside. *J. Infect. Dis.* **127:**639-647.

Kolvin, J. L. and Roberts, D. 1982. Studies on the growth of *Vibrio cholerae* biotype El Tor and biotype classical in foods. *J. Hyg. (Cambridge).* **89:**243-252.

Koonin, E. V. 1992. The second cholera toxin, Zot, and its plasmid-encoded and phage-encoded homologous constitute a group of putative ATPases with an altered purine NTP-binding motif. *FEBS Lett.* **312:**3-6.

Kovach, M. E., Phillips, R. W., Elzer II, P. H., Roop, R. M. and Peterson, K. M. 1994. pBBR1MCS: a broad-host-range cloning vector. *Biotechniques.* **16:**800-802.

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M. and Peterson. K. M. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance. *Gene.* **166:**175-176.

Kovach, M. E., Shaffer, M. D. and Peterson, K. M. 1996. A putative integrase gene defines the distal end of a large cluster of ToxR-regulated colonization genes in *Vibrio cholerae. Microbiol.* **142:**2165-2174.

Kozicki, M., Steffen, R. and Schar, M. 1985. "Boil it, cook it, peel it, or forget it" : does this rule prevent travelers diarrhoea? *Int. J. Epidemiol.* **14:**169-172.

Lan, R. and Reeves, P. R. 1998. Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of *Vibrio cholerae*. *Microbiol.* **144:**1213-1221.

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

Lawrence, J. G. and Ochman, H. 1997. A melioration of bacterial genomes : rates of change and exchange. *J. Mol. Evol.* **44:**383-397.

Lee, S. H., Hava, D. L., Waldor, M. K. and Camilli, A. 1999. Regulation and temporal expression patterns of *Vibrio cholerae* virulence genes during infection. *Cell.* **99:**625-634.

Lévesqué, C., Brassard, S., Lapointe, J. and Roy, P. H. 1994. Diversity and relative strength of tandem promoters for the antibiotic-resistance genes of several integrons. *Gene.* **142:**49-54.

Lévesqué, G., Lamarche, B., Murthy, M., Julien, P., Despres, J. P. and Deshaies, Y. 1994. Determination of changes in specific gene expression by reverse transcription PCR using interspecies mRNAs as internal standards. *Biotechniques.* **17.:**738-741.

Levin, B. R. and Tauxe, R. V. 1996. Cholera : Nice bacteria and bad viruses. *Curr Biol.* **6:**1389-1391.

Li, W. and Riddiford, L. M. 1992. Two distinct genes encode two major isoelectric forms of insecticyanin in the tobacco hornworm, *Manduca sexta*. *Eur. J. Biochem.* **205:**491-499.

Lorenzo, de., Herrero, M., Jakubzik, U. and Timmis, K. 1990. Mini-Tris transposon derivatives for insertional mutagenesis, promoter probing and chromosomal insertion of cloned DNA in Gram-negative DNA. *J. Bacteriol.* **172:**6568-6572.

Lugtenberg, B., Meijers, J., Peters, R., van der Hoek, P. and van Alphen, L. 1975. Electrophoretic resolution of the major outer membrane protein of *Escherichia coli* K-12 into four bands. *FEBS Lett.* **58:**254-258.

Lupski, J. R. and Weinstock, G. M. 1992. Short, interspersed repetitive DNA sequences in prokaryotic genomes. *J. Bacteriol.* **174:**4525-4529.

MacNamara, C. 1876. A history of Asiatic cholera. MacMillan and Co., London.

Maniatis, T., Fritsch, F. F. and Sambrook, J. 1982. Molecular cloning : a laboratory mannual. Cold Spring Harbor Laboratories, Cold Spring Harbor , New York USA.

Manning, P. A., Heuzenroeder, M. W., Yeadon, J., Leavesley, D. I., Reeves, P. R. and Rowley, D. 1986. Molecular cloning and expression in *Escherichia coli* K-12 of the O-antigens of the Inaba and Ogawa serotypes of *Vibrio cholerae* and development. *Infect. Immun.* **53:**272-277.

Manning, P. A. 1992. Molecular design of cholera vaccines. *Vaccine.* **10:**1015-1021.

Manning, P. A. 1994. Surface-associated and soluble components of *Vibrio cholerae* involved in bacteria-host interactions, Vol. 12. In, Dangl. J. L (ed.), *Current topics in Microbiology and Immunity.* Springer-Verlag, Germany.

Manning, P. A. 1995. A beginning to cell surface studies in *Vibrio cholerae* : a personal perspective. *Bull. Inst. Pasteur.* **93:**263-268.

Manning, P. A. 1997. The *tcp* gene cluster of *Vibrio cholerae*. *Gene.* **192:**63-70.

Manning, P. A., Clark, C. A. and Focareta, T. 1999. Gene capture in *Vibrio cholerae*. *Trends Microbiol.* **7:**93-95.

Marinaro, M., Di-Tommaso, A., Uzzau, S., Fasano, A. and De Magistris. M. T. 1999. Zonula occludens toxin is a powerful mucosal adjuvant for intranasally delivered antigens. *Infect. Immun.* **67:**1287-1291.

Marsh, J. W. and Taylor, R. K. 1999. Genetic and transcriptional analyses of the *Vibrio cholerae* mannose-sensitive hemagglutinin type4 pilus gene locus. *J. Bacteriol.* **181:**1110-1117.

Martin, J. L., Bardwell, J. C. A. and Kuriyan, J. 1993. Crystal structure of the DsbA protein required for disulphide bond formation *in vivo*. *Nature* (London). **365:**464-468.

Martinez, E. and de la Cruz, F. 1988. Transposon Tn*21* encodes a RecA-independent site-specific integration system. *Mol. Gen. Genet.* **211:**320-325.

