Copper Tolerance of
Listeria monocytogenes strain DRDC8

Francesca Y Bell, B. Biotech. (Hons.) (Adelaide)

A thesis submitted for the Degree of Doctor of Philosophy

School of Molecular and Biomedical Science

Faculty of Sciences,

The University of Adelaide

Adelaide, South Australia, Australia

May, 2010

Chapter 1: Introduction

1.1 Introduction

Listeria monocytogenes is a small, Gram-positive, facultative anaerobic intracellular pathogen. First discovered in 1924 as the cause of septicaemia in laboratory rabbits and guinea pigs (Murray *et al.*, 1926), this bacterium was originally named *Bacterium monocytogenes* due to its association with monocytes in peripheral blood (Murray *et al.*, 1926), and later renamed *Listeria monocytogenes* in 1940 (Farber & Peterkin, 1991). *L. monocytogenes* is a ubiquitous bacterium capable of survival and multiplication in diverse habitats and is thus widely distributed within the environment. Since the 1980's, *L. monocytogenes* has been recognized as a food-borne pathogen that causes listeriosis, a severe invasive illness in susceptible humans and animals that may result in death.

Listeria genome sequencing projects have enabled characterisation of the genetic mechanisms that enable this microorganism to adapt and withstand extreme and adverse environments. Common genetic themes between *Listeria* species include genes that encode proteins involved in cation transport and homeostasis. Indeed, the carriage of a large number of these genes on plasmid DNA suggests that plasmids may be important for survival of *L. monocytogenes* in conditions of severe cation stress. Tolerance to cation stress is likely to provide a selective advantage for the persistence of *Listeria* in the environment where cations are frequently present, but may also be implicated in survival and replication within the host where subtoxic or starvation levels of metal ions occur. Significantly, a gene (*ctpA*) that is involved in cation transport, is also significant for virulence of *L. monocytogenes* (Francis & Thomas, 1997b).

The *ctpA* gene encodes a copper transporting P-type ATPase (CtpA) that is apparently involved in maintenance of intra-cellular copper ion homeostasis in *L. monocytogenes* strain DRDC8 (Francis & Thomas, 1997a). Mutants unable to express *ctpA* are apparently important for persistence of *L. monocytogenes* in the liver and spleens of infected mice (Francis & Thomas, 1997b). Substantial evidence has been collated to show that *ctpA* is plasmid-encoded (Bell, 2002). In addition, sequence analysis of DNA flanking *ctpA* has identified several other genes that may be implicated in *ctpA* mediated copper ion

homeostasis in *L. monocytogenes*. However further investigation is required to substantiate this hypothesis. Interestingly, while *ctpA* is important for *L. monocytogenes* infection, its distribution within isolates of *L. monocytogenes* from clinical, environmental and food sources is confined (Francis, 1996; Webster, 2001). Thus the exact role played by *ctpA* and associated genes in *Listeria monocytogenes* isolates awaits further investigation.

1.2 Characteristics of *L. monocytogenes*

From a taxonomic perspective, *L. monocytogenes* is a facultative anaerobic, non-sporulating Gram-positive bacterium. Microscopically they appear as short, rod-shaped bacteria 0.5 - 2.0 μm and 0.4 - 0.5 μm in length and diameter respectively. When cultured on Blood agar, *L. monocytogenes* produces an incomplete β-hemolysis. *Listeria* are motile by means of one to five peritrichous flagella, when cultured at 25°C, but are non-motile at 37°C (Farber & Peterkin, 1991). *L. monocytogenes* is capable of growth and survival at temperatures ranging from < 1°C to < 50°C, a wide pH range (4.0 - 9.5) and at salt concentrations up to 12% (w/v) (Lovett, 1990). Optimal growth of *L. monocytogenes* is observed at 30°C - 37°C, however it is capable of survival in freezing conditions. *L. monocytogenes* ability to survive and multiply under conditions frequently used for food preservation renders it a significant problem for the food industry. This bacterium is a commonly contaminant of foods worldwide (Farber & Peterkin, 1991).

While classically regarded as a Gram-positive bacterium, older cultures of *L. monocytogenes* exhibit Gram variability (Martin & Fisher, 1999). In the Gram-positive state, a typical structure is observed where the cell membrane is surrounded by a thick cell wall, composed of peptidoglycan and lipoteichoic acids. Structural components of the cell wall are used as the basis for antibody typing of different strains. Currently, these typing schemes are based on identification of O (somatic) heat stable antigens and H (flagella) heat labile antigens (Martin & Fisher, 1999). Thirteen serotypes of *L. monocytogenes* (1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e, 7) have been identified, all potentially capable of causing human infection. However, epidemiological data has suggested that the majority of human and animal *L. monocytogenes* infections are associated with serotypes 1/2a, 1/2b and 4b (Wiedmann, 2002a). In particular, serotype 4b strains have been

associated with all major invasive listeriosis outbreaks and the majority of sporadic cases (Rocourt & Bille, 1997).

As of 2009, the *Listeria* genus comprised six species *viz. L. monocytogenes, L. ivanovii, L. seeligeri, L. innocua, L. welshimeri* and *L. grayi*. All members of this genus are generally aerobic or facultatively anaerobic, catalase positive and oxidase negative (Feresu & Jones, 1988). *Listeria* species hydrolyse esculin, ferment glucose without production of gas, are Voges-Proskauer and Methyl red positive, but are unable to hydrolyse urea or produce indole or H_2S (Feresu & Jones, 1988). Species identification is based upon the characteristic shape and size of the rods, motility and anaerobiosis of the organism. *L. monocytogene*s is the only member of the *Listeria* genus pathogenic to both humans and animals, while *L. ivanovii* remains exclusively an animal pathogen (Bille & Doyle, 1991). While all other members of the genus appear non-pathogenic, some rare cases of human infection involving *L. ivanovii* (Bille & Doyle, 1991), *L. innocua* and *L. seeligeri* have been documented (Martin & Fisher, 1999).

The *Listeria* genus is classified as the monospecific family *Listeriaceae* of the Bacillus/Clostridium group (Bortolussi *et al.*, 1985). Other members of this taxonomic group include *Clostridium, Lactobacilli, Bacilli, Streptococci* and *Staphylococci* (Seeliger & Jones, 1986). *Listeria* is antigenically related to a number of other organisms most particularly *Staphylococcus aureus*, *Streptococcus faecalis* and *Clostridium* (Sallen *et al.*, 1996). Mutual features of this group include factors such as low (G+C)% content (< 50 %), lack of mycolic acids and presence of lipoteichoic acids (Feresu & Jones, 1988).

1.3 Listeriosis

The incidence of human listeriosis has increased over the past two decades throughout the world. *L. monocytogenes* has been linked to sporadic episodes and large outbreaks of food-borne illness worldwide (Farber & Peterkin, 1991). White *et al*., (2002) has listed several previously documented, large outbreaks of listeriosis in England, Germany, Sweden, New Zealand, Switzerland, Australia, France and the USA. Since the first report of human listeriosis in 1929, the annual endemic disease rate has varied from 2-15 cases per million in developed countries (Farber & Peterkin, 1991). Currently, most European

Union countries report an annual incidence of human infection between two and ten cases per million (Jemmi & Stephan, 2006).

Listeriosis is one of the most deadly bacterial infections currently known, with a mean mortality rate in humans of 20 to 30% or higher despite antibiotic treatment (Schuchat *et al.*, 1991). Consequently, this organism is amongst the most frequent causes of death due to food-borne illness (Jemmi & Stephan, 2006). The high lethality rate and general severity of the disease makes *L. monocytogenes* a serious food-associated pathogen. A report by Mead *et al.*, (1999), showed that *L. monocytogenes* infections account for the highest hospitalisation rates (91%) amongst known food-borne pathogens in the USA.

While most pathogenic *Listeria* strains are susceptible to antibiotics such as ampicillin, rifampin, co-trimoxazole, vancomycin, erythromycin, or penicillin and gentamycin (White *et al.*, 2002), treatment of listeriosis in both humans and animals is difficult and typically unsuccessful (Charpentier *et al.*, 1995). Human patients are often slow to respond to antibiotic treatment and infection can persist despite therapy (Slutsker & Schuchat, 1999). In the case of animals, clinical signs are usually absent in early stages of infection and antibiotic treatment is often initiated too late to prove effective. In addition, symptoms of infection are commonly misdiagnosed and rapid death often follows onset of clinical signs (Wiedmann & Evans, 2002). Due to the high failure rate of treatment, prevention of listeriosis remains the best means of reducing morbidity and mortality.

1.3.1 Listeriosis in Animals

Several reviews have described the pathogenesis, epidemiology and clinics of animal listeriosis in detail (Hird & Genigeorgis, 1990; Low & Donachie, 1997; Wesley, 1999). As a facultative intracellular parasite of eukaryotic cells, *L. monocytogenes* has a exceptionally broad host range that ranges from mammals, birds, fish, crustaceans, ticks and flies (Kirk, 1993). Clinically significant disease is presented in a wide array of domestic animal species including poultry (Cooper, 1989), primates (Farber & Peterkin, 1991), cats (Decker *et al.*, 1976), pigs (Long & Dukes, 1972), horses (Jose-Cunilleras & Hinchcliff, 2001), and ruminants such as sheep (Wiedmann *et al.*, 1997), cattle (Schoder *et al.*, 2003) and goats. It has been estimated that in Northern America 80 - 90% of reported animal listeriosis cases occur in cattle, whilst the remaining cases generally occur in sheep

(Beauregard & Malkin, 1971; Gray, 1958; Gray & Killinger, 1966; Roberts & Wiedmann, 2003; Wesley, 1999).

Animal listeriosis is predominantly a feed-borne disease, largely being linked to ingestion of poorly prepared silage contaminated with high levels of *L. monocytogenes* (Crump *et al.*, 2002; Gray, 1960; Hinton, 2000; Low & Donachie, 1997; Wesley, 1999). During improper fermentation (at pH > 5.0), pathogenic *Listeria* naturally present in plant material used in silage, grow to high numbers (Fenlon, 1999). The low micronutrient composition of poorly prepared silage may also predispose livestock to infection (Wesley, 1999). Stress associated with the weather, transport or introduction of a new animal into a herd may be risk factors for listeriosis (Wesley, 1999).

Animal infection is primarily associated with disease of the central nervous system or the female reproductive tract. Infection may present as meningoencephalitis, gastroenteritis, while uterine infections are usually characterised by late-term abortions or septicaemia in neonates (Low & Renton, 1985; Schoder *et al.*, 2003; Wiedmann *et al.*, 1996; Wiedmann *et al.*, 1997). Meningoencephalitis or "circling disease", first described in New Zealand, is the most common clinical manifestation of listeriosis in ruminants (Gill, 1933; Gill, 1937). In addition, *L. monocytogenes* is also capable of causing eye infections and keratitis in ruminants (Wiedmann & Evans, 2002) as well as mastitis in both cattle and sheep (Bourry & Poutrel, 1996; Schoder *et al.*, 2003; Wagner *et al.*, 2005).

The encephalitic form of listeriosis is commonly characterised by depression, anorexia, decreased milk production and transient fever (Smith, 1994). This may be followed by neurological symptoms, including circling, excessive salivation and unilateral facial paralysis (Low & Donachie, 1997; Rebhun, 1995). The typical incubation period for encephalitic listeriosis is usually 2 - 3 weeks and the course of disease is approximately 1 - 4 days. In contrast, the incubation period for the septicemic/abortive form of listeriosis can be as short as 1 day (Roberts & Wiedmann, 2003). In the case of pregnancy, bacteria may colonise the placenta and cause abortion (Wesley, 1999). Often, infected animals are able to limit a primary septicemic infection, but become latent carriers, shedding *Listeria* in their faeces and milk (Smith, 1994; Wesley, 1999). In addition to acutely infected animals, a considerably larger number of animals within a herd may be asymptomatic carriers (Wesley, 1999). Asymptomatic carriage has also been reported in poultry (Wesley, 1999).

Dispersal of *L. monocytogenes* via animal faecal shedding is discussed in detail in Section 1.5.1.

1.3.2 Listeriosis in Humans

There are a number of reviews that describe the epidemiology and pathogenesis of human listeriosis in detail (Donnelly, 2001; Farber & Peterkin, 1991; Schuchat *et al.*, 1991; Slutsker & Schuchat, 1999; Vázquez-Boland *et al.*, 2001; Wing & Gregory, 2002). Listeriosis in humans occurs either as a mild non-invasive gastrointestinal illness, or as invasive disease (Roberts & Wiedmann, 2003). Non-invasive gastrointestinal listeriosis typically occurs in otherwise healthy adults and is primarily linked to the consumption of large numbers *L. monocytogenes* from contaminated food products. Symptoms of typical gastroenteritis include fever, diarrhoea and vomiting, with a mean incubation time *ca.* 18 – 20 h before onset of symptoms (Roberts & Wiedmann, 2003).

The most frequently observed (55 - 70%) clinical presentation of invasive listeriosis involves the central nervous system (CNS) (Vázquez-Boland *et al.*, 2001). Bacteremia or septicaemia, are also frequently observed forms of Listeriosis and can be associated with a mortality rate of up to 70% (Lorber, 1990). *L. monocytogenes* is currently the most common cause of bacterial meningitis. Infections in the form of meningoencephalitis are generally associated with a high mortality rate (30 - 60%) (Ortel, 1989). *Listeria* infection of the brain parenchyma are also observed in approximately 20% of all cases (Kirk, 1993). Other atypical clinical forms of Listeriosis include endocarditis, myocarditis, arteritis, pneumonia, aortic aneurysm, arthritis, osteomyelitis, cutaneous lesions, conjunctivitis, hepatitis, urethritis, cerebritis and brain abscess (Farber & Peterkin, 1991).

Listeriosis in neonates can occur as early or late onset (Slutsker & Schuchat, 1999). The majority (45 - 70%) of neonatal listeriosis occurs as early onset (Frederiksen & Samuelsson, 1992; McLauchlin, 1990). In this form, *L. monocytogenes* takes advantage of the natural localised immunosuppression at the maternal interface and infects the foetus *in utero* by transplacental transmission of bacteria from a septic mother, usually resulting in miscarriage or stillbirth (Roberts & Wiedmann, 2003). Late neonatal listeriosis infection involves meningitis, gastroenteritis or pneumonia (Farber & Peterkin, 1991). Despite modern antibiotic therapy, neonatal listeriosis has a mortality rate of approximately 36%

(Farber & Peterkin, 1991). These diverse clinical manifestations of Listeriosis are related to the ability of *L. monocytogenes* to cross three critical barriers in the human host. These include;

(1)   The intestinal barrier: ingestion of contaminated food leads to invasion of the intestinal epithelium followed by access to internal organs,

(2)   The blood-brain barrier: during severe infections *L. monocytogenes* can penetrate the meninges and the brain, and

(3)   The foetoplacental barrier: infection of the mother can lead to infection of the foetus via this barrier.

Susceptibility to infection and the outcome of disease caused by *L. monocytogenes* is dependant upon genetic variation between human hosts as well as the presence or absence of certain behavioural risk factors (Roberts & Wiedmann, 2003). Several susceptible, high risk groups that possess predisposing factors associated with *L. monocytogenes* infection include; the elderly, neonates, pregnant women, cancer patients, organ transplant recipients, individuals receiving immunosuppressive or radiation therapy and AIDS patients. Invasive human listeriosis most often occurs in adults with underlying immunosuppression, usually physiological or pathological defects that interfere with host T-cell mediated immunity (Vázquez-Boland *et al.*, 2001). AIDS is the underlying predisposing condition in 5 - 20% of listeriosis cases in non-pregnant adults, where risk of infection is estimated to be 300 to 1000 times than that of the general population (Vázquez-Boland *et al.*, 2001). Listeriosis in adult AIDS patients has also been strongly associated with malignancies such as sarcomas and leukaemia and immunosuppresent therapies. In addition, several studies have suggested a link between antacid therapy and development of listeriosis (Czuprynski & Faith, 2002; Schlech, 1993; Schuchat *et al.*, 1992). Similarly, it has been suggested that the consumption of alcohol may be a risk factor for infection (Pavia *et al.*, 2002). An abnormally high concentration of iron in the blood following transfusion may also be a risk factor for listeriosis (Vázquez-Boland *et al.*, 2001). In general, the outcome of infection is significantly influenced by the health status of the host, where immunocompetent individuals usually survive infection and higher

mortality rate is often observed in individuals with underlying susceptibility (Vázquez-Boland *et al.*, 2001).

1.3.3 Transmission of Infection to Humans

The illness and asymptomatic carriage of *L. monocytogenes* was originally considered a zoonosis, where an infected animal acts as a primary reservoir for human infection (Bojsen-Moller, 1972). However, it was commonly observed that human infection was established without contact with an infected animal and direct transmission from animals to humans was rarely reported. Reports of rare cases of human infection resulting from animal infection (Cain & McCann, 1986; Felsenfeld, 1951; Owen *et al.*, 1960) indicated that direct transmission of *L. monocytogenes* is possible, but occurs as a highly infrequent event. In addition, while not a common event, there is significant evidence to suggest that nosocomial outbreaks of neonatal listeriosis can occur (Campbell *et al.*, 1981; Filice *et al.*, 1978; Florman & Sundararajan, 1968; Larsson *et al.*, 1978; Lubani *et al.*, 1987; Nelson *et al.*, 1985; Simmons *et al.*, 1986). Typically, breaks in barrier nursing and hygiene in hospital nurseries lead to person-to-person transmission from an initially infected infant.

In the 1980's, a direct linkage between *Listeria* infections and ingestion of contaminated foodstuffs was established following investigation of several large outbreaks of human listeriosis (Bille & Doyle, 1991; Linnan *et al.*, 1988; Schlech *et al.*, 1983; Schuchat *et al.*, 1992). It is estimated that 99% of all human listeriosis cases are caused by consumption of contaminated food products (Mead *et al.*, 1999). *L. monocytogenes* is frequently found as a food contaminant and a major cause of food recalls, particularly in developed countries (Farber & Peterkin, 1991). It is a common contaminant in a variety of raw or processed foods including fresh vegetables, dairy products and meat products including seafood, poultry and fish (Bille & Doyle, 1991). For example, in 2002, a US poultry producer recalled 27.4 million pounds of fresh and frozen ready to eat poultry products due to *L. monocytogenes* contamination (Jacquet *et al.*, 2004). Several specific examples of animal derived *L. monocytogenes* contaminated food products have been implicated in human listeriosis cases. These include raw milk, pasteurised milk, chocolate milk, butter and soft cheeses as well as processed meats and poultry products (Dalton *et al.*, 1997; Headrick & Tollefson, 1998; Maijala *et al.*, 2001). The data presented in Table

1.1 details human listeriosis outbreaks linked to consumption of different food products in several different countries over the period 1976 - 2005. Perhaps one of the most documented outbreaks was linked to consumption of contaminated coleslaw in Maritime, Canada in 1981. Cabbage purchased from a farmer who had sheep with ovine listeriosis and had used raw sheep manure to fertilize a cabbage crop, was later identified as the source of contamination (Schuchat *et al.*, 1991). Another significant outbreak that occurred in Massachusetts, USA in 1983 was linked to consumption of pasteurized milk sourced from a herd of cattle suffering from listeriosis at the time of the outbreak (Fleming *et al.*, 1985). That outbreak involved 49 cases of human infection, 7 infants or foetuses and 42 immunosuppressed adults, of which 14 died (29% mortality).

Changes in food processing and production practices as well as an increase in consumption of minimally processed, ready to eat refrigerated or frozen convenience makes it likely that listeriosis will continue to be a significant human health threat (Tappero *et al.*, 1995). Furthermore, advances in the field of medicine have increased the survival of immunocompromised people and lifespan of the elderly and consequently a growing population of individuals susceptible to infection by *L. monocytogenes*.

1.4 Intracellular Invasion and Infection

The ability of *L. monocytogenes* to act as an enteric bacterium and invade, survive and multiply within both professional and non-professional phagocytic host cells, is critical for development of a systemic infection. These non-phagocytic cells include epithelial cells (Gaillard *et al.*, 1987; Mengaud *et al.*, 1996), fibroblasts (Kuhn *et al.*, 1988), hepatocytes (Dramsi *et al.*, 1995; Wood *et al.*, 1993), endothelial cells (Drevets *et al.*, 1995) and nerve cells including neurons (Dramsi *et al.*, 1995). The lifecycle of *L. monocytogenes* and the process by which infects host cells has been studied extensively and recently reviewed (Cossart, 2002; Hamon *et al.*, 2006; Lecuit, 2005; Vázquez-Boland *et al.*, 2001).