Martinez, E. and de la Cruz, F. 1990. Genetic elements involved in Tn*21* site-specific integration, a novel mechanism for the dissemination of antibiotic-resistance genes. *EMBO J.* **9:**1275-1281.

Mata, L. 1994. Cholera El Tor in Latin America, 1991-1993. *Ann. N. Y. Acad. Sci.* **15:**55-68.

Matsushita, S., Noguchi, Y., Yanagawa, Y., Igarashi, H., Morita, K., Wada, H., Watanabe, N., Shibata, M., Kanamori, M., Kudoh, Y. and Shimada, T. 1997. Sporadic diarrhea caused by *Vibrio cholerae* non-O1 and characteristics of the isolates. *Kansenshogaku Zasshi.* **71:**1204-1209.

Mazel, D., Dychinco, B., Webb, V. A. and Davies, J. 1998. A distinctive class of integron in the *Vibrio cholerae* genome. *Science.* **280:**605-608.

Mekalanos, J. J. 1983. Duplication and amplification of toxin genes in *Vibrio cholerae*. *Cell.* **35:**253-263.

Mekalanos, J. J. 1985. Cholera toxin genetic analysis, regulation, and role in pathogenesis. *Curr. Top. Microbiol. Immunol.* **118:**97-118.

Mekalanos, J. J. 1988. Production and purification of cholera toxin. *Methods Enzymol.* **165:**169-175.

Mekalanos, J. J., Rubin, E. J. and Waldor, M. K. 1997. Cholera: molecular basis for emergence and pathogenesis. *FEMS Immunol. Med. Microbiol.* **18:**241-248.

Miller, J. H. 1972. Experiments in Molecular Genetics. Cold Spring Harbour Laboratories, Cold Spring Harbor, NY, USA.

Miller, V. L. and Mekalanos, J. J. 1985. Genetic analysis ofthe cholera toxin-positive regulatory gene *toxR*. *J. Bacteriol.* **163:**580-585.

Miller, V. L. and Mekalanos, J. J. 1988. A novel suicide vector and its use in construction of insertion mutations : osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J. Bacteriol.* **170:**2575-2583.

Mintz, E. D., Popovic, T. and Blake, P. A. 1994. Transmission of *Vibrio cholerae* O1. pp. 345-356. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives*. ASM Press, Washington DC.

Miyada, C. G., Stolzfus, L. and Wilcox, G. 1984. Regulation of the *araC* gene of *Escherichia coli* : catabolite repression, autoregulation and effect on *araBAD* expression. *Proc. Natl. Acad. Sci. USA*. **81:**4120-4124.

Molitoris, E., Marii, M. A., Joseph, S. W., Krichevsky, M. I., Fanning, G. R., Last, G., El-Mishad, A. M., El Batawi, Y. A. and Colwell, R. R. 1989. Numericaltaxonomy and deoxyribonucleic acid relatedness of environmental and clinical *Vibrio* species isolated in Indonesia. *Int. J. Syst. Bacteriol.* **39:**442-449.

Morris, J. G., West, G. R., Holck, S. E., Blake, P. A., Echeverria, P. D. and Karaulnik, M. 1982. Cholera among refugees in Rangsit, Thailand. *J. Infect. Dis.* **145:**131-133.

Murley, Y. M., Carroll, P. A., Skorupski, K., Taylor, R. K. and Calderwood, S.B. Differential transcription of the *tcpPH* operon confers biotype-specific control of the *Vibrio cholerae* ToxR virulence regulon. *Infect. Immun.* **67:**5117-5123.

Naas, T., Sougakoff, W., Casetta, A. and Nordmann, P. 1998. Molecular characterization of OXA-20, a novel class D β-lactamase, and its integron from *Pseudomonas aeruginosa*. *Antimicrob. Agents. Chemother.* **42:**2074-2083.

Naas, T., Poirel, L., Karim, A. and Nordmann, P. 1999. Molecular characterization of In50, a class 1 integron encoding the gene for the extended-spectrum β-lactamase VEB-1 in *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* **176:**411-419.

Nagamune, K., Yamamoto, K., Naka, A., Matsuyama, J., Miwatani, T. and Honda, T. 1996. *In vitro* proteolytic processing and activation of the recombinant precursor of El Tor cytolysin/hemolysin (pro-HlyA) of *Vibrio cholerae* by soluble hemagglutinin/ protease of *Vibrio cholerae*, trypsin, and other protease. *Infect. Immun.* **64:**4655-4658.

Nagamune, K., Yamamoto, K. and Honda, T. 1997. Intramolecular chaperone activity of the pro-region of *Vibrio cholerae* El Tor hemolysin. *J. Biol. Chem.* **272:**1338-1343.

Nair, G. B., Bhattacharya, S. K. and Deb, B. C. 1994. *Vibrio cholerae* O139 in Bengal: the eighth pandemic strain of cholera. *Indian J. Public. Health.* **38:**33-36.

Nair, G. B., Ramamurthy, T., Bhattacharya, S. K., Mukhopadhyay, A. K., Garg, S., Bhattacharya, M. K., Takeda, T., Shimada, T., Takeda, Y. and Deb, B. C. 1994. Spread of *Vibrio cholerae* O139 Bengal in India. *J. Infect. Dis.* **169:**1029-1034.

Nalin, D., Daya, V., Reid, A., Levine, M. M. and Cisneros, L. 1979. Adsorption and growth of *Vibrio cholerae* on chitin. *Infect. Immun.* **25:**768-770.

Nesvera, J., Hochmannova, J. and Patek, M. 1998. An integron of class 1 is present on the plasmid pCG4 from gram-positive bacterium *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **169:**391-395.