Following consumption of contaminated foodstuffs, the gastrointestinal tract is considered to be the primary site of entry for *L. monocytogenes* via attachment to the intestinal mucosa of the gut. Evidence suggests that the precise point of entry and the mechanism of intestinal translocation is via two possible routes; invasion of intestinal villous epithelium (Racz *et al.*, 1972) or via invasion of specialised M-cells of Peyer's

patches (Jensen *et al.*, 1998). Following translocation across the intestinal barrier, internalisation of *L. monocytogenes* occurs in the underlying *Lamina propria* via active phagocytosis by macrophages, or induced phagocytosis by non-phagocytic cells such as epithelial cells (Farber & Peterkin, 1991). *L. monocytogenes* organisms are then carried by the lymph or blood directly to the mesenteric lymph nodes, the spleen and the liver (Marco *et al.*, 1992; Pron *et al.*, 1998). *In vivo* mouse models of infection have demonstrated that *L. monocytogenes* is then rapidly cleared from the bloodstream by resident macrophages in the spleen and liver (Conlan & North, 1991; Cousens & Wing, 2000; Mackaness, 1962). The majority (90%) of the bacterial load accumulates in the liver, where resident macrophages subsequently kill most of the ingested bacteria (Ebe *et al.*, 1999).

An intricate series of well characterized virulence factors function to facilitate the intracellular infection cycle and are crucial for establishment *L. monocytogenes* host cell infection and invasion (Tilney & Portnoy, 1989). Infection of host cells occurs in the following stages:

(1)   internalisation of the bacterium into the host cell following attachment to the host cell surface;

(2)   bacterial escape from the host cell vacuole via degradation of the endocytic membrane thus allowing entry into the host cytoplasm;

(3)   intracytoplasmic multiplication and movement via bacterial-induced polymerisation of host cell actin;

(4)   bacterial spread to neighbouring cells through pseudopod-like extensions of the host cell membrane and

(5)   bacterial escape from the resulting double-membrane vacuoles (Figure 1.1) (Tilney & Portnoy, 1989).

An intracellular life cycle allows *L. monocytogenes* to proliferate and evade antibody and complement mediated lysis as well as killing by professional phagocytes.

1.4.1 Virulence Determinants

Virulence genes are generally categorized as genes that are present in pathogenic strains, and absent (or at least mutated or not expressed) in non-pathogenic strains. Virulence

genes are expressed when the pathogen is in the host environment (Roberts & Wiedmann, 2003). Disruption of virulence genes results in reduction in virulence.

Eight key virulence genes that are absolutely essential for *L. monocytogenes* pathogenesis have been recently reviewed (Cossart, 2002; Dussurget *et al.*, 2004; Vázquez-Boland *et al.*, 2001). These virulence genes are linked on a 9kb chromosomal island; six of them are clustered together (*prfA, plcA, hly, mpl, actA* and *plcB*), *inlA* and *inlB* form a separate and distinct operon (Figure 1.2). Comprised of three transcriptional units, this virulence locus is coordinately regulated by the transcriptional activator protein PrfA which is encoded by the autoregulated gene *prfA* (Chakraborty *et al.*, 1992). This 27 kDa site-specific DNA binding protein is located downstream from and is cotranscribed with *plcA*. The phosphatidylinositol-specific phospholipase C (PI-PLC), encoded by *plcA*, plays an accessory role in disruption of the phagosomal membrane (Camilli *et al.*, 1993). While PrfA is the key regulatory element for the control of expression of most of the known virulence determinants, it remains unclear as to what other genes may be regulated directly or indirectly by PrfA.

Adjacent to, but downstream from *plcA* is the *hlyA* gene, which encodes listeriolysin O a sulfhydryl-activated pore-forming hemolysin. Listeriolysin, the primary virulence determinant of *L. monocytogenes,* is required for the lysis of the endocytic vacuolar membrane of infected cells and subsequent escape of the bacterium into the host cell cytoplasm (Tilney & Portnoy, 1989). A zinc-dependent metalloprotease encoded by *mpl* is thought to convert the inactive precursor form of phophatidylcholine-specific phospholipase C (PlcB) into its active mature form (Marquis *et al.*, 1997). Another virulence gene, *actA,* encodes for an actin-polymerizing protein. ActA promotes the assembly of polarized actin filaments essential for actin-based intra and intercellular motility (Suarez *et al.*, 2001). Downstream from *actA* is *plcB,* which encodes a broad range phoshatidylcholine-specific phospholipase C (PC-PLC) that aids in escape of bacteria from the primary endocytic vacuole (Marquis *et al.*, 1997). The *inlA* and *inlB* genes encode internalin proteins directly involved in the invasion of normally non-phagocytic host cells, such as epithelial cells (Dramsi *et al.*, 1995).

1.4.2 Environmental Control of Gene Expression

All bacteria coordinately regulate gene expression in response to environmental stimuli, usually detected by sensors located within the cell membrane. Intracellular pathogens such as *Listeria monocytogenes,* encounter dramatically different environments during passage through the stomach, intestine and establishment of host cell infection. Rapid adaptation to environmental stress encountered during intracellular host cell infection is not only essential for survival but also increases the opportunity for host infection by contributing to the virulence of the bacterium via direct activation of virulence genes. Environmental stimuli known to regulate virulence gene expression in pathogenic bacteria include iron, calcium, temperature, starvation, variations in pH and osmolarity, chemical stress, carbon levels, and competition with other microorganisms (Gross, 1993; Mekalanos, 1992; Miller *et al.*, 1989).

The ability of *L. monocytogenes* to respond to environmental stress has consequences for virulence (Marron *et al.*, 1997; O'Driscoll *et al.*, 1996; Rouquette *et al.*, 1996). Induction of transcription of the key *L. monocytogenes* virulence determinants has been attributed to iron limitation (Cowart & Foster, 1981; Geoffroy *et al.*, 1987), osmolarity (Myers *et al.*, 1993; Park *et al.*, 1992) and heat shock (Sokolovic & Goebel, 1989). Expression of the virulence determinants, internalin and *actA*, are also increased in response to elevated iron levels (Conte *et al.*, 1996). In addition, a linkage between stress response and virulence has been shown, where the expression of the stress response mediator protein ClpC was found to be negatively controlled by the virulence protein PrfA (Ripio *et al.*, 1998). ClpC is a chaperone protein important for stationary phase survival and resistance to low pH, oxidation, osmotic stress, high temperature and iron deprivation (Ripio *et al.*, 1998).

Heat stress has been linked to an increase in synthesis of several virulence genes, including the transcriptional activator protein PrfA and PrfA-dependant proteins such as ActA, listeriolysin O, PlcA and PlcB (Sokolovic *et al.*, 1993). PrfA-dependant transcription is weak below 30°C (environmental temperature), but active at 37°C (human body temperature) (Datta & Kothary, 1993; Ripio *et al.*, 1996). Temperature also governs the level of expression of flagellin. In addition, fermentable carbohydrates are involved in

virulence gene expression in *L. monocytogenes* (Behari & Youngman, 1998; Datta & Kothary, 1993; Milenbachs *et al.*, 1997; Ripio *et al.*, 1996).

Bacteria coordinately regulate virulence gene expression by transformation of an environmental stimulus into a cellular signal through a process called of signal transduction (Gross, 1993). This process involves sensor histidine kinases (HPK) and cognate response regulators (RR), which belong to a family of proteins known as "two component regulatory systems". Membrane-associated sensor HPKs allow detection of an environmental cue via a periplasmic or external domain. Signal transduction occurs via phosphate transfer to the cognate intracellular RR protein. The phosphorylated RR then binds to specific DNA sequences and facilitates up-regulation of transcription of specific gene(s) (Ninfa, 1996). HPKs typically contain five conserved domains, H, N, D, F and G boxes and one or more transmembrane domains (Ninfa, 1996). These proteins catalyse the ATP-dependent phosphorylation of a conserved histidine residue in the H-box domain. The phosphate is then transferred to a conserved aspartyl moiety in the RR.

Two component regulatory systems are a common theme in bacterial regulation and in *L. monocytogenes,* have been implicated in sensing environmental changes such as low pH, oxidative and ethanol stress. For example, the *lisR* and *lisK* genes encode for a RR and an HPK, respectively, which are involved in response to stresses such as low pH, regulation of virulence gene expression and *in vivo* survival of *L. monocytogenes* (Cotter *et al.*, 1999).

1.5 Ecology of *L. monocytogenes*

All pathogenic and non-pathogenic *Listeria* species are considered to be ubiquitous within nature and widely distributed in the environment (Hird & Genigeorgis, 1990). *Listeria* spp. are capable of survival and proliferation in numerous environments including soil, plants, animal faeces, decaying vegetation, silage, water and sewage. It is clear that *L. monocytogenes* has the capacity to mount an appropriate response to diverse environmental stimuli and persist in different environmental conditions. It has been shown that *L. monocytogenes* strains differ substantially in gene content. This may be due to evolutionary differentiation, specifically in genes that are most likely to offer a selective advantage within the environment or in an infected host (Doumith *et al.*, 2004). In view of

these characteristics, it remains to be determined how this saprophytic, environmental organism has acquired the capacity to invade eukaryotic cells, establish an intracellular life-style and evade host defences to establish human/animal infections.

While it is known that *L. monocytogenes* can be isolated from many environmental sources and host species, our current understanding of the dynamics of transmission and ecology of this bacterium remains unclear. It is unknown for example, whether *L. monocytogenes* has a host specific environmental reservoir. Furthermore, it not known if animal infection is necessary for dispersal of virulent *L. monocytogenes* into the environment, or whether they function as "dead-end" hosts that do not contribute to the survival and ecological success of this pathogen (Nightingale *et al.*, 2004). Based on the fact that *L. monocytogenes* is frequently carried in the intestinal tract of animals including cattle, poultry and pigs, several studies have suggested that livestock farms may function as a natural reservoir for human infection (Nightingale *et al.*, 2004; Wesley, 1999). Faecal shedding of *L. monocytogenes* by farm animals may represent a direct or indirect supply of the bacterium to the food processing environment.

Whether *L. monocytogenes* is a free-living organism, or one that exists in complex relationships with other microorganisms, such as free-living protozoa in the natural environment, remains to be established. Recent studies suggest that the maintenance of virulence genes in environmental pathogens is a result of their interaction with soil-borne organisms such as nematodes or amoebae (Molmeret *et al.*, 2005; Schulenburg & Ewbank, 2004; Steenbergen & Casadevall, 2003). Pathogenic bacteria may be capable of survival within protozoa and thus transmitted to humans by previously unrecognized routes (Harb *et al.*, 2000). If this scenario can be experimentally confirmed, then this raises the question about whether *L. monocytogenes* is a professional human pathogen, or if infection of human/animals is a consequence of evolution of *L. monocytogenes* virulence genetic systems in protozoa, or other simple organisms, that are essentially the primary reservoir.

Further research is required to gain a more comprehensive understanding of the natural transmission routes and ecology of *L. monocytogenes*. It is unclear why this opportunistic pathogen maintains genes required for an elaborate process of infection of mammalian cells, unless they are required for survival in a natural environment (Zhou *et al.*, 2007). Further research is also required to elucidate the dynamics of transmission from farm and

food-processing environments to the final food product in order to understand the persistence of *L. monocytogenes* in food (Slutsker & Schuchat, 1999).

1.5.1 *L. monocytogenes* in the Farm Environment

While most cases of animal listeriosis are caused by ingestion of poor quality silage (refer to Section 1.3.1), not all cases are feed-borne (Wiedmann *et al.*, 1999). Transmission of *L. monocytogenes* to farm animals may include initial contamination of crops, soil and other organic matter by wildlife that shed *Listeria* in their faeces or contaminated manure used as fertiliser (Smith, 1994). Gastrointestinal carriage and faecal shedding of *L. monocytogenes* has been reported as being relatively common in ruminants (Pell, 1997), particularly in cattle (33%) and sheep (8%) (Nightingale *et al.*, 2004; Wesley, 1999). Latent carrier farm animals may shed *Listeria* in their faeces and continually contaminate the farm environment (Roberts & Wiedmann, 2003).

Carriage and shedding of *L. monocytogenes* within the farm environment presents several important implications. Firstly, *L. monocytogenes* can survive significant periods of time in manure that has been spread on fields (Jiang *et al.*, 2004) and has been linked to the contamination of consumable goods such as fruit, vegetables and other food products (Beuchat & Ryu, 1997; Kathariou, 2002). Secondly, shedding by mastitic cows to milk, or environmental contamination of milk with *L. monocytogenes* during milking or milk collection, represents routes by which *L. monocytogenes* may contaminate the human food supply (Czuprynski, 2005). *L. monocytogenes* can be isolated from approximately 5% of unpasteurised milk samples (Lovett *et al.*, 1987). Thirdly, faecal shedding is often responsible for the contamination of raw meat during the slaughtering process (Jemmi & Stephan, 2006).

Although infected animals and contaminated agricultural environments have rarely been linked to human infections, unprocessed animal derived food products, such as unpasteurised milk or raw foods that have been contaminated by manure from infected or shedding animals, represent a direct association between human infections and *L. monocytogenes* in farm animals and farm environments (Nightingale *et al.*, 2004). In support, studies by Arimi *et al.*, (1997) identified common *L. monocytogenes* ribotypes in dairy-processing and farm environments. Several studies have also shown that

L. monocytogenes subtypes may differ in their virulence and transmission characteristics, thus also differing in their ability to infect animals and survive in farm environments (Kathariou, 2002; Nightingale *et al.*, 2004; Wiedmann, 2002b). In addition, ruminant farm ecosystems have been shown to maintain a high prevalence of *L. monocytogenes* isolates that have been linked to human infection (Nightingale *et al.*, 2004).

Cattle directly contribute to the amplification and dispersal of *L. monocytogenes* into the farm environment (Nightingale *et al.*, 2004). Maintenance and on-farm transmission of *L. monocytogenes* relies upon ingestion of contaminated feeds, amplification of the pathogen within cattle, followed by faecal dispersal of *L. monocytogenes* into the farm environment (Nightingale *et al.*, 2004). Farm environments may therefore represent an ecosystem where the emergence of novel, more virulent *L. monocytogenes* subtypes are selected as a result of growth in a high-transmission-frequency environment (Nightingale *et al.*, 2004).

1.5.2 *L. monocytogenes* in the Food-processing Environment

Commonly isolated from food-processing environments, particularly meat (Midelet & Carpentier, 2002) and dairy (Pritchard *et al.*, 1995) industries, *L. monocytogenes* can survive, and even thrive, in conditions frequently used for preservation of foods that are normally inhibitory to other food-borne pathogens. Extensive effort has been placed on assessment of the potential risk this organism has for significant public health impact. Consequently, risk analysis has emerged as a structured model for improvement of food control systems with the objectives of producing safer food, reducing the number of food-borne illnesses and facilitating domestic and international trade in food. Numerous risk assessments of *Listeria monocytogenes* have been described, including some in Australia (Bemrah *et al.*, 1998; FAO/WHO, 2004; Farber *et al.*, 1996; FSANZ, 2002a; FSANZ, 2002b; FSIS, 2003; Lindqvist & Westoo, 2000).

While typically originating from food raw materials, contamination of food products by *L. monocytogenes* may also occur during the production, processing, manufacturing or distribution phases (Nightingale *et al.*, 2004; Thevenot *et al.*, 2006). The unique growth and survival properties of *L. monocytogenes* contribute to the complexity of preventing contamination and thus human infection. The fact that human illness has been linked to

the consumption of refrigerated processed ready-to-eat foods such as milk, soft cheese, meat products and salads (Schuchat *et al.*, 1991; Schwarzkopf, 1996), highlights that current practices such as pasteurisation, refrigeration and freezing, may not be sufficient to ensure foods are free from *L. monocytogenes*. Studies by Junttila *et al.*, (1988) showed that *L. monocytogenes* is capable of growth at sub-refrigeration temperatures where the mean minimum growth temperature was found to be 1.1 (± 0.3)°C. Due to the fact that refrigeration is the most common way to increase the shelf life of food, the ability of *L. monocytogenes* to survive and multiple at these temperatures is amongst the many factors that make control of this organism problematic. In addition, the fact that *Listeria* are relatively heat resistant (Bearns & Girard, 1958) means that inappropriate pasteurization techniques are unlikely to eliminate *L. monocytogenes* contamination of dairy products.

Survival in the food-processing environment may also be achieved by adherence to various food contact surfaces via biofilms and adaptation to disinfectants and antimicrobial agents. Contaminated food processing equipment plays an important role in dispersal *L. monocytogenes* to processed foodstuffs (Lundén *et al.*, 2003a; Reij & Den Aantrekker, 2004). *L. monocytogenes* is capable of adhering to food handling, storage and processing surfaces such as conveyer belts and stainless steel equipment (Kumar & Anand, 1998; Wong, 1998). Contamination of meat may occur post slaughter, as a result of the use of contaminated equipment. Subsequently, numbers of *L. monocytogenes* cells can be amplified within the chilling and cutting environment (Nesbakken *et al.*, 1996; van den Elzen & Snijders, 1993). In contrast, it has also been suggested that raw meat contaminates meat processing environments (Giovannacci *et al.*, 1999). The ability of *L. monocytogenes* to maintain contamination of food processing plants is genotype associated, likely to be a consequence of natural selection (Autio *et al.*, 2003).

1.5.3 Association of *L. monocytogenes* with Protozoa

Free-living amoebae are present worldwide and have been isolated from soil, water, air insects and animals (Rodríguez-Zaragoza, 1994). These organisms are recognised as playing a significant role in the evolution of obligate and facultative bacteria pathogens (Cirillo *et al.*, 1997; Greub & Raoult, 2004; Molmeret *et al.*, 2005). Predation pressure

from amoebae has the potential to evolve and maintain survival traits in bacteria that inadvertently bestow their ability to initiate disease in animal hosts (Zhou *et al.*, 2007). This concept is supported by studies that have investigated the association between *Legionella pneumophilia* and amoebae. Bacterial mechanisms used by *Legionella pneumophila* for host recognition, cellular entrance and intracellular proliferation is similar for both amoeba and mammalian cells (Winiecka-Krusnell & Linder, 2001). In addition, it has been demonstrated that *L. pneumophila* acquires increased virulence and invasiveness following growth in protozoa (Roberts & Wiedmann, 2003). The fact that this organism harbours genes required for growth in human cells that are also required for intracellular growth in the protozoan host (Segal & Shuman, 1999), provides evidence to suggest that ability to infect humans and cause disease, is a consequence of adaptation to growth within protozoa.

Previous studies have provided evidence to suggest that *L. monocytogenes* shares comparable attributes with *L. pneumophila,* where by it not only has an ability to survive and multiply within eukaryotic cells to establish animal or human infection, it is capable of association with simple organisms such as protozoa (Ly & Müller, 1990). *Listeria* species have also been shown to associate with other simple organisms such as the nematode *Caenorhabditis elegans* (Anderson *et al.*, 2003). Studies by Ly and Müller, (1990) demonstrated that following ingestion, *L. monocytogenes* cells multiplied intracellularly within the protozoan species *Acanthamoeba* sp. and *Tetrahymena pyriformis*. Eight days post infection, apparently infected protozoan cells were lysed and viable *L. monocytogenes* were released into the extracellular environment. However, recent studies have provided contradictory evidence regarding the associative relationship shared between *L. monocytogenes* and protozoa. Zhou *et al.*, (2007) demonstrated that when co-cultured, *L. monocytogenes* does not actively kill *Acanthamoeba castellanii*. Furthermore, Huws *et al.,* (2008); Akya *et al.*, (2009a); Akya *et al.*, (2009b); and Akya *et al.*, (2010) have established that *Acanthamoeba polyphaga*, *castellanii* and *lenticulata* are able to efficiently kill phagocytosed *L. monocytogenes* cells. It remains uncertain whether protozoa represent an important environmental reservoir of this important human pathogen (Harb *et al.*, 2000).

1.6 Mechanisms of Adaptation and Survival

L. monocytogenes is an opportunistic pathogen that has evolved the capacity to inhabit various niches, infect a wide variety of organisms, thrive in a broad range of microenvironments and cause different diseases (Dussurget *et al.*, 2004). Important survival mechanisms allow *L. monocytogenes* to survive and tolerate a wide range of adverse environmental conditions, including temperature stress, osmotic stress, acid stress, biofilm formation, antimicrobial resistance and cation stress. Adaptation to such environmental stress is crucial for *L. monocytogenes* intracellular survival and establishment of successful host infection. For example, prior to establishing a successful infection in humans, *L. monocytogenes* must survive passage through the acidic stomach and following phagocytosis by intestinal epithelial cells, must be able to temporarily withstand the acid and oxidative stresses of the phagocytic vacuole. The mechanisms used by *L. monocytogenes* to adapt and survive a wide range of adverse environmental conditions have been reviewed in detail by (Gandhi & Chikindas, 2007).

Adaptation to altered water availability, salinity, pH and temperature allow *L. monocytogenes* to survive in certain food products. For example, *L. monocytogenes* strains isolated from cheeses were found to be more resistant to pH and sodium chloride stress than other food isolates analysed (Faleiro *et al.*, 2003). These observations are important for this industry because the making of cheeses involves exposure of the cheese to an acidic environment followed by salting. In addition, adaptation to extreme environmental pressure has been shown to provide cross-protection to other stress. Specifically, acid adaptation induces cross-protection to osmotic shock, heat shock (52°) and/or alcohol stress (Faleiro *et al.*, 2003; Farber & Pagotto, 1992; Gardan *et al.*, 2003; Lou & Yousef, 1997; O'Driscoll *et al.*, 1996; Phan-Thanh *et al.*, 2000). Salt adaptation (osmo-adaptation) and temperature stress adaptation may also be linked (Cole *et al.*, 1990). Cross-resistance has important implications for the food industry where food-stuffs commonly encounter these factors during food processing (van Schaik *et al.*, 1999). Understanding how *L. monocytogenes* adapts cellular physiology to overcome stress is important in preventing and controlling its contamination within farm and food processing environments, and thus ultimately preventing human infection (Jemmi & Stephan, 2006).

1.7 Genomics of *Listeria* species

Whilst the genes known to be absolutely essential for virulence of *L. monocytogenes* (refer to Section 1.4.1) have been intensively investigated, there is limited knowledge concerning the precise genetics that govern host infection by *L. monocytogenes*. Novel genes that enable survival under adverse environmental conditions or enhance the bacterium's ability to mount a successful host infection are also likely to exist. For example, the fact that the majority of food-borne outbreaks of listeriosis are linked to isolates of serovar 4b strains, suggests that strains of this serovar may possess unique virulence properties. The identification of genes that are specific for epidemic-associated *L. monocytogenes* strains is important in investigating human infection. From data obtained to date, it is apparent that speciation and diversity of *Listeria* strains has occurred in response to adaptation to niche specific environments. This is thought to have been achieved by lateral gene transfer of DNA encoding novel genes (Hain *et al.*, 2006a).

Genome sequencing projects have investigated the epidemiological and evolutionary relationships between pathogenic and non-pathogenic *Listeria* species to define variant-specific characteristics in strains capable of causing disease (Doumith *et al.*, 2004; Glaser *et al.*, 2001; Hain *et al.*, 2006a; Nelson *et al.*, 2004). These sequencing projects have highlighted the considerable genetic diversity between *Listeria* serovars, species and strains, where different subsets of genes that enable different lifestyles exist (Doumith *et al.*, 2004). Numerous *L. monocytogenes*-specific and several serovar-specific marker genes have been identified that can be used to differentiate the virulence of *L. monocytogenes* strains, and trace listeriosis outbreaks (Doumith *et al.*, 2004). From these studies, it has also become evident that *Listeria* genomes share a surprisingly high amount of synteny with the genomes of other ubiquitous bacterial species such as the soil bacterium *Bacillus subtilis* and the pathogenic bacterium *Staphylococcus aureus* (Glaser *et al.*, 2001). Genetic divergence between the species may have resulted from transformation, gene transfer or acquisition events mediated by plasmids or bacteriophages between cohabitating microorganisms (Buchrieser *et al.*, 2003).

To date (2009), the genomes of 27 *Listeria* species (partial and complete) have been sequenced and are available from GenBank databases (see Table 1.2)[1]. Partial and complete plasmid sequences for 7 of these strains are also available; only plasmids pLI100 (*L. innocua* CLIP11262) and pLM80 (*L. monocytogenes* H7858) have been completely annotated[2]. All *Listeria* chromosomes sequenced to date vary between 2.7 and 3.0 Mb in size, whilst complete plasmid sequences are approximately 80 Kb in size.

1.7.1 Genome Analysis

A comparative genome sequence analysis of *L. monocytogenes* and the non-pathogenic species *L. innocua* identified a high degree of genetic conservation between the two species (Glaser *et al.*, 2001). Common genes included those encoding for transport systems, transcriptional regulators, surface proteins and secreted proteins. The presence of 331 genes that encode transport-associated proteins correlates with the capacity of both *Listeria* species to inhabit and colonise a broad range of ecosystems. Furthermore, the identification of 209 and 203 transcriptional regulators in *L. monocytogenes* and *L. innocua*, respectively, is not surprising given the array of environmental conditions it is capable of surviving (Glaser *et al.*, 2001). The identification of strain-specific genes suggested that virulence in *Listeria* results from multiple gene acquisition and deletion events. Amongst the 270 predicted genes that were encoded by the *L. monocytogenes* genome, but absent from the *L. innocua* genome, were genes involved in acid resistance and bile salt degradation. This characteristic likely reflects the ability of *L. monocytogenes* to survive within the mammalian gut.

Studies of the genetic diversity between the epidemic serotype 4b strain *L. monocytogenes* CLIP80459 and *L. monocytogenes* EGDe have also identified common genetic content and genome organization (Doumith *et al.*, 2004). Common genes include those encoding proteins such as surface proteins, transport proteins and transcriptional regulators. Nevertheless, a degree of genetic diversity exists between the two strains.

---

[1] Since completion of this thesis, the genomes of 3 additional *Listeria* species have been sequenced and are available on the NCBI GenBank Nucleotide Database.
[2] Since completion of this thesis, *Listeria monocytogenes* 08-5578 plasmid pLM5578 has been completely sequenced and annotated (GenBank Accession CP001603).

Approximately 8% of the CLIP80459 genes were absent from *L. monocytogenes* EGDe genome (Doumith *et al.*, 2004). Furthermore, a comparative analysis of the genomes of three strains of *Listeria monocytogenes* associated with food-borne illness identified a total of 51, 97 and 69 strain specific genes in strains F2365 (serotype 4b), F6854 (serotype 1/2a) and H7858 (serotype 4b), respectively (Nelson *et al.*, 2004). Some 83 genes were found restricted to serotype 1/2a and 51 to serotype 4b strains. These strain- and serotype-specific genes are likely to contribute to observed difference in pathogenicity as well as the ability of the organisms to survive and growth in their respective environmental niches (Nelson *et al.*, 2004). In addition, considerable similarity in gene content and organization was observed between these genomes (Nelson *et al.*, 2004).

1.7.2 Plasmid Analysis

The presence of plasmids in *Listeria* species was first reported by Perez-Diaz *et al.*, (1982), who isolated a 38.5-MDa plasmid in 4 of 29 strains tested. A more extensive study by Lebrun *et al*., (1992) quantified the presence of plasmid DNA in 173 unrelated *L. monocytogenes* strains isolated from human, animal, environmental and food sources. Approximately 28% of isolates were found to harbor plasmids, ranging in size from 24 to 106 kb. The percentage of plasmid positive strains in *Listeria* spp. is found to be higher in strains of food and environmental origin than in clinical isolates (Perez-Diaz *et al.*, 1982; Peterkin *et al.*, 1992). With an overall frequency of 77%, Kolstad *et al.*, (1992) also found that the frequency of plasmids was the highest in environmental *L. monocytogenes* isolates, with clinical isolates showing the lowest plasmid percentage (28%).

Restriction analysis has revealed that homologous regions are shared between plasmids isolated from different *Listeria* (Flamm *et al.*, 1984; Kolstad *et al.*, 1991; Kolstad *et al.*, 1992; Margolles & de los Reyes-Gavilan, 1998). In addition, *L. monocytogenes* plasmid profiles may be serotype related, whereby plasmid DNA is more frequently isolated from serotype 1 strains than from serotype 4 (Flamm *et al.*, 1984; Kolstad *et al.*, 1991; Kolstad *et al.*, 1992; Lebrun *et al.*, 1992; Margolles & de los Reyes-Gavilan, 1998; McLauchlin *et al.*, 1997). Furthermore, plasmid profiling has also shown that common plasmids are shared amongst isolates of the same serotype; no identical plasmid was found to be shared by both serotype 1 and 4 (Lebrun *et al.*, 1992). Table 1.3 shows data obtained by

McLauchlin *et al*., (1997), who analysed the distribution of plasmid DNA in *L. monocytogenes* isolates from serotypes 1 and 4 from food and clinical sources. Whilst most isolated plasmids have been reported to be cryptic, research by Lebrun *et al.*, (1992) demonstrated through conjugation experiments that 95% of *L. monocytogenes* plasmids isolated confer cadmium resistance. Several studies have since shown that plasmids are involved in cadmium and antibiotic resistance in *Listeria* (Facinelli *et al.*, 1993; Hadorn *et al.*, 1993; Lebrun *et al.*, 1994a; Lebrun *et al.*, 1994b; Poyart-Salmeron *et al.*, 1990; Poyart-Salmeron *et al.*, 1992).

Whole-genome sequencing projects have identified 80 genes in the *L. innocua* CLIP11262 81.9 kbp plasmid pLI100 (Glaser *et al.*, 2001). The annotated sequence identified genes encoding for proteins involved in transcriptional regulation, plasmid replication and DNA repair. In particular, a large number of genes encoding transposases were identified. Most significantly, the plasmid encoded a high number of proteins involved in transport and resistance of cadmium, arsenic, potassium and other heavy metals. Several of these are associated with accessory proteins including negative repressors and two component regulatory systems. The *L. monocytogenes* strain H7858 plasmid pLM80 shares a high level of sequence similarity to pLI100 and have a similar genetic organisation (Nelson *et al.*, 2004). pLM80 also contains numerous genes that encode transposases, transcriptional regulators and proteins involved in metal ion homeostasis and resistance. In addition, several common genetic regions were found between the two plasmids, in particular a region encoding mobile genetic elements and proteins responsible for the detoxification of arsenate and cadmium (Nelson *et al.*, 2004). Both sequences have common replication regions. The organisation of these plasmids suggested that they have been constructed through gene insertion and deletion events via *Listeria*-specific mobile genetic elements. This is supported by the high number of transposable elements present (Nelson *et al.*, 2004). A summary of the protein categories encoded by plasmids pLI100 and pLM80 is shown in Table 1.4.

Based on comparative analysis of the two plasmids, it is clear that genes involved in metal ion homeostasis and resistance are a common genetic theme. Plasmid-encoded metal ion transport and resistance systems in other bacterial species have been reviewed and described by (Silver & Walderhaug, 1992).

1.8 Cation Transport and Homeostasis

Metal ions are present throughout different environments in various concentrations. Whilst metal ions are crucial for normal cell growth and metabolism, they are toxic when present in excess concentrations. Consequently, all cells have developed homeostatic mechanisms to maintain low intracellular ion concentrations. These mechanisms enable cell survival and proliferation in conditions of cation stress encountered in different environments. All bacteria possess highly specific systems that are responsible for transport and resistance of most cations and heavy metals (Silver, 1996a). Mechanisms of bacterial heavy metal resistance have been extensively reviewed (Silver & Walderhaug, 1992; Silver & Ji, 1994; Silver, 1996a; Silver & Phung, 1996; Silver, 1998).

Cation transporters and resistance systems are most frequently encoded on plasmids, but are also encoded by chromosomal genes. Bacterial plasmids have been reported to encode genes responsible for resistance to toxic heavy metals including; Ag^+, AsO_2^-, AsO_4^{3-}, Cd^{2+}, Co^{2+}, CrO_4^{2-}, Cu^{2+}, Hg^{2+}, Ni^{2+}, Pb^{2+}, Sb^{3+} and Zn^{2+} (Silver & Ji, 1994), while chromosomes have been reported to confer resistance to cations including; K^+, Mg^{2+}, Co^{2+}, Fe^{3+}, Mn^{2+}, Zn^{2+} (Silver, 1996a). Most cation resistance systems function by energy dependent efflux of toxic ions through ATPases (refer to Section 1.8.1) or though chemiosmotic cation/proton exchangers (Silver, 1996a). Metal ion resistance systems have been reported to exist on plasmids of every Eubacterial group tested from *E. coli* to *Streptomyces* (Silver, 1998). Within a bacterial cell, multiple transporters may exist to transport a single cation (influx and efflux) or a single transporter may exist that is responsible for transport of multiple ions (Silver, 1978). Mechanisms of resistance to cation stress may also vary between bacterial species or strains. For example, cadmium resistance involves an ATPase mediated transport in Gram-positive bacteria, chemiosmotic antiporters in Gram-negative bacteria and metallothionein proteins in cyanobacteria (Silver, 1998).

1.8.1 P-type ATPases

P-type ATPases form a large family of membrane proteins responsible for the ATP-driven translocation of cations across the cell membrane and confer resistance to toxic heavy metal ions (Lutsenko & Kaplan, 1995). P-type ATPases are found in almost all organisms

and are essential for cell survival. These ubiquitous membrane proteins have been classified into five groups according to ion specificity, biological occurrence and sequence (Axelsen & Palmgren, 1998). All P-type ATPases share the same basic structural characteristics, but vary in amino acid composition and these differences reflect functional diversity. Structurally, P-type ATPases consist of three to five membrane-spanning hydrophobic hairpins with 80% of the protein sequence in the cytoplasm and 20% within the plasma membrane (Silver *et al.*, 1989). Two large protein loops are usually located on the cytoplasmic side of the membrane.

Cation transporting P-type ATPases contain a number of conserved structural and functional domains, which include eight transmembrane helices, a CPC ion transduction motif, an ATP-binding site, a DKTG phosphorylation domain and an N-terminal metal binding motif (Smith *et al.*, 1993). This N-terminal motif is not required for the copper transporting properties of the copper exporting P-type ATPase, CopA, in *E. coli* (Fan *et al.*, 2001). Amino acid sequence alignment reveals significant degrees of homology indicating similar biochemical and structural properties, where closely related P-type ATPases appear to transport the same ions.

1.8.2 Copper Ion Homeostasis

Regulation of trace element concentrations within a cell is fundamental for survival. In particular, maintenance of intracellular copper ion concentration is crucial for cell growth, metabolism and viability in all life forms. Copper is an essential cofactor in over 30 various redox enzymes including cytochrome oxidase, superoxide dismuatse, and lysyl oxidase (Peña *et al.*, 1999). Whilst copper is required to sustain life, with roles in gene regulation, enzyme structure and catalysis, copper ions are toxic when present in excess. Redox cycling through between Cu^{2+} (oxidized) and Cu^{1+} (reduced) may catalyse the production of highly toxic hydroxyl free radicals that interfere with the functioning of intracellular macromolecules (Halliwell & Gutteridge, 1984). Reactive oxygen intermediates may be generated either endogenously during aerobic respiration or exogenously by neutrophils or macrophages (Agranoff & Krishna, 1998). Hydroxyl radicals can cause extreme cellular damage through direct oxidation of proteins, lipid peroxidation in membranes and cleavage of DNA or RNA molecules (Peña *et al.*, 1999).

Copper ion translocation systems such as P-type ATPase pumps play an integral role in bacterial survival, by maintaining a low, constant intracellular level of copper ions. There are many well documented examples of bacterial P-type ATPases involved in copper ion transport and homeostasis, several of which have been the subject of recent review (Magnani & Solioz, 2005; Osman & Cavet, 2008; Rensing & Grass, 2003; Solioz *et al.*, 2010). These homeostatic transport systems are known to be controlled in response to copper ion toxicity or starvation by two mechanisms of transcriptional regulation. These are CopY-like family of negative transcriptional regulators or copper-responsive two-component regulatory systems.

1.8.2.1 Regulation of Copper Homeostasis by CopY Transcriptional Repressors

One of the most studied and perhaps the most well understood bacterial copper homeostasis mechanisms is the *cop*YZAB operon in the Gram-positive bacterium *Enterococcus hirae* (refer to Figure 1.3) (Odermatt *et al.*, 1993; Odermatt & Solioz, 1995). Copper homeostasis in *E. hirae* has been extensively reviewed (Solioz & Stoyanov, 2003). CopA is a 727 amino acid P-type ATPase responsible for copper ion uptake and accumulation under conditions of copper limitations. CopB is a 745 amino acid P-type ATPase required for extrusion of copper ions when ambient concentrations reaches toxic levels. CopA and CopB function in synergy to maintain tolerable levels of copper inside the cell (Odermatt *et al.*, 1993). Expression of the operon is regulated in response to copper-starvation and excess conditions by the CopZ and CopY proteins (Odermatt & Solioz, 1995). Located upstream of *copA* and *copB, copZ* and *copY* encode hydrophilic proteins of 69 and 145 amino acids, respectively. CopZ is a copper chaperone responsible for delivery of Cu^{2+} ions to CopY (Cobine *et al.*, 1999). An N-terminal metal binding motif (MXCXXC) is characteristic of most CopZ-like Cu^{2+} chaperones proteins (Odermatt & Solioz, 1995). Mutational analysis by Odermatt and Solioz, (1995) showed that inactivation of *copZ* suppresses *copA* and *copB* expression to very low levels.

CopY functions as a copper-inducible transcriptional repressor which exhibits highly specific DNA protein interactions to regulate expression of the *cop* operon (Strausak & Solioz, 1997). Mutation in the *copY* gene results in an constitutive overexpression of *copA* and *copB* (Odermatt & Solioz, 1995). The CopY family of transcriptional regulators

typically contain three consensus motifs, a heavy-metal binding motif ($CXCX_4CXC$) at the C-terminus and two conserved sequences at the N-terminus ($IX_3EXEVMX_2W$ and $WX_3TX_2TX_3RLX_2K$) (Odermatt & Solioz, 1995). CopY is a homodimeric zinc protein ($[Zn(II)CopY]_2$) which is bound to a promoter region of the *cop* operon promoter in normal growth media, thus repressing transcription. When the cytoplasmic copper concentration rises, CopZ molecules deliver Cu^{2+} to CopY to displace the structurally required Zn^{2+}, converting $[Zn(II)CopY]_2$ to $[Cu(I)_2CopY]_2$, which is no longer able to bind to the promoter (Cobine *et al.*, 1999). CopY is released from the DNA, allowing transcription of the operon to proceed.

CopY dimers bind to two separate sites in the operator between nucleotides -71 to -11, relative to its own translational start site (Strausak & Solioz, 1997). These sites are hyphenated inverted repeats, called '*cop* boxes', designated by the consensus sequence (TACAXXTGTA) (Portmann *et al.*, 2004). This motif is found in the *cop* promoters of several different Gram-positive organisms including *Lactococcus lactis* and *Streptococcus mutans* and acts as the binding site for CopY-like copper-responsive repressors (Portmann *et al.*, 2004). Transcription is initiated between these repeats at nucleotide -42 (Strausak & Solioz, 1997). Figure 1.4 illustrates the $[Zn(II)CopY]_2$–DNA interaction with the *E. hirae cop* promoter *cop* boxes.

Comparative genome studies by Reyes *et al.*, (2006) have identified a conserved *cop*-like operon that involves CopY regulators, in 14 strains within nine species of the Order *Lactobacillale*. Two CopY binding sites designated by the consensus '*cop* box' motifs were identified in 11 of the 14 strains, these include all species analysed from the *Streptococcus* genus and *L. johnsonii*. Genes that are similar to *copA* were identified in 13 of the 14 strains, whilst a *copB*-like gene (copper efflux) was located in only one. The absence of *cop*B-like genes suggests that copper efflux may be provided by paralogs of CopA ATPases or alternatively, CopA-like ATPases may function in influx and/or efflux (Reyes *et al.*, 2006). In most cases the *copA*-like genes are located downstream from *copY*-like genes, while the position of *copZ*-like genes was variable, either preceding or following *copA*-like genes. *copZ*-like genes were identified in 9 of the 14 strains (Reyes *et al.*, 2006). In addition to orthologs of the *E. hirae cop* operon, other genes implicated in copper ion homeostasis were identified in some of the strains analysed. For example, a

putative cupredoxin-like protein (CuA) was identified in three strains (Reyes *et al.*, 2006). Cupredoxins, also known as blue copper proteins, are small (10 - 20 kDa) soluble copper proteins which shuttle electrons from a donor to an acceptor (Arnesano *et al.*, 2002; De Rienzo *et al.*, 2000).