Newcomer, M. E. and Ong, D. E. 1990. Purification and crytallyzation of a retinol acid-binding protein from rat epididymis. *J. Biol. Chem.* **265:**12876-12879.

Nicaud, J. M., Mackman, N., Gray, L. and Holland, I. B. 1985. Characterization of HlyC and mechanism of action and secretion of haemolysin from *Escherichia coli* 2001. *FEBS Lett.* **187:**339-344.

Nikaido, H. 1993. Transport across the bacterial outer membrane. *J. Bioenerg. Biomemb.* **25:**581-589.

Nishibuchi, M., Ishibashi, M., Takeda, Y. and Kaper, J. B. 1985. Detection of the thermostable direct haemolysin gene and related DNA sequences in *Vibrio parahaemolyticus* and other *Vibrio* species by the DNA colony hybridization test. *Infect. Immun.* **49:**481-486.

Nishibuchi, M., Fasano, A., Russell, R. G. and Kaper, J. B. 1992. Enterotoxigenicity of *Vibrio parahaemolyticus* with and without genes encoding thermostable direct hemolysin. *Infect. Immun.* **60:**3539-3545.

Nobechi, K. 1923. Immunological studies upon types of *Vibrio cholerae*. *Sci. Rep. Inst. Infect. Dis. Tokyo Univ.* **2:**43.

Norqvist, A. and Wolf-Watz, H. 1993. Characterization of a novel chromosomal virulence locus involved in expression of a major surface flagellar sheath antigen of the fish pathogen *Vibrio anguillarum*. *Infect. Immun.* **61:**2434-2444.

Ochman, H. and Wilson, A. C. 1987. Evolution in bacteria : evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* **26:**74-86.

Ochman, H. and Lawrence, J. G. 1996. Phylogenetics and the amelioration of bacterial genomes. In, Neidhardt, F. C., Curtiss III, R, Ingraham, J.L, Lin, E.C.C, Low, K.B, Magasanik, B, Reznikoff, W.S, Riley, M, Schaechter, M, Umbarger, H.E (eds.). *Escherichia coli* and *Salmonella* : *cellular and molecular biology*. 2nd ed. American Society for Microbiology, Washington, DC.

Ogawa, A. and Takeda, T. 1993. The gene encoding the heat-stable enterotoxin of *Vibrio cholerae* is flanked by 123-base pair direct repeats. *Microbiol. Immunol.* **37:**607-616.

Ogierman, M. A. and Manning, P. A. 1992. TCP pilus biosynthesis in *Vibrio cholerae* O1: gene sequence of *tcpC* encoding an outer membrane lipoprotein. *FEMS Microbiol. Lett.* **76:**179-184.

Ogierman, M. A. 1998. Molecular characterization of the haemolysin (HlyA) and the region downstream of *hlyA* in *Vibrio cholerae* O1. Ph. D. Thesis. Department of Microbiology and Immunology, The University of Adelaide.

Ølsvik, Ø., Popovic, T. and Fields, P. I. 1993. Polymerase chain reaction for detection of toxin genes in strains of *Vibrio cholerae* O1. pp. 573-583. In, White T.J. (ed.). *Diagnostic Molecular Microbiology*. Americann Society for Microbiology, Washington, D.C.

Ølsvik, Ø., Wahlberg, J., Petterson, B., Unlen, M., Popovic, T., Wachsmuth, I. K. and Field, P. I. 1993. Use of automated sequencing of PCR-generated amplicons to identify three types of cholera toxin subunit B in *Vibrio cholerae* O1 strains. *J. Clin Microbiol.* **31:**22-25.

Organization, WHO. 1993. Guidelines for cholera control. World Health Organization, Geneva, Switzerland.

Organization, WHO. 1995. Guidelines for cholera control. World Health Organization, Geneva, Switzerland.

Ørskov, I. and F. Ørskov. 1990. Serologic classification of fimbriae. pp. 71-90, Vol. 151. In, Jann, K. and Jann, B.K. (eds.). *Current topics in Microbiology and Immunity*. Springer-Verlag, Germany.

Osano, E., Arakawa, Y., Wacharotayankun, R., Ohta, M., Horii, T., Ito, H., Yoshimura, F. and Kato, N. 1994. Molecular characteriation of an enterobacterial metallo β-lactamase found in a clinical isolate of *Serratia marcescens* that shows imipenem resistance. *Antimicrob. Agents. Chemother.* **38:**71-78.

Osborn, M. J., Gander, J. E., Parisi, E. and Carson. J. 1972. Mechanism of assembly of the outer membrane of *Salmonella typhimurium. J. Bio. Chem.* **247:**3962-3971.

Osek, J., Jonson, G., Svennerholm, A.-M. and Holmgren, J. 1994. Role of antibodies against biotype-specific *Vibrio cholerae* pili in protection against experimental classical and El Tor cholera. *Infect. Immun.* **62:**2901-2907.

Ottemann, K. M., DiRita, V. J. and Mekalanos, J. J. 1992. ToxR proteins with substitutions in residues conserved with OmpR fail to activate transcription from the cholera toxin promoter. *J. Bacteriol.* **174:**6807-6814.

Ottemann, K. M. and Mekalanos, J. J. 1995. Analysis of *Vibrio cholerae* ToxR function by construction of novel fusion proteins. *Mol. Microbiol.* **15:**719-731.

Otto, B. R., Verweij-van Vught, A. M. J. J. and MacLaren, D. M. 1992. Transferrins and heme-compounds as iron sources for pathogenic bacteria. *Crit. Rev. Microbiol.* **18:**217-233.

Ouellette, M., Bissonnette, L. and Roy, P. H. 1987. Precise insertion of antibiotic resistance determinants into Tn*21*-like transposons : nucleotide sequence of the OXA-I β-lactamase gene. *Proc. Natl. Acad. Sci.* USA. **84:**7378-7382.

Ouellette, M. and Roy. P. H. 1987. Homology of ORFs from Tn*2603* and from R46 to site-specific recombinases. *Nucleic Acid Res.* **15:**10055.

Parsot, C., Taxman, E. and Mekalanos, J. J. 1991. ToxR regulates the production of lipoproteins and the expression of serum resistance in *Vibrio cholerae. Proc. Natl. Acad. Sci. USA.* **88:**1641-1645.

Patel, K. V., Sheth, H. G. and Schrey, M. P. 1997. Stimulation or endothelin-1 secretion by human breast cancer cells through protein kinase A activation: a possible novel paracrine loop involving breast fibroblast-derived prostaglandin E2. *Mol. Cell. Endocrinol.* **126:**143-151.

Payne, S. M. and Finkelstein, R. A. 1978. Siderophore production by *Vibrio cholerae. Infect. Immun.* **20:**310-311.

Pearson, G. D. N., DiRita, V. J., Goldberg, M. B., Boyko, S. A., Calderwood, S. B. and Mekalanos, J. J. 1990. New attenuated derivatives of *Vibrio cholerae. Res. Microbiol.* **141:**893-899.

Pearson, D. G. N., Woods, A., Chang, S. L. and Mekalanos, J. J. 1993. CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. *Proc. Natl. Acad. Sci. USA.* **90:**3750-3754.

Peitsch, M. C. and Boguski, M. S. 1990. Is apolipoprotein D a mammalian bilin-binding protein? *Nature New Biol.* **2:**197-206.

Peitsch, M. C., Amiguet, P., Guy, R., Brunner, J., Maizel, J. V. J. and Tschopp, J. 1990. Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin. *Mol. Immunol.* **27:**589-602.

Pemberton, J. M., Kidd, S. P. and Schmidt, R. 1997. Secreted enzymes of *Aeromonas*. *FEMS Microbiol. Lett.* **152:**1-10.

Petersen, A., Guardabassi, L., Dalsgaard, A. and Olsen, J. E. 2000. Class 1 integrons containing a *dhfrI* trimethoprim resistance gene cassette in aquatic *Acinetobacter spp. FEMS Microbiol. Lett.* **182:**73-76.

Peterson, K. M. and Mekalanos, J. J. 1988. Characterization of the *Vibrio cholerae* ToxR regulon : Identification of novel genes involved in intestinal colonization. *Infect. Immun.* **56:**2822-2829.

Pollitzer, R. 1959. Cholera. World Health Organization, Geneva, Switzerland.

Popovic, T., Fields, P. I. and Ølsvik, Ø. 1994. Detection of cholera toxin genes. pp 41-52. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives*. ASM Press, Washington DC.

Purins, L. 1997. Characterization of the VCR region in *Vibrio cholerae*. Honours thesis. Department of Microbiology and Immunology, The University of Adelaide.

Rabbani, G. H. 1986. Cholera. *Clin. Gastroenterol.* **15:**507-528.

Rabbani, G. H. and Greenough III, W. B. 1992. Pathophysiology and clinical aspects of cholera. pp. 209-228. In, Barua, D. and W. B. Greenough III. (eds.). *Cholera.* Plenum Medical Book, New York, USA.

Radstrom, P., Skold, O., Swedberg, G., Flensburg, J. and Roy, P. H. 1994. Transposon Tn*5090* of plasmid R751, which carries an integron, is related to Tn*7*, Mu, and the retroelements. *J. Bacteriol.* **176:**3257-3268.

Ramamurthy, T., Gang, S., Sharma, R., Bhattacharya, S. K., Nair, G. B., Shimada, T., Takeda, T., Karasawa, T., Kurazano, H., Pal, A. and Takeda, Y. 1993. Emergence of novel strain of *Vibrio cholerae* with epidemic potential in southern and eastern India. *Lancet.* **341:**703-704.

Recchia, G. D., Stokes, H. W. and Hall, R. M. 1994. Characterization of specific and secondary recombination sites recognized by the integron DNA integrase. *Nucleic Acid Res.* **22:**2071-2078.

Recchia, G. D. and Hall, R. M. 1995a. Plasmid evolution by acquisition of mobile gene cassettes : plasmid pIE723 contains the *aadB* gene cassette precisely inserted at a secondary site in the IncQ plasmid RSF1010. *Mol. Microbiol.* **15:**179-187.

Recchia, G. D. and R. M. Hall. 1995b. Gene cassettes : a new class of mobile element. *Mol. Microbiol.* **141:**3015-3027.

Recchia, G. D. and Hall, R. M. 1997. Origins of the mobile gene cassettes found in integrons. *Trends in Microbiol.* **5:**389-394.

Redmond, J. W., Korsch, M. J. and Jackson, G. D. 1973. Immunochemical studies to the O-antigens of *Vibrio cholerae* : Partial characterization of an acid-labile antigenic determinant. *Aust J. Exp. Biol. Med. Sci.* **51:**229-235.

Reed, L. J. and Muench, H. 1938. A simple method for estimating fifty percent end points. *Am J. Hyg.* **27:**493-497.

Richardson, K. 1991. Roles of motility and flagellar structure in patrhogenecity of *Vibrio cholerae* : analysis of motility mutrants in three animal models. *Infect. Immun.* **59:**2727-2736.
Richardson, S. H. 1994. Animal models in cholera research. pp. 203-226. In Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Richardson, S. H. 1994. Host susceptibility. pp. 273-289. In, Wachsmuth, I. K, Blake, P.A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Roberts, D. 1992. Growth and survival of *Vibrio cholerae* in foods. *Microbiol. Dig.* **9:**24-31.