1.8.2.2 Regulation of Copper Homeostasis by Two-Component Regulatory Systems

Although CopY proteins play an important role in regulation of expression of other proteins involved in maintenance of copper homeostasis, there are examples of other regulatory controls that affect expression of copper homeostasis genes and gene products. For example, the *cop* and *cin* operons of *Pseudomonas* spp. are regulated by two-component regulatory systems.

This plasmid-encoded copper resistance operon of *P syringae* comprises the *copABCD* genes. This operon is metalloregulated by the *copRS* two component regulatory system (Cha & Cooksey, 1991; Mellano & Cooksey, 1988; Mellano & Cooksey, 1988a; Mills *et al.*, 1993). CopB functions as an outer membrane protein, whilst CopA and CopC function as periplasmic copper-binding proteins (Cha & Cooksey, 1991). These proteins mediate copper ion sequestration in the periplasm and outer membrane and may prevent toxic levels of copper from entering the cytoplasm (Cha & Cooksey, 1991). CopD functions as a cell membrane protein involved in copper transport (Cha & Cooksey, 1993).

The *copRS* genes are located 3' to the *copABCD* operon and whilst transcribed in the same orientation, they are expressed from a separate, constitutive promoter 5' to the *copR* gene (Mills *et al.*, 1993). CopRS are responsible for activation of expression of the *cop* operon in a copper-sensitive fashion (Mills *et al.*, 1993). CopS protein is a membrane bound copper binding protein that functions as a sensor HPK to detect high levels of free copper ions in the periplasm (Mills *et al.*, 1993). Phosphorylated CopS transfers the phosphate to the RR protein, CopR and converts it from an inactive to an active state. The active phosphorylated CopR binds to a conserved site within the promoter region of the *copABCD* operon and induces expression (Mills *et al.*, 1994).

In *Pseudomonas putida*, the CinS-CinR two component regulatory system activates transcription of *cinAQ* in the presence of copper (Quaranta *et al.*, 2009). The *cinAQ* genes

are located adjacent to the *cinRS* two-component regulatory system. These copper-inducible genes, encode a copper-containing azurin-like protein (*cinA*) and a pre-Q$_0$ reductase (*cinQ*) (Quaranta *et al.*, 2007). The *cinRS* operon encodes a typical bacterial HPK (CinS) and a DNA binding RR (CinR). Together, the CinS-CinR two component regulatory system was shown to be responsive to 0.5 M copper, but is also specifically activated by silver ions (Quaranta *et al*., 2009). Two hisitidines located within the periplasmic domain of CinS are essential for induction of transcription from the *cinAQ* promoter. This observation suggests that these hisitidines may act as a periplasmic copper sensing site of the copper sensing HPK. In addition, the CinR-CinS two component regulatory system may be involved in detection of physiological levels of copper as part of the process of activation of copper utilization pathways.

1.8.2.3 Copper Homeostasis in Copper Tolerant Strains of *E. coli*

Several chromosomally-determined systems are involved in copper tolerance in *E. coli*. CopA is a P-type ATPase that functions in copper excess conditions to pump Cu^{2+} ions out of the cytoplasm (Rensing *et al.*, 2000). CueO is a 516 amino acid multicopper oxidase that functions to protect the periplasm from copper-mediated damage (Grass & Rensing, 2001b). CueO is a homolog of the plasmid-encoded CopA protein of *Pseudomonas syringae* (Cha & Cooksey, 1991). The *cueO* and *copA* genes are both regulated by CueR, a copper activated homologue of the metallo-regulatory protein, MerR (Outten *et al.*, 2000).

The chromosomal *cus* determinants, consisting of two operons (*cusRS* and *cusCFBA*), also function in copper ion resistance in *E. coli* (Grass & Rensing, 2001a; Munson *et al.*, 2000). The *cusRS* operon encodes a two component regulatory system; CusR functions as a RR and CusS an HPK (Munson *et al.*, 2000). This two-component regulatory system regulates transcription of *cusCFBA* in a copper-dependent fashion (Franke *et al.*, 2001). The promoter region of *cusCFBA* includes a motif with sequence homology to the CopR binding motif in *P. syringae* (Franke *et al.*, 2003; Mills *et al.*, 1994). The *cusCBA* genes belong to a family of proton-cation antiporter complexes involved in the export of metal ions, whilst CusF is a periplasmic copper-binding protein (Franke *et al.*, 2001). The genes encoded by the *cusCFBA* operon form the Cus complex which is responsible for copper

ion efflux by transport of Cu(I) from the periplasm, directly across the outer cell membrane (Franke *et al.*, 2003).

Additional copper tolerance in *E. coli* may be provided by the copper resistance *pco* operon encoded on plasmid pRJ1004 (Tetaz & Luke, 1983). The *pco* operon (*pcoABCDRSE*) governs inducible copper resistance by decreasing the net accumulation of copper within the cell (Brown *et al.*, 1995; Lee *et al.*, 1990; Rouch *et al.*, 1985; Tetaz & Luke, 1983). Its expression is regulated by a two-component activator/sensor system encoded by the *pcoR* (HPK) and *pcoS* (RR) genes (Brown *et al.*, 1995). PcoA is predicted to be a multicopper oxidase that interacts with PcoB, an outer membrane protein which may function in copper binding. PcoC is predicted to be a periplasmic protein involved in intracellular copper binding responsible for delivery of periplasmic copper to PcoD, an inner membrane protein thought to be involved in copper uptake. An additional gene, *pcoE* is expressed from a separate copper-inducible promoter regulated by the *cusRS* two-component regulatory system (Brown *et al.*, 1995; Munson *et al.*, 2000). PcoE is a periplasmic protein which is found to reduce the time required for *E. coli* strains to recover from copper ion stress by binding to copper ions in the periplasm (Munson *et al.*, 2000; Outten *et al.*, 2000).

1.9 Copper Transport in *L. monocytogenes* strain DRDC8

For *L. monocytogenes* dairy strain DRDC8, the *ctpA* gene encodes a 653 amino acid polypeptide (CtpA) that functions as a copper transporting P-type ATPase to maintain intracellular copper homeostasis (Francis & Thomas, 1997a). CtpA shares 42.7% and 48.4% sequence identity to two well-characterized examples of P-type ATPases involved in the transport of copper ions, CopA of *E. hirae* (Odermatt *et al.*, 1993) and PacS of *Synechococcus* (Kanamaru *et al.*, 1993), respectively (Francis & Thomas, 1997a). CtpA also shares 43% sequence identity with two well characterized P-type ATPases associated with the human copper metabolism disorders, Menkes (Chelly *et al.*, 1993; Mercer *et al.*, 1993; Vulpe *et al.*, 1993) and Wilson disease (Bull *et al.*, 1993).

CtpA comprises six putative membrane spanning α-helices and contains all of the key conserved structural domains typical of cation transporting P-type ATPases, including ion transduction, phosphorylation, ATP-binding and hinge domains. However, CtpA lacks the

consensus N-terminal motif (CXXC) (Bull & Cox, 1994) typically responsible for copper binding (Francis & Thomas, 1997b). This motif is thought to act by sensing and binding copper ions in the cytosol for transfer to the translocation domain. Other examples of copper transporting P-type ATPases that do not carry the CXXC motif have been identified. For example, the CopB copper transporting P-type ATPase of *E. hirae* contains an alternative copper binding site consisting of HXXMXGM repeats (Odermatt *et al.*, 1993). Sequence analysis of *ctpA* and flanking regions thus far has failed to identify any motif or accessory protein pertaining to metal binding function (Bell, 2002; Francis, 1996). However, the *ctpA* gene shares significant similarity with other cation transporting P-type ATPases from *Lactobacillus* strains, which are also reported to lack the CXXC motif. Indeed, the N-terminus of the CopA Cu(I)-translocating P-type ATPase is not required for resistance or transport of copper in *E. coli* (Fan *et al.*, 2001). This may also be the case for CtpA. Alternatively, an undefined copper binding motif may be involved in copper binding by CtpA.

Growth studies of *L. monocytogenes* DRDC8 variants C185 (containing a transposon insertion in *ctpA* (*ctpA*::Tn917-*lacZ-cat86*)) and DSE201 (containing a *ctpA*::*erm* mutation) in conditions of copper ion stress, provided direct evidence that implicated CtpA in copper ion transport (Francis & Thomas, 1997a). The *ctpA* mutants showed significantly restricted growth and hypersensitivity to low copper ion conditions induced in the presence of the Cu^{2+} chelating agent 8-hydroxyquinoline compared to wild type *ctpA*$^+$ strains (Figure 1.5). The same effects could not be demonstrated for other cations. In support, unpublished studies (Mok, 2003) showed that growth of *ctpA*$^-$ wild type strains was severely inhibited in broth supplemented with 10mM copper ion concentration in comparison to *ctpA*$^+$ strains (Figure 1.6). Furthermore, Northern blot analysis showed that presence or absence of Cu^{2+} ions directly influences induction of transcription of the *ctpA* gene (Francis & Thomas, 1997a). Growth of strain DRDC8 under conditions of copper starvation and high copper concentrations resulted in induction of transcription and concomitantly elevated levels of *ctpA* mRNA molecules. However, *ctpA* mRNA was barely detectable in the absence/presence of divalent metal ions other than copper. Given that *L. monocytogenes* is ubiquitously distributed throughout the environment, these findings suggested CtpA may assist the survival of strain DRDC8 in its natural

environment by establishing copper homeostasis through transport of Cu^{2+} ions into the cell (Francis & Thomas, 1997b).

1.9.1 Significance of *ctpA* for Virulence of *L. monocytogenes*

Studies by Francis and Thomas, (1997b) investigated the role of CtpA in the *in vivo* growth and survival of *L. monocytogenes*. When used to infect mice, the mutant strain (DSE201) showed significantly restricted growth and persistence and was more rapidly cleared from the liver and to a lesser extent, the spleen in comparison to wild-type strains (Figure 1.7). Competition experiments involving intravenous infection of mice with a combination of wild type and mutant strains provided supporting results. Organ tissue (liver and spleen) was harvested from infected mice during the course of a six-day infection. The persistence and recovery of mutant bacteria in the livers of infected mice was shown to be dramatically reduced when compared to wild-type organisms. Wild type strains were able to persist in organs of infected mice for the duration of the experiment. These results showed that *ctpA* was implicated in establishment of persistent *L. monocytogenes* infections in mice. *ctpA* is the first described copper transport protein involved in pathogenesis of *L. monocytogenes*.

Trace elements such as copper, iron and zinc are essential for cell viability. Importantly, infection can dramatically alter the concentration of trace elements within humans and laboratory animals in response to systemic inflammation (Beisel, 1977). During parasitic infection, Cu^{2+} levels are significantly reduced in the liver (Crocker *et al.*, 1992; Matousek de Abel de la Cruz *et al.*, 1993) and significantly increased in the spleen of infected rats in comparison to normal uninfected animals (Matousek de Abel de la Cruz *et al.*, 1993). Infected animals became more susceptible to infection as a result of these ion changes. It is possible that *ctpA* mutants were more rapidly cleared from the liver of infected mice in response to reduced Cu^{2+} concentrations in the liver induced by infection (Francis, 1996). Furthermore, the fact that during normal host infection, the liver and the spleen function as primary sites for macrophage recruitment to clear *L. monocytogenes* from the bloodstream (refer to Section 1.4), is important in understanding why *ctpA*⁻ mutants were more rapidly cleared than wild type strains in the liver and spleen in mice. The results suggested that *ctpA* mutants exhibit attenuated virulence in response to low Cu^{2+} availability and that CtpA is necessary to establish a tropism of *L. monocytogenes* for

the liver of an infected host (Francis, 1996). *L. monocytogenes* may also require other determinants that complement the pathogenic process. In addition, the fact that an inhibition of growth and persistence of *ctpA⁻* mutants is reflective of Cu^{2+} ion availability suggested that CtpA is responsible for Cu^{2+} influx in *L. monocytogenes* (Francis & Thomas, 1997a).

More recent studies by Osman *et al.*, (2010) identified two copper exporting P_{1B}-type ATPases that are important for survival of pathogenic *Salmonella enterica sv. Typhimurium*. In this study, mutants that lack the CopA and GolT copper-transporting P-type ATPases, hyperaccumulated copper and exhibited reduced survival inside cultured macrophages. Significantly, these studies also monitored copper-availability to *Salmonella* in macrophage-phagosomes. This work showed that elevated copper is a feature of this bacterial compartment. These findings suggested that copper resistance aids survival of the bacterium in this compartment. CopA and GolT are proposed to contribute to *Salmonella enterica sv. Typhimurium* copper resistance by pumping copper from the bacterias cytosolic compartment to the periplasm. On this basis, it was hypothesised that elevated copper within macrophages may select for specialised copper-resistance systems in pathogenic microorganism including *S. enterica sv. Typhimurium* (Osman *et al.*, 2010).

In addition, increased killing of *Escherichia coli* in activated macrophages is also associated with increased copper uptake and trafficking of the copper exporting P1B-type ATPase, ATP7A, from the Golgi apparatus to phagosome associated vesicles (White *et al.*, 2009). Furthermore, an *E. coli* mutant lacking the CopA copper-exporter, exhibited both copper-sensitivity and reduced viability in macrophages. Together, these studies are consistent with a model in which copper-toxicity contributes to pathogen killing within macrophage phagosomes. Whether this is the case for *L. monocytogenes* has yet to be confirmed, although Francis and Thomas, (1997b), showed that strain DRDC8 requires a copper homeostasis mechanism to establish persistent infection in the liver and spleen of mice.(White *et al.*, 2009)

1.9.2 Genes Accessory to *ctpA* mediated Copper Transport

Sequence analysis of clones carrying DNA fragments from *L. monocytogenes* strain DRDC8 encoding *ctpA* has identified several genes flanking *ctpA* that may be implicated in copper ion homeostasis in *L. monocytogenes* (Bell, 2002). Based on predictions of function and proximity to the *ctpA* gene (pCT0020), open reading frames (ORFs) pCT0017, pCT0018 and pCT0019 (GenBank Accession U15554, presented in Appendix A, page 310) (see Table 1.5 for ORF description and Figure 1.8 for location)[3] are likely to encode proteins that function in *ctpA*-mediated copper ion transport (Bell, 2002). Furthermore, ORFs pCT0017, pCT0018, pCT0019 and *ctpA* may represent a *cop*-like operon similar to that described in *E. hirae* (see Section 1.8.2.1).

ORF pCT0017 encodes a polypeptide similar to CopY copper-responsive transcriptional repressor proteins commonly found in *cop*-like operons of Gram-positive bacteria, including the CopY negative transcriptional regulator of *Streptococcus mutans* (Vats & Lee, 2001) and *Streptococcus pneumoniae* R6 (Accession NP_358233). The three consensus motifs characteristic of the CopY family of transcriptional regulators; a heavy-metal binding motif ($CXCX_4CXC$) at the C-terminus and two conserved sequences at the N-terminus ($IX_3EXEVMX_2W$ and $WX_3TX_2TX_3RLX_2K$) (Vats & Lee, 2001), have been identified in the deduced polypeptide sequence of pCT0017. This family of regulators are all homologues of the CopY copper responsive transcriptional repressor of the *cop* operon in *E. hirae* (Odermatt & Solioz, 1995) (refer to Section 1.8.2.1).

ORF pCT0018 shares significant similarity to hypothetical proteins of unknown function in *L. salivarius* strain UCC118 (Accession YP_536591) and *S. pneumoniae* R6 (Accession NP_358234). These proteins are located upstream of cation/copper transporting-ATPases. In addition, the deduced polypeptide sequence encoded by this ORF has a putative C-terminal CXMXMXH metal binding motif that may be involved in binding copper ions (Dancis *et al.*, 1994; Puig & Thiele, 2002).

ORF pCT0019 is located immediately upstream of *ctpA*. It encodes a small 79 amino acid protein sharing significant similarity to a small ORF located immediately upstream the

---

[3] A "pCT" prefix is used to conform to the naming scheme used in the GenBank submission of these ORFs (Accession Number U15554).

same a cation transporting P-type ATPase in *L. salivarius* strain UCC118 (Accession YP_536590) and the N-terminal region of the CtpA copper transporter of *S. pneumoniae* strain R6 (Accession NP_358235) (Bell, 2002). The arrangement of ORFs pCT0017, pCT0018, pCT0019 and *ctpA*, is consistent with genes upstream of cation/copper transporting-ATPases in *L. salivarius* strain UCC118 and *S. pneumoniae* R6 (Figure 1.9). This is of significance due to the fact like CtpA, both of these cation transporters also lack the consensus cation binding motifs typical of copper P-type ATPases.

Figure 1.10 shows a comparison of the organisation of the putative *cop*-like operon of DRDC8 comprising ORFs pCT0017, pCT0018, pCT0019 and *ctpA* (pCT0020), 15 *cop*-like operons in species of the *Lactobacillale* order characterised by Reyes *et al*., (2006) as well as the *cop* operon of *E. hirae*. *copY*-like genes and *copA*-like genes are found in most of the bacterial strains examined. CopY-like proteins invariably regulate expression of CopA proteins involved in copper uptake, however, the structure of the regulon is dependent on the organism of interest (Reyes *et al.*, 2006). On this basis, it is possible that ORFs pCT0017, pCT0018, pCT0019 and *ctpA* (pCT0020) comprise a *cop*-like operon regulated by the protein encoded by ORF pCT0017.

In addition to a putative CopY-like regulatory protein, a putative two-component regulatory system encoded by ORFs pCT0023 and pCT0024 (GenBank Accession U15554, presented in Appendix A, page 313) (refer to Table 1.5 for ORF description and Figure 1.8 for location) has been identified within DNA flanking *ctpA* (Webster, 2001). The deduced polypeptide sequence of pCT0023 shares significant identity to bacterial histidine kinases (HPK). In addition, this peptide sequence contains four of the five conserved domains characteristic of histidine kinases. pCT0024 is a partial ORF that shares significant similarity to DNA-binding response regulators (RR). The functional domains associated with response regulators were not located within the available pCT0024 sequence, however these domains are usually located at the N-terminal of the protein which is absent from region sequenced. Whilst there are many examples of bacterial copper homeostasis mechanisms that involve two-component regulatory systems (refer to Section 1.8.2.2), pCT0023 and pCT0024 do not share high similarity with two-component regulatory systems involved in copper ion homeostasis. However, in view of the proximity of the putative CopY and the putative two-component regulatory system to

the coding region of *ctpA*, either or both systems may function in regulation of expression of *ctpA*. Further investigation is required to determine the role of these ORFs (pCT0017, pCT0023 and pCT0024) in regulation of expression of *ctpA* in response to copper ion concentration. It also remains to be investigated if ORFs pCT0018 and pCT0019 encode Cop-like accessory proteins involved in *ctpA* mediated copper transport.

1.9.3 Distribution of *ctpA* in *L. monocytogenes* isolates

Given that expression of *ctpA* is important for pathogenicity and survival of *L. monocytogenes* during host infection, it is reasonable to assume that *ctpA* would be ubiquitously expressed in all *L. monocytogenes* strains. Intriguingly, data exists that indicates the prevalence and distribution of the *ctpA* gene is restricted to comparatively few *L. monocytogenes* isolates; generally of environmental origin (Francis, 1996; Webster, 2001). Studies by Bell, (2002) confirmed these observations; only 15 % (21/144) randomly selected *L. monocytogenes* isolates from environmental, laboratory and clinical sources tested positive for the *ctpA* gene by PCR (Table 1.6) and sequence analysis. From these results, it was clear that *ctpA* gene was not ubiquitously distributed amongst isolates of *L. monocytogenes* and its presence was typically limited to Australian non-clinically derived environmental isolates, mainly of dairy and poultry origin.

Figure 1.11 shows an example of typical PCR amplification of *ctpA* from different *L. monocytogenes* isolates (Bell, 2002). Interestingly, qualitative estimates of the amount of *ctpA* specific PCR product generated by some isolates appeared greater than that amplified by other *ctpA* positive strains when visualised by agarose gel electrophoresis. The fact that identical conditions and substrate concentrations were used for each reaction implied multiple copies of *ctpA* in selective *L. monocytogenes* isolates. This observation awaits further investigation.

1.9.4 *ctpA* is encoded on Plasmid DNA

Bell, (2002) proposed that *ctpA* is encoded on a large plasmid carried by *L. monocytogenes* strain DRDC8. Evidence to confirm this hypothesis was provided by differential PCR amplification of *ctpA* and the chromosomal marker gene *hly*, from serial dilutions of plasmid DNA extracted from *L. monocytogenes* strain DRDC8. The aim of this approach

was firstly to detect *L. monocytogenes* DNA contaminants and secondly to dilute chromosomal DNA contamination to extinction. Amplification of *ctpA* and *hly* was found to be differentially affected by dilution, where DNA preparations diluted to 1/1000 allowed amplification of *ctpA* but minimal dilution was required to eliminate amplification of the chromosomal marker *hly*. This data suggested that plasmid DNA was likely to be the template for amplification of *ctpA*.

Sequence analysis of clones carrying DNA fragments from *L. monocytogenes* strain DRDC8 encoding *ctpA* provided supporting data for this conclusion (Bell, 2002). A comparative analysis of the sequence data obtained showed that it shares a high level of sequence and gene organisation similarity to plasmid pLI100 from *L. innocua* CLIP11262 and plasmid pLM80 from *L. monocytogenes* strain H7858 (Bell, 2002). Specifically, five ORFs (pCT0001, pCT0002, pCT0003, pCT0004 and pCT0005 (GenBank Accession U15554, presented in Appendix A, page 302) (refer to Table 1.5 for ORF description and Figure 1.8 for location)) were identified that encode polypeptide sequences that share ≥ 93% sequence identity to proteins from plasmids pLI100 and pLM80. In addition, the analysis showed that *ctpA* and flanking ORFs (pCT0017, pCT0018, pCT0019) (refer to Section 1.9.2) was not identified in any of the *L. monocytogenes* genomic sequences available on the GenBank database (refer to Table 1.2). Other studies have also examined partial sequence data of DRDC8 DNA clones encoding *ctpA* (Accession U15554) (Akya, 2007; Francis, 1996; Webster, 2001). Collectively, ORFs encoding polypeptides that share similarity to proteins involved in; transcriptional regulation, DNA transposition, plasmid DNA damage repair mechanisms, plasmid DNA replication, and copper and cadmium ion transport, have been identified (refer to Table 1.5 for ORF description and Figure 1.8 for location).

In conclusion, ample evidence is available to suggest that *ctpA* is located on plasmid DNA in *L. monocytogenes* isolate DRDC8. Genes involved in metal ion homeostasis and resistance have been identified as a common theme in plasmid DNA in other *Listeria* species (refer to Section 1.7.2). Numerous examples of bacterial copper transport and resistance systems conferred by plasmids have been previously described, including in *P. syringae*, *E. coli, Xanthomonas and Enterococcus* (Brown *et al.*, 1995; Cooksey, 1993; Cooksey, 1994; Hasman, 2005; Lee *et al.*, 1994). Indeed, the limited distribution of *ctpA*

in environmental, mainly of dairy and poultry origin, Australian *L. monocytogenes* isolates may be explained by carriage of this gene on plasmid DNA. Indeed, several studies have shown that a limited number of clinical isolates carry plasmid DNA (refer to Section 1.7.2). Nevertheless, further investigation is required to show definitively that *ctpA* is encoded by plasmid DNA.

1.10 Hypothesis and Aims

The work described in this thesis was designed to identify and characterise genes involved in copper ion homeostasis and tolerance for *L. monocytogenes* strain DRDC8. The specific hypotheses addressed and the associated aims of the work were:

Hypothesis 1.   Plasmid DNA from strain DRDC8 carries *ctpA* and flanking ORFs that encode proteins involved in copper homeostasis.

    Aim 1:   Provide confirmatory evidence to demonstrate that the *ctpA* gene is encoded on plasmid DNA.

    Aim 2:   Identify additional accessory genes involved in copper homeostasis by further sequence analysis of plasmid DNA isolated from *L. monocytogenes* strain DRDC8.

Hypothesis 2.   *L. monocytogenes* strain DRDC8 carries chromosomal and plasmid genes that encode proteins involved in copper tolerance.

    Aim 1:   Provide evidence to show that ORFs pCT0017, pCT0018, pCT0019 and *cutR* are involved in copper tolerance by analysis of the effects of independent mutations (pCT0017::*erm*, pCT0018::*erm*, pCT0019::*erm* and *cutR*::*erm*) and loss of plasmid DNA for DRDC8 and a *cutR*::*erm* variant on the ability of *L. monocytogenes* to tolerate copper ion stress.

Hypothesis 3.   ORF pCT0017 encodes a CopY-like copper-responsive transcriptional repressor protein that regulates expression of *ctpA* and associated DNA.

Aim 1:   Provide experimental evidence to show that ORF pCT0017 encodes a CopY-like protein that is able to specifically bind *L. monocytogenes* DRDC8 DNA in a copper dependent manner.

Aim 2:   Characterise the DNA binding sites of pCT0017.

Table 1.1: Outbreaks of human food-borne listeriosis.

Country	Year	Food	Cases	Deaths	Serotype
USA	1976	Raw salad (?)	20	5	4b
New Zealand	1980	Fish (?)	22	7	1/2a
Canada	1981	Coleslaw	41	18	4b
USA	1983	Milk	49	14	4b
USA	1985	Soft cheese	142	30	4b
Switzerland	1983-1987	Soft cheese	122	34	4b
United Kingdom	1987-1989	Pate	355	94	4b
France	1993	Pork tongue in aspic	279	NK	4b
France	1993	Pork rillettes	38	10	4b
USA	1994	Milk	45	0	1/2b
Sweden	1994-1995	Fish	9	2	4b
France	1995	Soft cheese	17	4	4b
Canada	1996	Crab meat	2	0	1/2a
Italy	1997	Salad	1566	0	4b
USA	1998-1999	Hot dogs	50	>8	4b
Finland	1998-1999	Butter	25	6	3a
Finland	1999	Fish	5	NK	1/2a
France	1999-2000	Pork rillettes	10	2	4b
France	1999-2000	Pork tongue in jelly	32	10	4b
USA	2000	Turkey meat	29	7	NK
Switzerland	2005	Soft cheese	3	1	NK

NK: Not known

Table based on data from McLauchlin *et al*., (2004).

Table 1.2: Sequence data for *Listeria* species available in the GenBank database.

Organism	Serovar	Genome Accession Number	Plasmid Accession number/name	Reference
Listeria grayi DSM 20601	NA	ACCR00000000	NA	Qin *et al.*, (unpublished)
Listeria innocua Clip11262	6a	NC_003212	NC_003383 /pLI100	Glaser *et al.*, (2001)
Listeria monocytogenes EGD-e	1/2a	NC_003210	NA	Glaser *et al.*, (2001)
Listeria monocytogenes Finland 1988	NA	AART00000000	NA	Borowsky *et al.*, (unpublished)
Listeria monocytogenes str. 4b F2365	4b	NC_002973	NA	Nelson *et al.*, (2004)
Listeria monocytogenes str. 1/2a F6854	1/2a	AADQ01000008 AADQ01000065	NA	Nelson *et al.*, (2004)
Listeria monocytogenes str. 4b H7858	4b	AADR01000011 AADR01000054	AADR01000010 /pLM80 AADR01000058 /pLM80	Nelson *et al.*, (2004)
Listeria monocytogenes HCC23	4a	CP001175	NA	Steele *et al.*, (unpublished)
Listeria monocytogenes FSL F2-515	NA	AARI00000000	AARI01001405* AARI01001406* AARI01001407* AARI01001408* AARI01001409* AARI01001410* AARI01001411*	Birren *et al.*, (unpublished)

.

			AARI01001412* AARI01001413* AARI01001414* AARI01001415*	
Listeria monocytogenes FSL R2-503	NA	AARR00000000	AARR01000052* AARR01000053* AARR01000054*	Birren *et al*., (unpublished)
Listeria monocytogenes FSL R2-561	NA	AARS00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes FSL J1-194	NA	AARJ00000000	AARJ01000044*	Birren *et al*., (unpublished)
Listeria monocytogenes J0161	NA	AARW00000000	AARW01000050* AARW01000051*	Birren *et al*., (unpublished)
Listeria monocytogenes FSL N1-017	NA	AARP00000000	AARP02000075* AARP02000076* AARP02000077*	Birren *et al*., (unpublished)
Listeria monocytogenes HPB2262	NA	AATL00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes J2818	NA	AARX00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes F6900	NA	AARU00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes FSL N3-165	NA	AARQ00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes 10403S	NA	AARZ00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes FSL J1-175	NA	AARK00000000	NA	Birren *et al*., (unpublished)
Listeria monocytogenes FSL J2-003	NA	AARM00000000	NA	Birren *et al*., (unpublished)

Listeria monocytogenes FSL J1-208	NA	AARL00000000	NA	Birren *et al.*, (unpublished)
Listeria monocytogenes FSL J2-064	NA	AARO00000000	NA	Birren *et al.*, (unpublished)
Listeria monocytogenes FSL J2-071	NA	AARN00000000	NA	Birren *et al.*, (unpublished)
Listeria monocytogenes LO28	NA	AARY00000000	NA	Birren *et al.*, (unpublished)
Listeria seeligeri serovar 1/2b str. SLCC3954	1/2b	NC_013891	NA	Steinweg *et al.*, (2010)
Listeria welshimeri serovar 6b str. SLCC5334	6b	NC_008555	NA	Hain *et al.*, (2006b)

NA: Not available

* Since completion of this thesis, this record has been removed and superceded by a new assembly.

Table 1.3: Detection of plasmid DNA in isolates of *Listeria monocytogenes*.

Serovar	Origin	Total number of cultures	Plasmid DNA detected (%)
1/2a	Clinical	77	30(39)*
	Food	21	17(81)*
1/2b	Clinical	44	27(61)
	Food	6	3(50)
1/2c	Clinical	11	10(90)**
	Food	6	2(33)**
4b	Clinical	113	10(9)
	Food	15	3(20)
Other	Clinical	18	4(22)
	Food	11	5(45)
Total		**322**	**111(34)**

Fisher's exact test: $*P$=0.001, $**P$=0.028

Table based on data from McLauchlin *et al*., (1997).

Table 1.4: Summary of proteins encoded by *Listeria* plasmid genes.

Protein function	Species (Plasmid)	
	L. innocua (pLI100)	*L. monocytogenes* strain 4b H7858 (pLM80)
Hypothetical/unknown	20	56
DNA Transposition/Recombination	19	13
Transport/Resistance	13	5
Plasmid DNA replication	7	3
Transcriptional regulation	6	3
DNA restriction endonuclease/methylase	2	6
DNA repair	1	1
Competence/conjugation	1	1
Other	4	5
Total	**80**	**93**

Table 1.5: ORFs flanking *ctpA* for *L. monocytogenes* strain DRDC8.

Putative ORF[4]	Putative Protein Function	Reference
pCT0001	UV-damage repair protein (partial ORF)	Bell, (2002)
pCT0002	Hypothetical protein	Bell, (2002)
pCT0003	Hypothetical protein	Bell, (2002)
pCT0004	DNA transposition-associated protein	Bell, (2002)
pCT0005	Transposase (partial ORF)	Bell, (2002)
pCT0008	Cadmium efflux ATPase	Akya, (2007)
pCT0009	Transposase	Akya, (2007)
pCT0010	Hypothetical protein	Akya, (2007)
pCT0011	Transposase	Akya, (2007)
pCT0012	Restriction modification system	Akya, (2007)
pCT0013	Plasmid DNA replication protein	Akya, (2007)
pCT0014	Plasmid DNA replication protein	Akya, (2007)
pCT0017	CopY negative transcriptional regulator	Bell, (2002)
pCT0018	Hypothetical Protein	Bell, (2002)
pCT0019	Hypothetical Protein	Bell, (2002)
pCT0020	Copper-translocating P-type ATPase (CtpA)	Francis, (1996)
pCT0021	Transposase	Francis, (1996)
pCT0022	Hypothetical Protein	Webster, (2001)
pCT0023	Two-Component Sensor Histidine Kinase	Webster, (2001)
pCT0024	Two-Component Response Regulator (partial ORF)	Webster, (2001)

---

[4] A "pCT" prefix is used to conform to the naming scheme used in the GenBank submission of these ORFs (Accession Number U15554).

Table 1.6: Distribution of *ctpA* positive *L. monocytogenes* isolates identified by PCR.

Source of *L. monocytogenes* isolate	No. positive/No. tested (% positive)
Poultry	4/55 (7)
Dairy	12/54 (22)
Clinical	1/28 (3)
SLCC[*]	4/7 (57)
Total	**21/144 (15)**

Table from Bell, (2002).

[*] Seeliger's *Listeria* Culture Collection.

Figure 1.1: The intracellular life cycle of *L. monocytogenes* in a host cell.

The genes required at each stage of the infectious process are indicated. Abbreviations: *inlAB* locus, internalin; *hly*, listeriolysin O; *plcA*, phospholipase C; *actA*, actin polymerisation; *mpl*, metalloprotease; *plcB*, phosphatidylcholine-phospholipase C (lethicinase). Shading represents initial nucleation and polymerisation of host cell actin. Figure adapted from Tilney and Portnoy, (1989).
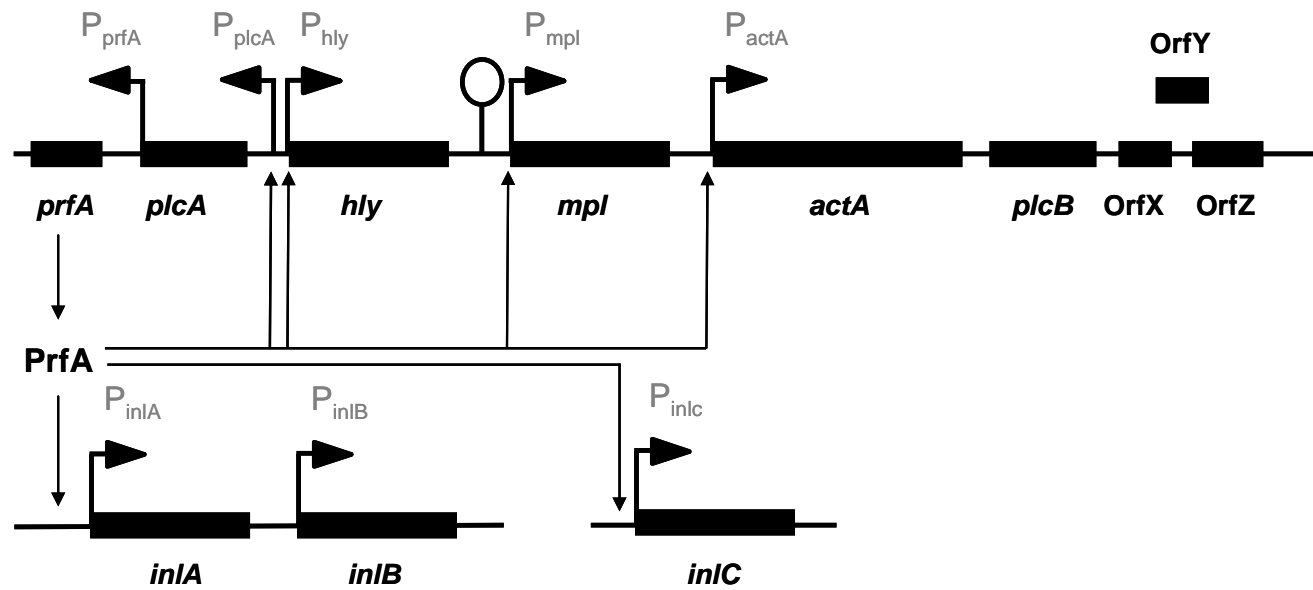
Figure 1.2: Genetic organisation of the *L. monocytogenes* pathogenicity island.

Figure from Cossart, (2002).

Figure 1.3: Model of the regulation of the *cop* operon in *E. hirae*.

Figure reproduced from Odermatt and Solioz, (1995).
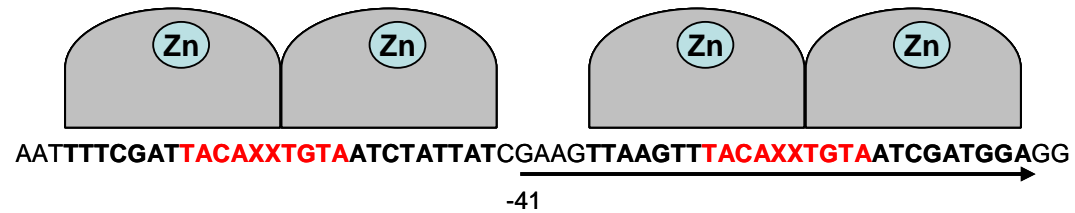
Figure 1.4: Schematic representation of the [Zn(II)CopY]$_2$ - DNA interaction with the *E. hirae cop* operon promoter region.

The conserved *cop* box inverted repeats (TACAXXTGTA) are shown in red. A [Zn(II)CopY]$_2$ dimer binds to each of the two *cop* boxes and protects the DNA regions shown in bold. Transcription begins at nucleotide -41, indicated by the arrow.

Figure adapted from Portmann *et al.*, (2004).
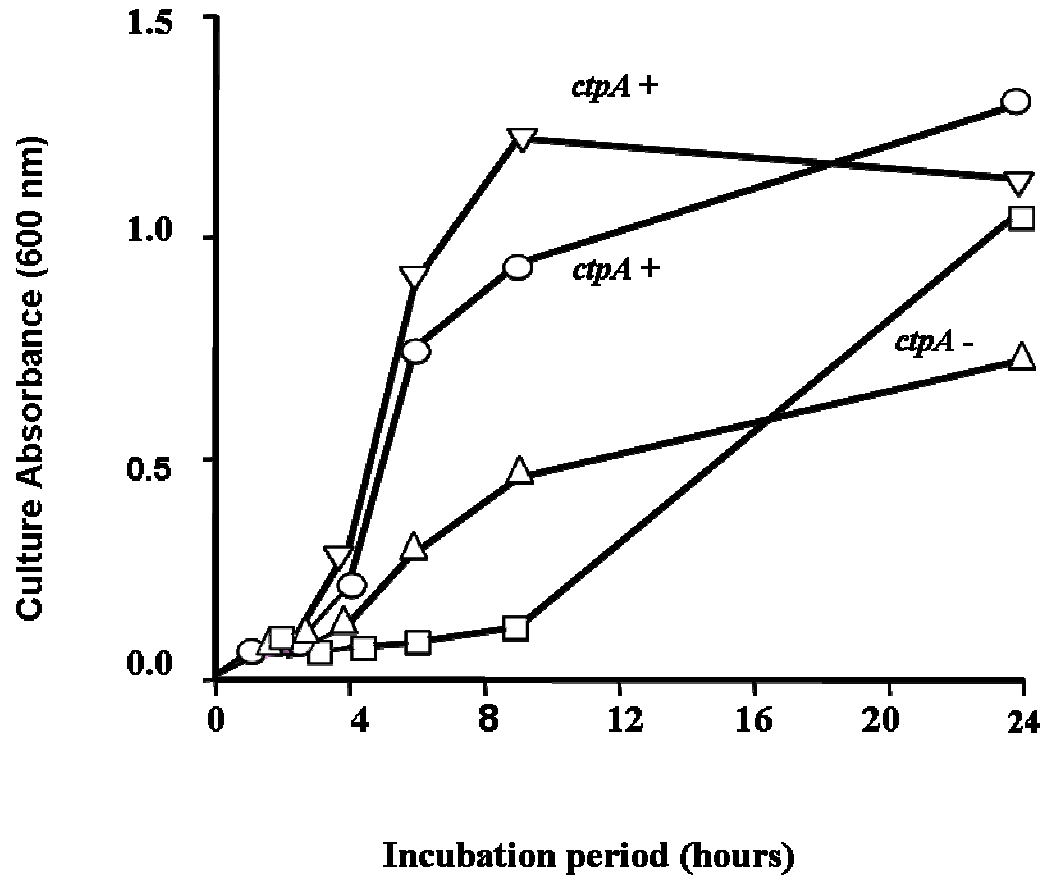
Figure 1.5: Growth of *L. monocytogenes ctpA* mutants in BHI broth containing 5μM of the cation chelating agent, 8-hydroxyquinoline.

Figure adapted from Francis and Thomas, (1997a).