Rodrigue, D. C., Popovic, T. and Wachsmuth, I. K. 1994. Nontoxigenic *Vibrio cholerae* O1 infections in the United States. pp. 69-76. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Rosser, S. J. and Young, H. K. 1999. Identification and characterization of class 1 integrons in bacteria from an aquatic environment. *J. Antimicrob. Chemother.* **44:**11-18.

Rowe-Magnus, D. A. and Mazel, D. 1999. Resistance gene capture. *Curr. Opin. Microbiol.* **2:**481-488.

Rowe-Magnus, D. A., Guerout, A.-M. and Mazel, D. 1999. Super-integrons. *Res. Microbiol.* **150:**641-651.

Russell, C. B., Stewart, R. C. and Dahlquist, F. W. 1989. Control of transducer methylation levels in *Escherichia coli* : investigation of components essential for modulation of methylation and demethylation reactions. *J. Bacteriol.* **171:**3609-3618.

Sack, R. B. 1992. Colonization and pathology. pp 189-197. In, Barua, D. and Greenough, W. B. III (eds.). *Cholera.* Plenum Medical Book, New York, USA.

Saito, H. and Shimoda, M. 1997. Insecticyanin of *Agrius convolvuli*: purification and characterization of the biliverdin-binding protein from the larval hemolymph. *Zoolog. Sci.* **14:**777-783.

Saito, H. 1998. Purification and characterization of two insecticyanin-type proteins from the larval hemolymph of the Eri-silkworm, *Samia cynthia ricini*. *Biochim. Biophys. Acta.* **1380:**141-150.

Sakazaki, R. and Tamura, K. 1971. Somatic antigen variation in *Vibrio cholerae*. *Jpn. J. Med. Sci. Biol.* **24:**93-100.

Sakazaki, R. 1992. Bacteriology of *Vibrio* and related organisms. pp 37-56. In, Barua, D. and Greenough, W.B. III (eds.). *Cholera*. Plenum Medical Book, New York, USA.

Salyers, A. A. and Whitt, D. D. 1994. Bacterial pathogenesis : A molecular approach. ASM Press, Washington DC.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular cloning : A laboratory manual. Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, USA.

Sanchez, D., Ganfornina, M. D. and Bastiani, M. J. 1995. Developmental expression of the lipocalin Lazarillo and its role in axogonal pathfinding in the grasshopper embryo. *Development.* **121:**135-147.

Sanger, F., Nicklen, S. and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74:**5463-5467.

Sansom, C. E., North, A. C. T. and Sawyer, L. 1994. Structural analysis and classification of lipocalins and related proteins using a profile-search method. *Biochim. Biophys. Acta.* **1208:**247-255.

Sanyal, S. C. 1992. Epidemiology and pathogenicity of non-O1 *Vibrio* species and related organisms, p. 57-65. In, Barua, D. and Greenough, W.B. III (eds.). *Cholera*. Plenum Medical Book Company, New York.

Sathyamoorthy, V., Huntley, J. S., Hall, A. C. and Hall, R. H. 1997. Biochemical and physiological characteristics of HlyA, a pore-forming cytolysin of *Vibrio cholerae* serogroup O1. *Toxicon.* **35:**515-527.

Sawaddiwudhipong, W., Akarasewi, P., Chayaniyayodhin, T., Kunasol, P. and Foy, H. M. 1990. A cholera outbreak associated with eating uncooked pork in Thailand. *J. Diarrhoeal Dis. Res.* **8:**94-96.

Schmidt, F. and Klopfer-Kaul, I. 1984. Characterization of specific and secondary recombination sites recognised by the integron DNA integrase. *Nucleic Acid Res.* **22:**2071-2078.

Schoenberg, B. E. 1974. Snow on the water of London. *Mayo Clin. Proc.* **49:**680-684.

Schuhmacher, D. A. and Klose, K. E. 1999. Environmental signals modulate ToxR-dependent virulence factor expression in *Vibrio cholerae*. *J. Bacteriol.* **181:**1508-1514.

Sciortino, C. V. and Finkelstein, R. A. 1983. *Vibrio cholerae* expresses iron-regulated outer membrane proteins *in vivo*. *Infect. Immun.* **42:**990-996.

Sciortino, C. V., Yang, Z. and Finkelstein, R. A. 1985. Monoclonal antibodies to outer membrane antigens of *Vibrio cholerae*. *Infect. Immun.* **49:**122-131.

Sciortino, C. 1996. Protection against mortality due to *Vibrio cholerae* infection in infant rabbits caused by immunization of mothers with cholera protective antigen. *J. Diarrhoeal Dis. Res.* **14:**16-26.

Sechi, L. A., Zanetti, S., Dupre, I. Delogu, G. and Fadda, G. 1998. Enterobacterial repetitive intergenic consensus sequences as molecular target for typing of *Mycobacterium tuberculosis* strains. *J. Clin. Microbiol.* **36:**128-132.

Senda, K., Arakawa, Y., Ichiyama, S., Nakashima, K., Ito, H., Ohsuka, S., Shimokata, K., Kato, N. and Ohta, M. 1996. PCR detection of metallo-β-lactamase gene (bla_{IMP}) in Gram-negative rods resistant to broad-spectrum β-lactams. *J. Clin. Microbiol.* **34:**2909-2913.

Sharma, D. P., Stroeher, U. H., Thomas, C. J., Manning, P. A. and Attridge, S. R. 1989. The toxin-coregulated pilus TCP of *Vibrio cholerae* : molecular cloning of the genes involved in pilus biosynthesis and evaluation of TCP as a protective antigen in the infant mouse model. *Microb. Pathog.* **7:**.

Sharma, C., Maiti, S., Mukhopadhyay, A. K., Basu, A., Basu, I., Nair, G. B., Mukhopadhyay, R., Das, B., Kar, S., Ghosh, R. K. and Ghosh, A. 1997a. Unique organization of the CTX genetic elements in *Vibrio cholerae* O139 strains which reemerged in Calcata, India, in September, 1996. *J. Clin. Microbiol.* **35:**3348-3350.

Sharma, C., Nair, G. B., Mukhopadhyay, A. K., Bhattacharya, S. K., Ghosh, R. K. and Ghosh, A. 1997b. Molecular characterization of *Vibrio cholerae* O1 biotype El Tor strains Isolated between 1992 and 1995 in Calcutta, India : Evidence for the emergence of a new clone of the El tor. *J. Infec. Dis.* **175:**1134-1141.

Sharples, G. T. and Lloyd, R. G. 1990. A novel repeated DNA sequence located in the intergenic region of bacterial chromosomes. *Nucleic Acids Res.* **18:** 6503-6508.

Shaw, C. E., Peterson, K., Sun, M, D., Mekalanos, J. J. and Taylor, R. K. 1988. TCP pilus expression and biogenesis by classical and El Tor biotypes of *Vibrio cholerae* O1, pp. 23-35. In, Switalski, Hook and Beachey (eds.). *Molecular mechanism of microbial adhesion.* Springer-Verlag, New York.

Sheehy, T. W., Sprintz, H., Augerson, W. S. and Formal, S. B. 1966. Laboratory *Vibrio cholerae* infection in the United States. *J. Am. Med. Assoc.* **197:**321-325.

Shimada, T. 1980. A new selective isolation medium for *Vibrio cholerae*, El Tor-PMT agar. *Media Circle.* **25:**6-9.

Shukla, B. N., Singh, D. V. and Sanyal, S. C. 1995. Attachment of non-culturable toxigenic *Vibrio cholerae* O1 and non-O1 and *Aeromonas spp.* to the aquatic arthropod Gerris spinolae and plants in the River Ganga, Varanasi. *FEMS Immunol. Med. Microbiol.* **12:**113-120.

Sigel, S. P. and Payne, S. M. 1982. Effect of iron limitation on growth, siderophore production, and expression of outer membrane proteins of *Vibrio cholerae. J. Bacteriol.* **150:**148-155.

Simon, R., Hote, B., Klauke, B. and Kosier, B. 1991. Isolation and characterization of insertion sequence elements from Gram-negative bacteria by using new broad-host-range, positive selection vectors. *J. Bacteriol.* **173:**.

Skorupski, K. and Taylor, R. K. 1997. Control of the ToxR virulence regulon in *Vibrio cholerae* by environmental stimuli. *Mol. Microbiol.* **25:**1003-1009.

Southern, E. M. 1975. Detection of specific sequence among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **97:**503-517.

Sperandio, V., Giron, J. A., Silveira, W. D. and Kaper, J. B. 1995. The OmpU outer membrane protein, a potential adherence factor of *Vibrio cholerae. Infect. Immun.* **63:**4433-4438.

Spreyer, P., Schaal, H., Kuhn, G., Rothe, T., Unterbeck, A., Ølek, K. and Muller. H.W. 1990. Regeneration-associated high level expression of apolipoprotein D mRNA in endoneutral fibroblasts of peripheral nerve. *EMBO J.* **9:**2479-2484.

St. Louis, M. E., Porter, T. D., Helal, A., Drame, K., Hargrett-Bean, N., Wells, J. G. and Tauxe, R. V. 1990. Epidemic cholera in West Africa : the role of food handling and high-risk foods. *Am. J. Epidemiol.* **131:**719-727.

Stern, M. J., Ames, G. F.-L., Smith, N. H., Robinson, E. C. and Higgins, C. F. 1984. Repetitive extragenic palindromic sequences : a major component of the bacterial genome. *Cell.* **37:**1015-1026.

Stoebner, J. A. and Payne, S. M. 1988. Iron-regulated hemolysin production and utilization of heme and hemoglobin by *Vibrio cholerae. Infect. Immun.* **56:**2891-2895.

Stokes, H. W. and Hall, R. M. 1989. A novel family of potentially mobile DNA elements encoding site-specific gene-integration functions : integrons. *Mol. Microbiol.* **3:**1669-1683.

Stokes, H. W., Gorman, D. B. O., Recchia, G. D., Parsekhian, M. and Hall, R. M. 1997. Structure and function of 59-base element recombination sites associated with mobile gene cassettes. *Mol. Microbiol.* **26:**731-745.

Stroeher, U. H., Karageorgos, L. E., Morona, R. and Manning, P. A. 1992. Serotype conversion in *Vibrio cholerae* O1. *Proc. Natl. Acad. Sci. USA.* **89:**2566-2570.

Stroeher, U. H., Jedani, K. E. and Manning, P. A. 1998. Genetic organization of the regions asociated with surface polysaccharide synthesis in *Vibrio cholerae* O1, O139 and *Vibrio anguillarum* O1 and O2 : a review. *Gene.* **223:**269-282.

Studier, F. W. and Moffant, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189:**113-130.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. 1990. Use of T7 RNA polymerase to direct the expression of cloned genes. *Methods Enzymol.* **185:**60-89.

Stugard, C. E., Daskaleros, P. A. and Payne, S. M. 1989. A 101-kilodalton haem-binding protein associated with Congo red binding and virulence in *Shigella flexneri* and enteroinvasive *Escherichia coli* strains. *Infect. Immun.* **57:**3534-3539.

Sun, F., Mekalanos, J. J. and Taylor, R. K. 1990. Antibodies directed against the toxin-coregulated pilus isolated from *Vibrio cholerae* provide protection in the infant mouse experimental cholera model. *J. Infect. Dis.* **161:**1231-1236.