Figure 1.6: Growth of *ctpA* positive and *ctpA* negative *L. monocytogenes* isolates in BHI broth containing 10mM CuSO$_4$.

Figure from Mok, (2003).
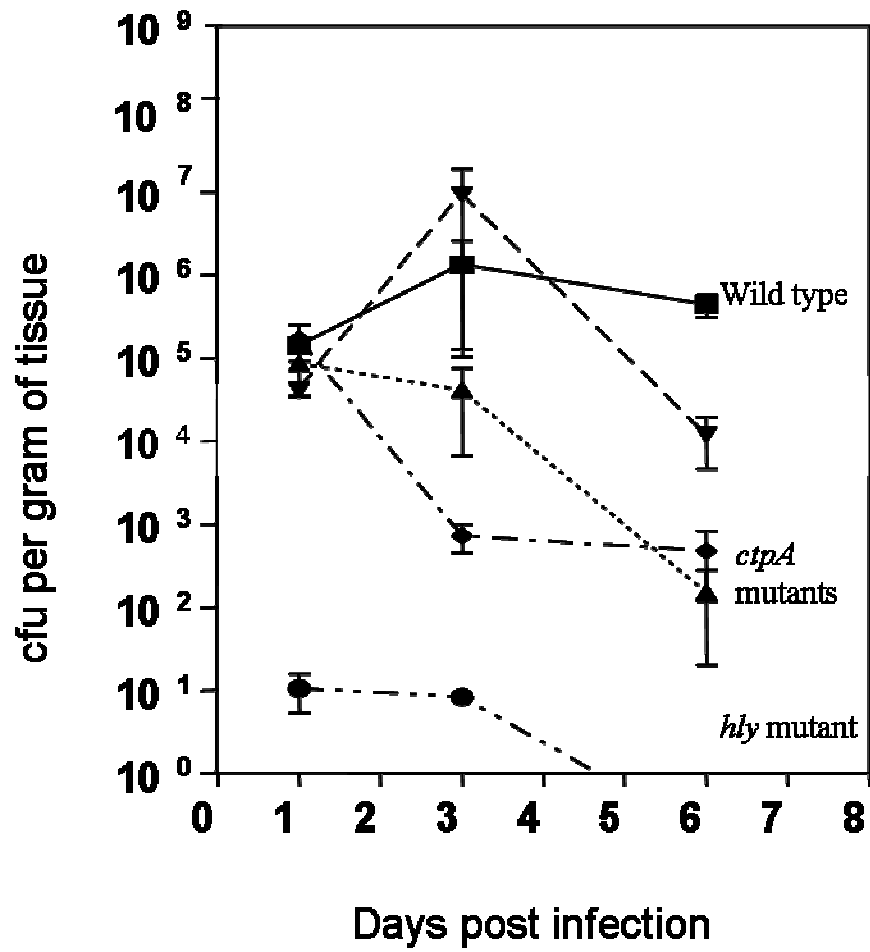
Figure 1.7: Course of infection by _L. monocytogenes ctpA_ mutants in the liver of mice.
Figure from Francis and Thomas, (1997b).
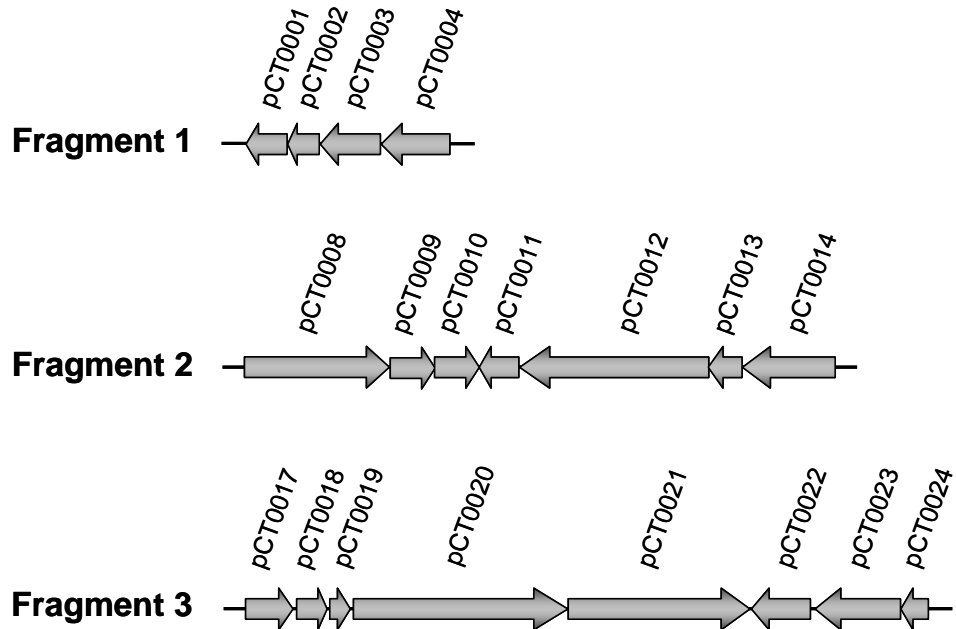
Figure 1.8: Genetic organisation of putative ORFs encoded by DRDC8 plasmid DNA.

Three non-contiguous fragments of sequence data generated by Francis, (1996); Webster, (2001) and Bell, (2002). The orientation of each putative ORF is shown accordingly. Refer to for putative ORF description.
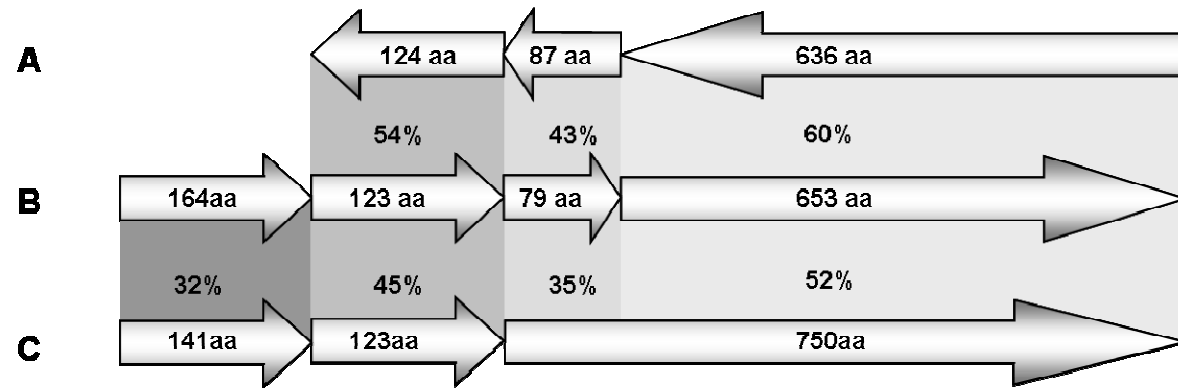
Figure 1.9: Comparison of the genetic arrangement of ORFs pCT0017, pCT0018, pCT0019 and *ctpA* of *L. monocytogenes* strain DRDC8 and genes that encode similar proteins in *Lactobacillus salivarius* subsp. UCC118 and *Streptococcus pneumoniae* R6.

A. Genetic arrangement of genes encoding two hypothetical proteins (124 aa and 87 aa) and a cation transporting P-type ATPase (636 aa) in *Lactobacillus salivarius* subsp. UCC118.

B. Genetic arrangement of ORFs pCT0017 (164 aa), pCT0018 (123 aa), pCT0019 (79 aa), and *ctpA* (653 aa) in *L. monocytogenes* strain DRDC8 the CtpA copper-transporting P-type ATPase.

C. Genetic arrangement of genes encoding a COPAB ATPase metal-fist type repressor (141 aa), a hypothetical protein (123 aa) and a P-type ATPase copper transporter (750 aa) in *Streptococcus pneumoniae* R6.

The percentage of amino acid sequence similarity shared between the polypeptides encoded by pCT0017, pCT0018, pCT0019 and *ctpA* and proteins from *L. salivarius* UCC118 and *S. pneumoniae* R6 are indicated between the respective genes/ORFs. Similar proteins are identified by the shaded regions.
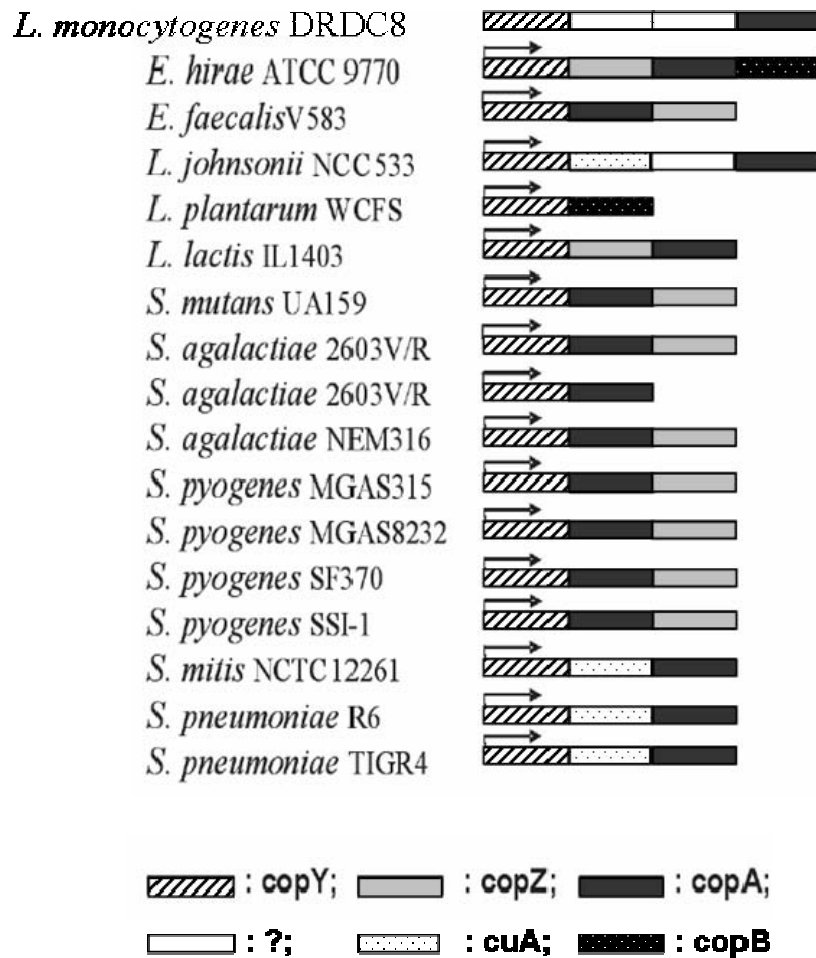
Figure 1.10: Organisation of *cop*-like operons in different Gram-positive bacteria.

The organisation of the putative *cop*-like operon of *L. monocytogenes* strain DRDC8 is compared with *cop*-like operons in 15 strains (10 species) of *Enterococcus*, *Lactobacillus* and *Streptococcus* as characterised by Reyes *et al.*, (2006) and the *cop* operon of *E. hirae*. Arrows represent transcriptional start sites. The names of the genes within the operons are indicated; genes of unknown function are indicated by a question mark.

The order of ORFs of *L. monocytogenes* strain DRDC8 is pCT0017, pCT0018, pCT0019 and pCT0020. ORF pCT0017 is predicted to encode a *copY* homolog. The function of ORFs pCT0018 and pCT0019 is unknown. ORF pCT0020 (*ctpA*) is a *copA*-like gene.

Figure modified from Reyes *et al.*, (2006).
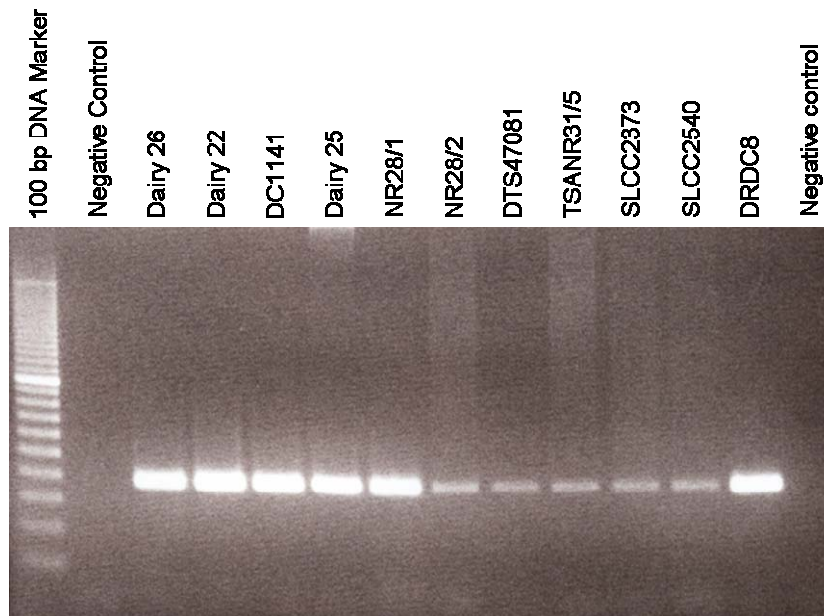
Figure 1.11: Typical PCR amplification of *ctpA* from *L. monocytogenes* isolates.

396 bp PCR products indicative of *ctpA* were amplified from DNA extracted from various *L. monocytogenes* isolates using the FB001/LM2004 oligonucleotide pair. Variation in DNA band intensity is obvious between some strains analysed. Figure from Bell, (2002).

Chapter 2: Materials and Methods

2.1 Chemicals and Reagents

All chemicals were AnalaR grade or higher and were purchased from Sigma Chemical Company (St Louis, Missouri, USA), Ajax Chemicals (Auburn, New South Wales, Australia), BDH Laboratory Supplies (Poole, Dorset, England) or Bio-Rad (Richmond, California, USA), unless otherwise stated. All molecular weight markers, enzymes and immuno-conjugants were purchased from New England Biolabs (Beverly, Massachusetts, USA), Roche Diagnostics (GmbH, Germany), Promega Corporation (Madison, USA), Sigma-Aldrich (St Louis, Missouri, USA), Applied Biosystems (Foster City, California, USA) and Difco (Detroit, Michigan, USA), unless otherwise stated.

2.2 Bacterial strains and Plasmids

All *Listeria* spp. used in this study are listed in Table 2.1. *Escherichia coli* strains used are listed in Table 2.2. Plasmid vectors and clones in this study are described in Table 2.3.

2.3 Bacterial Growth Media

Listeria spp. were cultured in Brain Heart Infusion (BHI) broth (Difco) or *Listeria* enrichment broth (LEB) base [30% (w/v) Bacto Tryptic Soy broth (Difco) and 6% (w/v) Bacto Yeast extract (pH 7.3 ± 0.2) (Difco)]. BHI agar (Bacto Agar BBL technical grade, 15 g L^{-1}), or LEB agar (Bacto Agar BBL technical grade, 15 g L^{-1}) was used as solid media for standard cultivation.

E. coli strains were typically cultivated in Luria Broth (LB) or Nutrient broth (NB). LB consisted of Bacto-Tryptone (Difco) (10 g L^{-1}), Bacto Yeast extract (Difco) (5 g L^{-1}), and NaCl (5 g L^{-1}) as described by Miller (1972). NB consisted of Lab Lemco (Oxoid Ltd., London, England) (10 g L^{-1}), Bacto peptone (Oxoid) (10 g L^{-1}) and NaCl (5 g L^{-1}). Solid media used was Nutrient agar (NA) or Luria agar (LA), consisting of NB or LB, respectively, with the addition of agar (BBL technical grade) (15 g L^{-1}).

All bacterial strains were incubated at 37°C, unless otherwise indicated and liquid cultures were grown in 20 mL McCartney bottles or in 50 - 2000 mL flasks with aeration unless otherwise stated.

Where appropriate, antibiotics were added to broth and solid culture media at the following final concentrations: Ampicillin (Amp), 100 μg mL^{-1}; chloramphenicol (Cm), 25 μg mL^{-1}; kanamycin (Kan), 50 μg mL^{-1}; streptomycin (Sm), 50 μg mL^{-1}; spectinomycin (Sp), 100 μg mL^{-1}; rifampicin (Rif), 100 μg mL^{-1}. Erythromycin (Em) was added to media used to culture *E. coli* at 150 μg mL^{-1}, and 8 μg mL^{-1} for media used to culture *L. monocytogenes*.

Cultures of *L. monocytogenes* used for tissue culture infection assays were grown on BHI media supplemented with 0.2% (w/v) activated charcoal as described by Ripio *et al.*, (1996) and Vega *et al.*, (2004).

X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) medium was used to identify colonies containing DNA inserts cloned into plasmid vectors encoding for *lac*Z blue/white colour selection. This medium was composed of LA, X-gal (10 μg mL^{-1}) and IPTG (isopropyl-β-D-thiogalactopyranoside) (24 μg mL^{-1}).

2.4 Maintenance of Bacterial strains

Stock cultures were maintained and stored as concentrated bacterial suspensions in Wheaton vials (Millville, New Jersey, USA) in a 1 mL solution of glycerol (30% (v/v)) and Bacto Peptone (1% (w/v) (Difco)) at -80°C. Fresh cultures of *Listeria* and *E. coli* strains were prepared by streaking a loopful of frozen glycerol stock onto the appropriate medium and incubated O/N at the optimal temperature.

For routine use, cultures of bacterial strains were maintained on plates of nutrient medium at 4°C.

2.5 Preparation of Tris-HCL Buffered Phenol

An equal volume of 1M Tris-HCL, (pH 8.0) and 500 mg of 8-hydroxyquinoline was added to 500 g of melted phenol, mixed thoroughly and allowed to equilibrate. The upper

aqueous phase was removed by aspiration and further equilibrated twice with equal volumes of 1M Tris-HCL (pH 8.0), followed by three additions of an equal volume of 0.1 M Tris HCL (pH 8.0). Approximately 1/5 volume of 50 mM Tris-HCL (pH 8.0) was used to cover the phenol. This was stored at 4°C in a foil-covered container.

2.6 Centrifugation

The following conditions were used unless otherwise stated. Cells or DNA were collected in small volumes by centrifugation in 1.5 mL reaction tubes (Sarstedt Australia Pty Ltd, Mawson Lakes, South Australia) in an Eppendorf centrifuge 5417R (Netheler, Hinz GmbH Hamburg, Germany) at 20,000 × g for 1 min at 4°C. Alternatively, centrifugation was performed in sterile glass 10 mL McCartney bottles, 5 mL or 10 mL plastic conical tubes (Sarstedt Australia Pty Ltd, Mawson Lakes, South Australia) or 50 mL plastic centrifuge tubes (Becton Dickinson Labware, Franklin Lakes, NJ USA) in a Sigma CE 4-15 Laboratory Centrifuge (Quantum Scientific, Queensland, Australia) at 3,000 × g for 10 min. Larger volumes (50 mL to 500 mL) were centrifuged using a Beckman J2-21M ultracentrifuge (GMI Minnesota, USA) in JA-20, JA-14 or JA-10 rotor at the maximum speed specified for the rotor in use. CsCl gradient centrifugation was performed in 5.1 mL Quick-Seal® ultracentrifuge tubes in a TLA-100.4 or an SW40 rotor in a Beckman Optima^TM TLX Ultracentrifuge (GMI).

2.7 DNA Extraction Procedures

2.7.1 Bacterial Plasmid DNA Isolation

2.7.1.1 Small Scale Preparations

Method 1:   Small scale quantities of plasmid DNA (3 to 5 μg mL^{-1}) were purified from *E. coli* by a modification of the three step alkali lysis method (Sambrook *et al.*, 1989). An O/N culture (1.5 mL), was collected by centrifugation and the pellet resuspended in 100 μL of Solution 1 [50 mM glucose, 25 mM Tris-HCL (pH 8.0), 10 mM EDTA, 5 μg mL^{-1} RNase A (Sigma)]. Following the addition of 200 μL of Solution 2 [0.2 M NaOH, 1% (w/v) SDS], the sample was inverted gently and incubated on ice for 5 min. The addition

of 150 μL of Solution 3 [5 M potassium acetate 2 M acetic acid (pH 4.8)] was also followed by 5 min incubation on ice. Cell debris was collected by centrifugation (20,000 × g, 10 min) and supernatant extracted with an equal volume of Tris-HCL buffered phenol (refer to Section 2.5) in a fresh tube. The aqueous phase was extracted with an equal volume of phenol:chloroform (1:1) solution and finally with chloroform. The aqueous phase was transferred to a fresh reaction tube and the plasmid DNA precipitated on ice in two volumes of cold 100% ethanol. DNA was collected by centrifugation (20,000 × g, 10 min), and the DNA pellet washed with 70% (v/v) ethanol and dried *in vacuo* (Speed Vac SVC100, Savant Instruments Inc., Farmingdale, New York, USA). The DNA pellet was then resuspended in 1 × TE and stored at 4°C.