Sundstrom, L. 1998. The potential of integrons and connected programmed rearrangements for mediating horizontal gene transfer. *APMIS Suppl.* **84:**37-42.

Sunstrom, N. A., Acheson, N. H. and Hassell, J. A. 1991. Determination of the origin-specific DNA-binding domain of polyomavirus large T antigen. *J. Virol.* **65:**6998-7003.

Suresh, S., Yan, Z., Patel, R. C., Patel, Y. C. and Patel, S. C. 1998. Cellular cholesterol storage in the Niemann-Pick disease type C mouse is associated with increased expression and defective processing of apolipoprotein D. *J. Neurochem.* **70:**242-251.

Svennerholm, A. M., Johnson, G. and Holmgren, J. 1992. Immunity to *Vibrio cholerae* infection. pp 257-271. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.) *Vibrio cholerae and cholera : Molecular to global perspectives.* ASM Press, Washington DC.

Swerdlow, D. L. and Isaacson, M. 1994. The epidemiology of cholera in Africa. pp. 297-307. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.) *Vibrio cholerae and cholera: Molecular to global perspectives.* ASM Press, Washington DC.

Tabor, S. and Richardson, C. C. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA.* **82:**1074-1078.

Tabor, S. and Richardson, C. C. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* **84:**4767-4771.

Takeda, Y. 1986. Thermostable direct haemolysin of *Vibrio parahaemolyticus*. pp.183-205. In, Dorner, F. and Drews, J. (eds.). *Pharmacology of bacterial toxins*. Pergamon Press, Oxford.

Takeshita, S., Sato, M., Toba, M., Masahashi, W. and Hashimoto-Gotoh, T. 1987. High-copy-number and low-copy-number plasmid vectors for *lacZ* α-complementation and chloramphenicol. *Gene*. **61:**63-74.

Tashima, K. T., Carroll, P. A., Roger, M. B. and Calderwood, S. B. 1996. Relative importance of three iron-regulated outer membrane proteins for *in vivo* growth of *Vibrio cholerae*. *Infect. Immun*. **64:**1756-1761.

Tauxe, R., Seminario, L., Tapia, R. and Libel, M. 1994. The Latin American epidemic. pp. 321-344. In, Wachsmuth, I. K, Blake, P. A. and Ølsvik, Ø. (eds.). *Vibrio cholerae and cholera : Molecular to global perspectives*. ASM Press, Washington DC.

Taylor, R. K., Miller, V. L., Furlong, D. B. and Mekalanos, J. J. 1987. Use of *phoA* gene fusion to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci. USA*. **84:**2833-2837.

Taylor, D. N., Echeverria, P., Pitarang, C., Seriwatana , J., Sethabutr, O., Bodhidatta, L., Brown, C., Herrmann, J. E. and Blucklow, N. R. 1988. Application of DNA hybridization techniques in the assessment of disease among refugees in Thailand. *Am. J. Epidemiol.* **127:**179-187.

Taylor, J. L., Tuttle, J., Pramukul, T., O'Brien, K., Barrett, T. J., Jolbitado, B., Lim, Y. L., Vugia, D., Morris, J. G., Tauxe, R. V. and Dwyer, D. M. 1993. An outbreak of cholera in Maryland associated with imported commercial frozen fresh coconut milk. *J. Infect. Dis.* **167:**1330-1335.

Taylor, R. K. 1999. Virus on virus infects bacterium. *Nature*. **399:**312-313.

Teppema, J. S., Guinee, P. A. M., Ibrahim, A. A., Paques, M. and Ruitenberg, E. J. 1987. *In vivo* adherence and colonization of *Vibrio cholerae* strains that differ in hemagglutination activity and motility. *Infect. Immun*. **55:**2093-2102.

Terai, A., Baba, K., Shirai, H., Yoshida, O., Takeda, Y. and Nishibuchi, M. 1991. Evidence for insertion-sequence-mediated spread of the thermostable direct haemolysin gene among *Vibrio* species. *J. Bacteriol*. **173:**5036-5046.

Texier, J. P., Huxtable, J., Rhodes, E., Miallier, D. and Ousmoi, M. 1988. New data on the chronological situation of the aterian of Morocco and their implications. *Comp. Rend. De L Acad. Des Sci. Serie Ii Mecanique Phys. Chim. Sci. De L Univ. Sci. De La Terre*. **88:**124-140.

Thomas, S., Williams, S. G. and Manning, P. A. 1995. Regulation of *tcp* genes in classical and El Tor strains of *Vibrio cholerae* O1. *Gene.* **166:**43-48.

Towbin, H., Staehlin, T. and Gordon. J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets : procedure and some applications. *Proc. Natl. Acad. Sci. USA.* **90:**5267-5271.

Trucksis, M., Galen, J. E., Michalski, J., Fasano, A. and Kaper, J. B. 1993. Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. *Proc. Natl. Acad. Sci. USA.* **90:**5267-5271.

Trucksis, M., Conn, T. L., Fasano, A. and Kaper, J. B. 1997. Production of *Vibrio cholerae* accessory cholera enterotoxin (Ace) in the yeast *Pichia pastoris*. *Infect. Immun.* **65:**4984-4988.

Trucksis, M., Michalski, J., Deng, Y. K. and Kaper. J. B. 1998. The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc. Natl. Acad. Sci. USA.* **95:**14464-14469.

van Dongen, W. M. A. M. and De Graaf, F. K. 1986. Molecular of a gene coding for a *Vibrio cholerae* haemagglutinin. *J. Gen. Microbiol.* **132:**2225-2234.

van Dongen, W. M. A. M., Vlerken, V. and De Graff, F. K. 1987. Nucleotide sequence of a DNA fragment encoding a *Vibrio cholerae* haemagglutinin. *Mol. Gen. (Life Sci. Adv.).* **6:**85-91.