Method 2: Small scale plasmid DNA isolation from *E. coli* strains was performed using the QIAprep Spin Miniprep Kit (QIAGEN, Pty Ltd, Doncaster, Victoria, Australia), according to the manufacturer's instructions.

2.7.1.2 Large Scale Preparations

Method 1: Large scale plasmid DNA isolation from *L. monocytogenes* was performed by a modification of the Birnboim and Doly, (1979) technique. Briefly, 250mL of O/N culture was pelleted by centrifugation in a Beckman J2-21M ultracentrifuge (JA-10 rotor, 11,000 × g, 10 min). The pellet was washed in 12.5 mL of 0.1 × SSC [15 mM NaCL, 1.5 mM sodium citrate], resuspended in 11.25 mL of lysozyme solution [0.01 M sodium phosphate buffer pH 7.0, 20% (w/v) sucrose and 5 mg mL^{-1} lysozyme], followed by incubation at 37°C for a minimum of 45 min. Plasmid DNA was then extracted by a modification of the three step alkali lysis method (Section 2.7.1.1, Method 1), using 3 mL of Solution 1 [50 mM glucose, 25 mM Tris-HCL (pH 8.0), 10 mM EDTA, 5 μg mL^{-1} RNase A (Sigma)], 5.5 mL of Solution 2 [0.2 M NaOH, 1% (w/v) SDS] and 2.8 mL of Solution 3 [5 M potassium acetate 2 M acetic acid (pH 4.8)]. The cell debris was removed by centrifugation (JA-20 rotor, 15,000 × g, 20 min) and the supernatant extracted with an equal volume of Tris-HCL buffered phenol in a fresh tube. The aqueous phase was extracted with an equal volume of phenol:chloroform (1:1) solution and finally with chloroform. The DNA was transferred to a fresh reaction tube, precipitated as described in Section 2.8.6, then resuspended in 1 × TE and stored at 4°C.

This method was also scaled up to extract large plasmid DNA from 2 L of O/N culture. Harvested cells were washed in 100 mL of 0.1 × SSC and resuspended in 90 mL of lysozyme solution. DNA was extracted from cells extracted with 50 mL of Solutions 1, 2 and 3. The quality and quantity of plasmid DNA preparations was analysed by electrophoresis in 1% (w/v) agarose gels (Section 2.8.3).

Method 2: Isolation of *L. monocytogenes* plasmid DNA was performed using gentle shearing of chromosomal DNA and precise neutralisation of alkali denatured DNA by a modification of the method previously described by Currier and Nester, (1976). Briefly, 500 mL of BHI broth was inoculated with 10 mL of O/N bacterial culture and was incubated at 37°C with agitation for 16 h. The cells were pelleted by centrifugation, washed twice and resuspended in 25 mL of TE buffer [50 mM Tris-HCL (pH 8.0), 20 mM EDTA]. Lysozyme was added to a final concentration of 5 mg/mL and incubated at 37°C for 30 min. Following the addition of a further 75 mL of TE buffer, Pronase and SDS were added to the final concentrations of 500 μg mL^{-1} (w/v) and 1.0% (v/v), respectively and incubated at 37°C for 30 - 45 min. The viscosity of the lysate was then reduced by gentle shearing in a Sorvall Omnimixer for 2 min. The pH of the lysate was adjusted to 12.1 - 12.3 by addition of 3 M NaOH solution with gentle stirring for 10 min. The pH was then reduced to 8.5 - 9.0 by the addition of 2 M Tris-HCL pH 7.0, stirred for an additional 3 min and adjusted to 3% (w/v) NaCl. An equal volume of buffered phenol (refer to Section 2.5) was added and the lysate stirred for 5 min at 100 - 150 rpm. The aqueous phase was separated by centrifugation (JA-10 rotor, 11,000 × g, 5 min), followed by removal of residual phenol by extraction with an equal volume of chloroform:isoamyl alcohol (24:1). The aqueous phase was concentrated by the addition of 15 mM MgCl$_2$, 5 mM sodium phosphate buffer and 0.7 vol of cold 95% (v/v) EtOH. Following incubation at -20°C O/N, the precipitated DNA was collected by centrifugation, washed in 10 mL of 70% (v/v) ethanol and resuspended in 2 mL of 1 × TE buffer.

Plasmid DNA was purified by caesium chloride-ethidium bromide density gradient centrifugation. To the DNA, 4 g of CsCl$_2$ and 0.4 mL of Ethidium bromide (EtBr) solution (10 mg mL^{-1}) were added and the mixture transferred to a 5 mL Quick-Seal$^®$ ultracentrifuge tube. The tube was sealed using a Beckman heat sealing device and centrifuged in a Beckman Optima^{TM} TLX Ultracentrifuge (GMI) (TLA-100.4 rotor,

240,000 × g, 17 h, 4ºC). The plasmid DNA band was extracted using a sterile 18 gauge syringe. An equal volume of isoamyl alcohol was added and mixed by inversion to extract the EtBr. The aqueous layer was removed and the extraction repeated twice. The plasmid DNA preparation was then dialysed in 1 × TE buffer and the $CsCl_2$-free DNA ethanol/glycogen precipitated and resuspended in 1 × TE. Plasmid DNA digested with restriction endonucleases and the fragments separated by agarose gel electrophoresis in a 1% (w/v) agarose gel in TAE buffer (see Section 2.8.3).

2.7.2 Preparation of Bacterial Genomic DNA

Method 1:   Genomic DNA from *L. monocytogenes* was isolated by a modification of the method previously described by Flamm *et al.*, (1984). Briefly, 10 mL of O/N culture was pelleted by centrifugation (JA-20 rotor, 15,000 × g, 7 min,) and the cells washed in 1 mL of 0.1 × SSC [15 mM NaCl, 1.5 mM sodium citrate].   Washed cells were then resuspended in 0.6 mL of lysozyme solution [0.01 M sodium phosphate buffer (pH 7.0), 20% (v/v) sucrose and 2.5 mg mL^{-1} lysozyme] and incubated for 1 h at 37°C. Pronase solution (5.4 mL) [10 mM Tris-HCL (pH 8.0), 1 mM EDTA, 1% (v/v) SDS and 0.5 mg/mL of pronase] was added and incubated for a minimum of 1 h at 37°C. The solution was extracted with an equal volume of phenol, followed by phenol:chloroform (1:1) and finally with chloroform. The DNA was precipitated as described Section 2.8.6.

Method 2:   *Listeria monocytogenes* genomic DNA was isolated using the Wizard® Genomic DNA Purification Kit (Promega), protocols as per manufacturer's instructions.

Method 3:   Crude extracts of genomic DNA from *E. coli* or *L. monocytogenes* was isolated from a cell suspension by the following method. Briefly, loopfuls of cells grown on the appropriate solid growth medium were resuspended in either 500 μL of Milli-Q H_2O or 0.85% (w/v) saline in a 1.5 mL reaction tube. The tube was mixed by vortexing, then heated in a boiling water bath for 3 min, cooled on ice, and the cell debris pelleted by centrifugation. The supernatant was removed and (2 μL) used as a crude source of chromosomal DNA.

2.7.3 Bacteriophage DNA extraction

To 1 mL of bacteriophage preparation (see Section 2.22.1), a solution containing ETDA (20mM), Pronase (0.5 mg/mL), and SDS (0.5% (v/v)) was added and mixed by inversion. Following incubation at 37°C for 1 h, an equal volume of phenol was added and mixed by vigorous inversion for 30 sec. The aqueous phase was separated by centrifugation (20,000 $\times$ g, 5 min), mixed with an equal volume of phenol:chloroform (1:1) and spun for a further 5 min. Residual phenol was extracted with an equal volume of chloroform. The aqueous phase was separated by centrifugation and the DNA precipitated by the addition of 0.1 volume of 3M sodium acetate (pH 5.2) and 0.6 volume of 100% isopropanol. Following incubation at -20°C for 1 h, the DNA was collected by centrifugation, washed in 70% (v/v) ethanol and resuspended in 1 $\times$ TE Buffer. The DNA was then analysed by gel electrophoresis.

2.8 DNA Analysis and Manipulation

2.8.1 DNA Quantification

The concentration of DNA was determined by measuring the A_{260} of the sample using a BioRad Smart Spec^{TM} 3000 (BioRad Laboratories Pty Ltd., NSW, Australia). Samples with an A_{260} of 1.0 were assumed to contain 50 µg dsDNA mL^{-1}. Agarose gel electrophoresis was also used to qualitatively estimate DNA concentration by comparison of the fluorescence intensity of DNA fragments stained with EtBr to that obtained from DNA fragments of known concentration.

2.8.2 Restriction Endonuclease Digestion of DNA

Cleavage of DNA with restriction enzymes was performed using enzymes according to the reaction conditions and supplied buffers as specified by the manufacturer (New England Biolabs).

Restriction endonuclease digests involving more than one enzyme, were performed either simultaneously or sequentially, following DNA precipitation (refer to Section 2.8.6), depending on the compatibility of buffering conditions.

2.8.3 Agarose Gel Electrophoresis of DNA

Separation and visualisation of DNA fragments by electrophoresis was performed at room temperature on horizontal 0.6% - 2.0% (w/v) agarose gels, the percentage dependent upon the size of the fragment of DNA to be analysed. Agarose gels (Seakem) were prepared in $1 \times$ TAE buffer [40 mM Tris, 1.5 mM EDTA, 0.012% (v/v) glacial acetic acid] and allowed to set at RT in a gel casting mould. Tracking dye (0.1 vol) [20% glycerol (v/v), 0.01% bromophenol blue (w/v) or 0.6% sucrose (w/v), 0.01% (w/v) bromophenol blue)] was added to DNA samples and the mixture loaded onto the gel. Gels were electrophoresed in horizontal electrophoresis gel tanks (EASY-CAST^{TM}, Owl Scientifics, Inc.) in $1 \times$ TAE buffer [40 mM Tris, 1.5 mM EDTA, 0.012% (v/v) glacial acetic acid] at a maximum of 100 V for 1 - 3 h. DNA was visualised by trans-illumination on a UV transilluminator (UVP Inc., Upland, California USA) after staining with an aqueous solution of EtBr (2 μg mL^{-1}) for 10 - 15 min. Stained gels were documented using a Tractel GDS 2 Gel Documentation System (Vision Systems, Salisbury, South Australia, Australia).

2.8.4 Determination of DNA Fragment Size

The size of plasmids, DNA fragments and PCR products were calculated by the comparison of their relative mobility to several different DNA markers. Markers used are shown in Table 2.4.

2.8.5 Isolation of DNA Fragments from Agarose Gels

Method 1:   DNA was extracted from 1% (w/v) agarose gel slices using the QIAGEN QIAEX$^®$ II Gel Extraction Kit (QIAGEN) or the QIAquick Gel Extraction Kit according to the manufacturer's instructions.

Method 2:   The DNA was electrophoresed on 1.5% (w/v) Seakem low melting temperature agarose gel (FMC BioProducts, Rockland, Maine, USA) at a maximum of 80V. Following excision from the gel, the gel slice was placed in an Eppendorf tube and heated to 65°C for 5 - 10 min in $1 \times$ TE ($2 \times$ volume of gel slice). The mix was extracted by the addition of an equal volume of phenol. This extraction procedure was repeated on the aqueous phase and the DNA precipitated and resuspended in $1 \times$ TE buffer

2.8.6 Precipitation of DNA

All DNA was precipitated with 100% isopropanol, ethanol acetate [100% ethanol, 1M sodium acetate (pH 5.2)] or 100% ethanol. In some cases precipitation was assisted by the addition of 0.025 vol of 50 mg mL^{-1} (w/v) glycogen and incubation at -20°C for at least 30 min. In each case the precipitated DNA was collected by centrifugation, (20,000 × g, 30 min, 4°C), and the pellet washed with 70% (v/v) ethanol, dried *in vacuo* or air dried and then resuspended in the appropriate volume of 1 × TE buffer.

2.8.7 Dephosphorylation of DNA

Antarctic phosphatase (New England Biolabs) treatment of restriction digested vector DNA was performed to dephosphorylate the ends of vector DNA linearised by digestion with restriction endonucleases. Reactions (10 μL) consisted of vector DNA (1 μg), 1 × Antarctic Phosphatase Reaction Buffer (supplied by the manufacturer) and 5 units of Antarctic Phosphatase. The reaction was incubated at 37°C for 15 min. Phosphatase was inactivated by heating at 65°C for 5 min.

2.8.8 Ligation Reactions

Ligation reactions were generally performed in 20 μL volumes using T4 DNA ligase (New England Biolabs). Vector and insert DNA were mixed in a reaction tube at a ratio of *ca.* 1:3, and where necessary, Milli-Q H$_2$O was added to make the final volume. 2 μL of 10 × ligation buffer (supplied by the manufacturer) and T4 DNA ligase (1 units) were added to the reaction and incubated at 15°C for approximately 2 h or left at 4°C ON. For blunt end ligations, 4 units of T4 DNA ligase was used per reaction and PEG$_{600}$ (15% w/v) was added to the reaction to assist in ligation. PEG$_{600}$ was prepared as a 60% (w/v) stock solution and autoclaved prior to use.

Ligations involving insertions into pET21 vectors consisted of 50 - 100 ng of restriction endonuclease digested pET vector, 2 μL of 10 × Ligase buffer, 2μL of 100 mM DTT, 1 μL of 10 mM ATP, 1 μL of T4 DNA ligase (0.2 - 0.4 Weiss units/μL), 0.2 pmol target gene insert and MilliQ water to a final volume of 20 μL.

2.8.9 *In vitro* Cloning

Unless otherwise specified, cloning of purified PCR products (refer to Section 2.15) or DNA fragments was achieved using the pGEM®-T Easy Vector System (Promega Corporation, Madison, USA), as per the manufacturer's protocols. Plasmid DNA was used to transform competent *E. coli* strains (Table 2.2) using chemical transformation methods (refer to Section 2.9.2).

2.9 Chemical Transformation of *E. coli*

2.9.1 Preparation of Competent Cells

Competent cells were prepared by a modification of the method described by Hanahan, (1983). An O/N culture of *E. coli* DH5-α cells was diluted 1/20 into 50 mL of LB broth and was grown with agitation at 37°C to an OD_{600} reading of 0.6. The culture was incubated on ice for 20 min and centrifuged (JA-14 rotor, 15,000 × g, 7 min) and the cell pellet resuspended in 10 mL of cold 100 mM $MgCl_2$. The solution was centrifuged (JA-14 rotor, 15,000 × g, 7 min) and the cells resuspended in 2 mL of cold 100 mM $CaCl_2$, followed by incubation on ice for a minimum of 1 h. The cells were used immediately.

2.9.2 Transformation Procedure

Transformation was performed by the addition of pre-chilled DNA to the cells and incubation on ice for 1 h. The cells were then heat shocked at 42°C for 2 min and incubated immediately on ice for 5 min. LB broth (750 µL) was added to the cells followed by incubation at 37°C for 1 h with gentle agitation. The cells were collected by centrifugation, and resuspended in a small quantity of supernatant fluid. Unless otherwise specified, transformants were recovered by plating suitable dilutions of cells on LA plates containing Ampicillin, IPTG (24 mg mL^{-1}) and X-Gal (50 mg mL^{-1}).

2.10 Electro-transformation of *Listeria monocytogenes*

2.10.1 Preparation of Electro-competent Cells

Electroporation-mediated transformation of *L. monocytogenes* cells was performed by a modification of the method previously described by Park and Stewart, (1990). Briefly, 100 mL of BHI containing sucrose (0.5 M) was inoculated with 15 mL of O/N culture and incubated at 37°C with shaking until an A_{600} of 0.2 was reached. Penicillin G (10 µg mL^{-1}) was added the culture and incubated for a further for 2 h. Bacterial cells were then harvested by centrifugation, washed once with 100 mL of electroporation buffer (1 mM HEPES pH 7.0, 0.5 M sucrose), twice with 50 mL of buffer and then resuspended in 2.5 mL of buffer. All suspensions of electro-competent cells were used within 30 min of preparation.

2.10.2 Electroporation Procedure

Plasmid DNA (1 µg) was mixed with 100 µL of electro-competent *L. monocytogenes* cells (containing approximately 1×10^{11} cell mL^{-1}) on ice and transferred to a pre-chilled electroporation cuvette (0.2 cm electrode gap) (Bio-Rad). Electroporation was carried out at 2500 V, 25 µF capacitance, 200 Ω resistance and 5 ms in a Gene Pulser Xcell transformation apparatus (Bio-Rad). Cells were recovered from the cuvette, and immediately resuspended in 1 mL of BHI, followed by incubation at 37 °C for 1 h then plated onto the BHI agar containing appropriate antibiotics. Selection plates were incubated at 37°C for 48 h.

2.11 Bacterial Conjugation

Bacterial conjugation was performed by a modification of the method described by Trieu-Cuot *et al.*, (1991). Briefly, from O/N bacterial cultures, 0.1 vol of donor and recipient strains were each subcultured into 10 mL of fresh BHI and the culture incubated for approximately 5 h or until late logarithmic phase. Cultures were incubated without agitation to avoid damage to the sex pili of the donor strain. The cells were harvested by centrifugation and washed once and gently resuspended in 10 mL of fresh broth. Suspensions of donor (2 mL) and recipient (2 mL) cells were gently mixed. A sample of

this cell suspension (0.2 mL) was spread onto a 5 cm diameter, 0.45 micron nitrocellulose membrane filter (Millipore Corporation), placed onto a agar plate and incubated for 14 - 18 h at 37°C. Bacteria were resuspended in 2 mL of fresh broth. 200μL aliquots of neat and serially diluted culture were plated onto the appropriate selective medium and incubated O/N at 37°C.

2.12 Southern Hybridisation Analysis

2.12.1 Southern Transfer

DNA preparations were separated by electrophoresis on a 1% (w/v) agarose gel in TAE buffer. Unidirectional transfer of DNA from an agarose gel to a positively charged nylon membrane (Roche Diagnostics) was performed by the protocol of Reed, (1990). Following transfer, the DNA was irreversibly bound to the membrane by UV-crosslinking (254 nm, 5 min) and the membrane was washed briefly in distilled H_2O.

2.12.2 DIG-11-UTP Labelling of DNA Probes

Digoxigenin labelled probe DNA was generated by PCR using specific oligonucleotide pairs. The amplified DNA was purified (refer to Section 2.15) and DIG-11-dUTP labelled using a DIG-Nick Translation Mix (Roche Diagnostics) according to the manufacturer's instructions. The standard labelling reaction mixture consisted of 1 μg of template DNA, 16 μL of Milli-Q H_2O and 4 μL of DIG-Nick Translation Mix, with incubation for 90 min at RT. The reaction was terminated by the addition of 1 μL of 0.5 M EDTA (pH 8.0) and heating to 65°C for 10 min. Unincorporated nucleotides were removed by ethanol precipitation (Section 2.8.6).

2.12.3 Hybridisation

Nylon membranes were incubated in prehybridisation solution [6 × SSC, 5 × Denharts solution, 0.5% (w/v) SDS, 250 μg mL^{-1} single stranded herring sperm DNA (Sigma)] for a minimum of 1 h at 42°C in an Extron HI 2001 hybridisation oven (Bartelt Instruments Pty. Ltd. Heidelberg West, Victoria, Australia). The double stranded, labelled probes were diluted in 50 μL of Milli-Q water and boiled for 10 min to denature the DNA. The sample

was immediately placed on ice to prevent the DNA from re-annealing. The denatured probe DNA was then combined with the appropriate volume of fresh prehybridisation buffer, to form hybridisation buffer and incubated with membranes O/N at 42°C. The membranes were washed $2 \times$ for 5 min in $2 \times$ SSC [60 mM NaCl, 60 mM sodium citrate (pH 7.0)] and 0.1% (w/v) SDS at RT, followed by 2×20 min washes with $0.1 \times$ SSC containing 0.1% (w/v) SDS at 42°C.