Voss, E. and Attridge, S. R. 1993. *In vitro* production of toxin-coregulated pili by *Vibrio cholerae* El Tor. *Microb. Pathog.* **15:**255-268.

Voss, E., Manning, P. A. and Attridge, S. R. 1996. The toxin-coregulated pilus is a colonization factor and protective antigen of *Vibrio cholerae* El Tor. *Microb. Pathog.* **20:**141-153.

Waldor, M. K. and Mekalanos, J. J. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science.* **272:**1910-1915.

Waldor, M. K., Rubin, E. J., Pearson, G. D. N., Kimsey, H. and Mekalanos, J. J. 1997. Regulation, replication, and integration functions of the *Vibrio cholerae* CTXϕ are encoded by region RS2. *Mol. Microbiol.* **24:**917-926.

Waldor, M. K. and RayChaudhuri, D. 2000. Treasure trove for cholera research. *Nature.* **406:**469-470.

Ward, J. M. and Grinsted, J. 1982. Physical and genetic analysis of the Inc-W group plasmids R388, Sa, and R7K. *Plasmid.* **7:**239-250.

Watnick, P. I., Butterton, J. R. and Calderwood, S. B. 1998. The interaction of the *Vibrio cholerae* transcription factors, Fur and IrgB, with the overlapping promoters of two virulence genes, *irgA* and *irgB*. *Gene*. **209:**65-70.

Wiedemann, B., Meyer J. F. and Zuhlsdorf, M. T. 1987. Insertion of resistance genes into Tn*21*-like transposons. *J. Antimicrob. Chemother*. **18:**85-92.

Williams, P. H., Knutton, S., Brown, M. G. M., Candy, D. C. A. and McNeish, A. S. 1984. Characterization of non fimbrial mannose-resistant protein hemagglutinins of two *Escherichia coli* strains isolated from infants with enteritis. *Infect. Immun*. **44:**592-598.

Williams, S. G. and Manning, P. A. 1991. Transcription of the *Vibrio cholerae* haemolysin gene, *hlyA*, and cloning of a positive regulator, *hlyU*. *Mol. Microbiol*. **5:**2031-2038.

Williams, S. G., Attridge, S. R. and Manning, P. A. 1993. The transcriptional activator HlyU of *Vibrio cholerae* : nucleotide sequence and role in virulence gene expression. *Mol. Microbiol*. **9:**751-760.

Wong, S. M., Carroll, P. A., Rahme, L. G., Ausubel, F. M. and Calderwood, S. B. 1998. Modulation of expression of the ToxR regulon in *Vibrio cholerae* by a member of the two-component family of response regulators. *Infect. Immun*. **66:**5854-5861.

Wu, Z., Milton, D., Nybom, P., Sjo, A. and Magnusson, K. E. 1996. *Vibrio cholerae* hemagglutinin/protease (HA/protease) causes morphological changes in cultured epithelial cells and perturbs their paracellular barrier function. *Microb. Pathog*. **21:**111-123.

Wu, Z., Nybom, P., Sundqvist, T. and Magnusson, K. E. 1998. Endogenous nitric oxide in MDCK-I cells modulates the *Vibrio cholerae* haemagglutinin/protease. *Microb. Pathog*. **24:**321-326.

Wyckoff, E. E., Duncan, D., Torres, A. G., Mills, M., Maase, K. and Payne, S. M. 1998. Structure of the *Shigella dysenteriae* haem transport locus and its phylogenic distribution in enteric bacteria. *Mol. Microbiol*. **28:**1139-1152.

Yamaghuchi, K., Yu, F. and Inouye, M. 1988. A single amino acid determinant of the membrane localization of lipoprotein in *Escherichia coli. Cell*. **53:**423-432.

Yamamoto, K., Ichinose, Y., Nakasone, N., Tanabe, M., Nagahama, M., Sakurai, J. and Iwanaga, M. 1986. Identity of hemolysins produced by *Vibrio cholerae* non-O1 and *Vibrio cholerae* O1 biotype El Tor. *Infect. Immun*. **51:**927-931.

Yang, Y. and Ferro-Luzzi Ames, G. 1990. The family of repetitive extragenic palindromic sequences : interaction with DNA gyrases and histonelike protein HU. pp 211-225. In, Riley, M. and Drlica, K. (eds.). *The bacterial chromosome*. ASM, Washington DC.

Yang, C. Y., Gu, Z. W., Blanco-Vaca, F., Gaskell, S. J., Yang, M., Massey, J. B., Gotto, A. M. and Pawnall, H. J. 1994. Structure of human apolipoprotein D : locations of the intermolecular and intramolecular disulfide links. *Biochem.* **33:** 12451-12455.

Yansura, D.G. and Henner, D.J. 1983. Development of an inducible promoter system for *Bacillus subtilis*. In, Ganesan, A.T. and Hoch, J.A. (eds.). *Biology and Biotechnology of the Bacilli*. Academic, New York.

Yansura, D.G. and Henner, D.J. 1984. Use of the *Escherichia coli lac* repressor and operator to control gene expression in Bacillus subtilis. *Proc. Natl. Acad. Sci. USA.* **81:**439-443.

Yoshida, J., Takamura, S., Suzuki, S. and Nishio, M. 1996. Streptococcal glycoprotein-induced tumour cell growth inhibition involves the modulation of a pertussis toxin-sensitive G protein. *Br. J. Cancer.* **73:**917-923.

Zitzer, A., Palmer, M., Weller, U., Wassenaar, T., Biermann, C., Tranum-Jensen, J. and Bhakdi, S. 1997. Mode of primary binding to target membranes and pore formation induced by *Vibrio cholerae* cytolysin (hemolysin). *Eur. J. Biochem.* **247:**209-216.