2.12.4 Detection

Method 1:   After a 1 min wash in Buffer 1 [0.1 M Tris-HCL, 0.15 M NaCl (pH 7.5)], nylon membranes were incubated in Buffer 2 [5% (w/v) skim milk in Buffer 1] for 30 min. The membrane was briefly washed with Buffer 1, followed by a 30 min incubation with anti-DIG-AP Fab fragments (diluted 1:5000 in Buffer 1). Membranes were then washed 2 $\times$ 15 min in Buffer 1, then neutralised by a 2 min wash in Buffer 3 [0.1 M Tris-HCL, 0.1 M NaCl, 50 mM $MgCl_2$ (pH 9.5)]. Labelled hybrid DNA was detected colourimetrically by addition of 0.34 mg/mL NBT and 0.18 mg/mL X-Pho in Buffer 3. The enzyme reaction was terminated by a 5 min wash in Buffer 4 [0.01 M Tris-HCL, 1 mM EDTA (pH 8.0)].

Method 2:   Filter hybridisation and probe hybridisation detection was performed as described in the DIG Application Manual for Filter Hybridisation (Roche Diagnostics). Visualisation of probe-target hybrids was performed using the DIG Luminescent Detection Kit (Roche Diagnostics) as per the manufacturer's instructions and using the supplied reagents and buffers.

2.13 Oligonucleotides

All synthetic oligonucleotide used in this study are described in Table 2.6, Table 2.7, Table 2.8 and Table 2.9. Oligonucleotides were designed using the Primer Designer Version 2 software package (Scientific and Education Software) and were synthesised at 40 nmol concentration, sequencing grade purity from GeneWorks (Adelaide, South Australia). Primers were dissolved in Milli-Q water and stored at -20°C.

2.14 Polymerase Chain Reaction (PCR)

2.14.1 Standard PCR Conditions

All PCR were performed using AmpliTaq Gold® PCR Master Mix (Applied Biosystems) using a standard protocol. Each reaction consisted of 12.5 µL of 2 × AmpliTaq Gold® PCR Master Mix, forward and reverse primers (1 µM of each), template DNA (100 ng) and Milli-Q water to a final volume of 25 µL. Thermal cycling conditions were carried out in a GeneAmp PCR system 2400 (Perkin Elmer, US instrument division, Norwalk, USA). Each cycle consisted of denaturation at 95°C for 30 sec, primer annealing at 50 - 65°C for 30 sec and an extension step at 72°C for 30 sec/500bps for 30 - 50 cycles, dependent upon the required product yield. PCR products were maintained at 4°C or stored at -20°C.

2.14.2 Amplification of Plasmid DNA using TempliPhi^{TM}

Large plasmid DNA was amplified using the TempliPhi^{TM} Large Construct DNA Amplification Kit (Amersham Biosciences, Piscataway, NJ, USA) as per the manufacturer's instructions.

2.15 Purification of PCR Products and other DNA fragments

PCR products and other DNA fragments were purified using the QIAquick® PCR Purification Kit (QIAGEN), as per the manufacturer's instructions. Alternatively PCR products were electrophoresed and purified using methods outlined in Section 2.8.5.

2.16 Dye Terminator Sequence Analysis

2.16.1 PCR Products and Plasmid DNA

All DNA sequencing performed in this study was done so using vector specific custom oligonucleotides, primer walking or PCR product sequencing. BigDye Terminator Ready Reaction Mix Version 3 (Applied Biosystems) was used to sequence DNA. A standard reaction mix, in a total of 14 µL made up with Milli-Q water, consisted of 4 µL of Big-Dye V3 terminator mix (Version 3), 1 µL of forward/reverse primer and 1 µg of template DNA. The thermo-cycling conditions were set at 25 cycles of 95°C for 30 sec, 55°C for 30 sec,

followed by 60°C for 4 min. The DNA was precipitated at RT for 1 - 4 h with 50 µL of 75% propan-2-ol in a 1.5 mL reaction tube. The precipitate was collected by centrifugation (20,000 × g, 20 min, RT) and washed with 250 µL of 75% propan-2-ol. The DNA pellet was dried *in vacuo* (Speed Vac SVC100) and sent to the IMVS sequencing centre for analysis using an Applied Biosystems 377 DNA Analyser.

Optimised sequencing reactions consisted of 1 µL of BigDye Terminator Ready Reaction Mix Version 3 (Applied Biosystems), 3 µL of 5 × Big Dye Buffer (Applied Biosystems), 4 µL of oligonucleotide (final concentration of 3.2pmol/µL), template DNA (200 ng µL^{-1}) and Milli-Q water to a final volume of 20 µL. The thermo-cycling conditions were set at 25 cycles of 95°C for 30 sec, 50°C for 15 sec, followed by 60°C for 4 min. DNA was isopropanol precipitated (see Section 2.8.6) and submitted to the DNA Sequencing Facilities at the Institute for Medical and Veterinary Science (IMVS) (Adelaide, South Australia, Australia) for analysis using an Applied Biosystems 3730 DNA Analyser.

2.16.2 TempliPhi^{TM} Amplified DRDC8 Plasmid DNA

TempliPhi^{TM} amplified DNA standard sequencing reaction mix consisted of the following: 5 µL of product, 4 µL Big-Dye V3 terminator mix (Version 3), 1 µL of forward/reverse primer, 2 µL of 5 × Big Dye Buffer (Applied Biosystems), and 8 µL of Milli-Q water. This reaction mix was subjected to the following thermo-cycling conditions, 60 cycles of 95°C for 20 sec, 50°C for 30 sec, followed by 60°C for 4 min. The DNA was precipitated by the addition of 1 µL of 0.5 M EDTA and 80 µL of precipitation mix [0.7 M Sodium Acetate (v/v) and 70% EtOH (v/v)] and incubated at RT for 15 min. The DNA was then collected by centrifugation, washed with 250 µL of 70% EtOH, dried *in vacuo* (Speed Vac SVC100) and submitted to the DNA Sequencing Facilities at the Institute for Medical and Veterinary Science (IMVS) (Adelaide, South Australia, Australia) for analysis using an Applied Biosystems 3730 DNA Analyser.

2.16.3 Computer Analysis of Sequence Data

Non-contiguous sequence was aligned and joined manually using the ClustalX 1.81 (Thompson *et al.*, 1994) multiple alignment program and BLASTN, BLASTX and

BLASTP alignment algorithms (Altschul *et al.*, 1990). ORF prediction and gene family identification were completed using methodology described previously (Nelson *et al.*, 1999; Nelson *et al.*, 2004). ORFs were identified using the NCBI ORF Finder analysis tool (National Institutes of Health (NIH); U.S.A.). The BLASTX alignment algorithm was used to identify sequence similarity to data in the GenBank database (Altschul *et al.*, 1990).

2.17 Listeria Antibodies

2.17.1 Immuno-fluorescence Microscopy of *L. monocytogenes* Cells

Overnight cultures of *L. monocytogenes* strains were harvested by centrifugation and the cells resuspended in cold PBS to achieve a concentration of *ca.* 10^7 cells mL^{-1}. The bacterial cells were smeared on a microscope slide and fixed in 10% (v/v) formalin for 30 min. The cells were washed with Milli-Q water and allowed to air dry. Anti *Listeria* antibody (Difco^{TM} *Listeria* O Antisera Type 1 and *Listeria* O Antisera Type Poly 1,4) was added to the fixed cells (final dilution of 1 in 200) and incubated at RT for 30 min followed by washing 6 × with PBS. FITC-labelled antibody (Anti-Rabbit Immunoglobulin F(ab)$_2$ Fraction Affinity Isolated Fluorescein Conjugated, Silenus, USA) to a final dilution of 1 in 100 in 10% (v/v) foetal bovine serum in 1 × PBS, was placed on the cell smear and incubated for 30 min at RT in the dark. The cell smear were washed a further 6 × with 1 mL PBS, drained on tissue paper and allowed to air dry. A 3 μL drop of Mowiol 4-88 containing an anti-bleaching agent (*p*-phenylene diamine) was placed onto the microscope slide, covered with a coverslip, allowed to stand for 10 - 15 min. Cells were analysed using an Olympus BH-S microscope with a BH2-RFC UV fluorescent attachment (Olympus Optical Co. Ltd.).

2.17.2 *L. monocytogenes* Serotype Analysis

A loopful of bacterial growth was removed from an agar plate and resuspended into 1 mL of 1 × PBS in a reaction tube. Following heating at 80°C for 1 h, the cells were collected by centrifugation and the majority of supernatant removed. A drop of the resuspended heated cells was mixed with 20 μL of diluted antibody for 1 - 2 min on a glass slide and visualised for agglutination reaction.

2.18 Curing *L. monocytogenes* Plasmid DNA

L. monocytogenes strains were cured of plasmid DNA by a modification of the method previously described by Miller, (1972). Briefly, approximately 10^4 cells were added to 10 mL of BHI broth (pH 7.6) containing a sub-inhibitory concentration of acridine orange (10 μg mL^{-1}) and serially cultured O/N at 37°C in the dark with shaking. Cells were plated onto BHI agar supplemented with acridine orange (10 μg mL^{-1}) and incubated O/N at 37°C. Plasmid cured variants were isolated on BHI supplemented with acridine orange (10 μg mL^{-1}) and an appropriate selective marker.

2.19 MIC of Heavy Metal Cations

The minimal inhibitory concentrations (MIC) of $CuSO_4$, $CdSO_4$ and 8-hydroxyquinoline for *L. monocytogenes* strains were determined and calculated as previously described (Sahm & Washington II, 1991). Briefly, a 5 μL sample of an O/N culture washed in saline was spotted onto BHI agar plates containing copper sulphate (0 - 25 mM), cadmium sulphate (0 - 1 mM) ions or 8-hydroxyquinoline (0 - 80 μM). Plates were then incubated at 37°C for 48 h. The MIC values were defined as the lowest concentration that inhibited visible growth of *L. monocytogenes*.

2.19.1 Statistical Analysis

MIC values of $CuSO_4$, $CdSO_4$ and 8-hydroxyquinoline for *L. monocytogenes* strains were compared using unpaired Student's *t*-tests in Prism (GraphPad Software Inc.) software, with $P < 0.05$ considered significant.

2.20 Bacterial Growth Experiments

O/N cultures of *L. monocytogenes* strains were subcultured (0.01 vol) in 100 mL of BHI broth. Cultures were incubated at 37 °C with agitation. The OD_{600nm} was measured in a Spectronic 20D$^+$ Spectrophotometer (Milton Roy, Ivyland, PA, USA) for a period of 24 h.

2.21 Cell Culture

2.21.1 Cell Lines

The J774 murine macrophage-like cell line (Rathjen & Geczy, 1986), was used for tissue culture studies

2.21.2 Growth Conditions and Culture Media

J774 cells were routinely cultured in 75 cm^3 plastic tissue culture flasks (Corning, New York, USA) in GIBCO DMEM supplemented with 5% (v/v) foetal calf serum, Penicillin G (100 μg mL^{-1}) and Streptomycin (10 μg mL^{-1}). Cells were maintained in a 37°C humidified incubator containing 5% CO_2. J774 cell cultures were split as necessary.

2.21.3 Preparation of Glass Coverslips

Coverslips (12 mm diameter) were washed O/N in a solution containing 0.34 M potassium dichromate and 1 M sulphuric acid. The acid solution was removed by repeated washing in cold Milli-Q water. Coverslips were air dried, baked for 2 h at 80°C and stored in methanol.

2.21.4 Tissue Culture Monolayer Invasion Assay

Semi-confluent monolayers were prepared by subculture of cell monolayers into 24-well tissue culture plate (Falcon, Becton Dickinson Labware), containing prepared coverslips. Approximately 5×10^5 J774 cells were added to each well. The trays were then incubated at 37°C in 5% CO_2 O/N. Prior to infection with *L. monocytogenes* DRDC8Sm* (refer to Table 2.1) cells, J774 cell monolayers on coverslips were washed 2× with 1 mL of pre-warmed (37°C) antibiotic free DMEM culture medium.

Cultures of logarithmic phase *L. monocytogenes* (10 mL) were prepared in BHI supplemented with 0.2% charcoal (see Section 2.3). The cells were harvested by centrifugation and were washed $2 \times$ in 2 mL of D-PBS [$1 \times$ PBS, 0.1% $CaCl_2$ (v/v), 0.1% $MgCl_2$ (v/v)] and resuspended in 1 mL of D-PBS. The resultant cell suspension was diluted in fresh DMEM culture media and added to the J774 monolayers to achieve a MOI of 1 bacterial cell to 300 J774 cells (1:300). The mixtures were incubated for 30 min at

37°C in 5% CO_2 to allow for the bacteria to invade J774 cells. Infected J774 cell monolayers were then washed 3 × with antibiotic free DMEM medium followed by addition of fresh DMEM supplemented with 40 μg mL^{-1} gentamycin and incubated for a 60 min to kill extracellular bacteria. Gentamycin treated monolayers were then washed with 2 × 1 mL fresh antibiotic free DMEM medium, covered with 1 mL of fresh DMEM medium and incubated for a further 4 h at 37°C in 5% CO_2. Intracellular bacteria were recovered by lysis of the monolayers with 500 μL ice-cold 0.1% (v/v) Triton X-100 in 1 × D-PBS. Numbers of viable bacterial cells was determined by plating appropriate serial dilutions of the lysates on BHI agar supplemented with the appropriate antibiotics.

2.22 Isolation of Bacteriophage from *L. monocytogenes* DRDC8

2.22.1 Induction of Bacteriophage

Mitomycin C induction of bacteriophage from cultures of *L. monocytogenes* was performed by a modification of methods as described by Maniatis *et al.*, (1982). An O/N culture of *L. monocytogenes* was used to inoculate 200 mL of fresh BHI broth. The culture was incubated at 37°C with agitation for 3 h. Mitomycin C was added to a final concentration of 1 μg mL^{-1} and the culture incubated for an additional 20 h. Chloroform was added to the culture (1% (v/v)) and incubated for a further 15 min. The aqueous, supernatant phase was collected following centrifugation (JA-14 rotor, 15,000 × g, 20 min).

2.22.2 Purification of Bacteriophage

The supernatant from induced cultures was treated with DNase (10 μg mL^{-1}) and RNase (5 μg mL^{-1}) at RT for 30 min. Solid NaCl was then added to a final concentration of 1 M and dissolved by stirring and the mixture incubated on ice for 1 h. Following centrifugation (JA-14 rotor, 15,000 × g, 10 min), the supernatant was collected. Solid PEG_{600} (10% (w/v)) was added and dissolved by slow stirring at RT and left to stand on ice for 1 h. The supernatant was discarded following centrifugation and the remaining pellet resuspended in 3.2 mL of SM buffer [0.58% (w/v) NaCL, 0.2% (w/v) $MgSO_4.7H_2O$, 50mM Tris-HCL (pH 7.5), 0.01% (w/v) gelatin]. An equal volume of chloroform was

added, the mixture vortexed for 30 sec and the aqueous phase containing bacteriophage recovered by centrifugation. The bacteriophage suspensions were then stored at 4°C or used immediately.

Bacteriophage particles were recovered by CsCl gradient [35% (w/w), 41% (w/w), 47% (w/w) CsCl in SM buffer] centrifugation in 5.1 mL Quick-Seal® ultracentrifuge tubes. A sample of prepared bacteriophage suspension (1 mL) (see Section 2.22.1) was added to the gradient and centrifuged in a Beckman Optima^{TM} TLX Ultracentrifuge (GMI) (SW 40 Ti rotor, 155,000 × g, 3 h, 20°C). Visible bands containing bacteriophage particles were extracted with a 21 gauge needle and dialyzed in SM buffer.

The presence of bacteriophage particles was confirmed by electron microscopy following staining. Staining was performed by placing a formvar coated copper electron microscope grid onto a 20 μL drop of bacteriophage preparation, and excess liquid was removed with filter paper. The grid was placed in 4% uranyl acetate and excess stain removed with filter paper, and the grid examined using a Philips CM 200 Transmission Electron Microscope.

2.22.3 Transfection of *L. monocytogenes* with Bacteriophage Preparations

O/N cultures of *L. monocytogenes* strains were subcultured (0.01 vol) in 10 mL into fresh media and grown for approximately 4.5 h. 1 mL aliquots of bacterial culture was transferred to reaction tubes and MgSO$_4$ added to a final concentration of 0.01 M. Bacteriophage particles (Section 2.22.2) were added to the tubes as 10 μL volumes of either neat or serially diluted preparations. The tubes were shaken at 37°C for 2 h, then plated out onto the appropriate selective BHI agar and incubated O/N.

2.23 Protein Overexpression

Vector construction and protein overexpression was performed using the pET-21c(+) plasmid expression system (Novagen) according to the Novagen pET system manual (11th edition).

Strains for overexpression were grown in 10 mL of LB broth supplemented with Amp and incubated O/N at 37°C. This culture (10 mL) was used to inoculate 500 mL of fresh

LB broth, which was incubated with shaking at 37°C until a culture OD_{600} 0.6 - 0.9 was reached. Induction of protein expression was achieved by the addition of IPTG to a final concentration of 1 mM and further incubation with shaking for a minimum of 6 h at 30°C. Cells were pelleted by centrifugation and stored at -20°C O/N.

2.24 Preparation of whole Cell Lysates

Cell pellets prepared in Section 2.23 were resuspended in 12.5 mL of lysis buffer [50 mM potassium phosphate pH 7.8, 400 mM NaCl, 100 mM KCl, 10% (v/v) glycerol, 0.5% (v/v) Triton X-100, 10 mM imidazole] and incubated on ice for 10 min. (0.01 vol) of PMSF (100 mM) or Protease Inhibitor Cocktail (Sigma). Cells were then sonicated using a Branson Sonifier Cell Disrupter Model B15. Sonication of cells was performed on ice with a microtip on a pulsed cycle (50%) for 30 sec followed by 30 sec to cool on ice for a minimum of 10 min. Alternatively, cell lysates were prepared by several passages through a pre-cooled French pressure cell (Aminco-SLM, Urbana, IL). Cell debris and any whole cells were pelleted by centrifugation (20,000 × g, 1 h, 4°C). The supernatant was used immediately, kept on ice for short-term storage or long term stored at -20°C in 50% glycerol.

2.25 SDS-PAGE

SDS-PAGE was performed using a modification of the method of Lugtenberg *et al.*, (1975) as described by Achtman *et al.*, (1978). Gels were prepared, electrophoresed and transferred using a Mini-protean® II Dual Slab Cell Apparatus (Bio-Rad). Stock solutions of acrylamide (40%) and bis-acrylamide (2%) (BioRad) were used to prepare acrylamide gels, as per the manufacturer's instructions. Gels were constructed as a separating gel (15% (v/v) poly acrylamide 37.5:1 acrylamide:bis) that was overlayed with a stacking gel (4% (v/v) polyacrylamide 37.5:1 acrylamide:bis). Electrophoresis was performed in GTS [SDS-PAGE Electrophoresis Buffer (25 mM Tris-HCL, 192 mM glycine) (BDH AnalaR®, Merck, Australia), 0.1% SDS (Sigma)]. All protein samples were denatured at 95°C for 5 min in 5 × LaemmLi loading reducing sample buffer [0.25 mM Tris-HCL, pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 15% (w/v) bromophenol blue] immediately prior to loading onto the gel. Broad range,

prestained protein marker, (New England Biolabs) (Table 2.5), was also heat denatured and electrophoresed in conjunction with protein samples for size and molecular weight comparison. Electrophoresis was carried out at 100 V through the stacking gel, then run at 200 V for approximately 45 min through the running gel until adequate dye front migration was achieved. The gel was then stained by gentle agitation in Coomassie stain [0.09% (w/v) Comassie Brilliant Blue G250, 40% (v/v) methanol, 10% (v/v) acetic acid] for 2 - 24 h, followed by destaining in 5% (v/v) acetic acid, 10% (v/v) methanol.

2.26 Purification of His-tagged Protein

Recombinant His-tagged protein was purified from whole cell lysate using by gravity flow column chromatography using HisLink^{TM} Protein Purification Resin (Promega), according to the manufacturer's instructions. A HisLink^{TM} resin column was prepared by loading 1 mL of resin slurry into a Disposable 5 mL Polypropylene Column (Pierce). Prior to loading onto the equilibrated HisLink^{TM} resin column, freshly prepared cell lysate (refer to Section 2.24) was treated with DNase I (10 μg mL^{-1}) (Roche Diagnostics) and RNase A (5 μg mL^{-1}) (Sigma) at RT for 15 min. This was followed by centrifugation at 10,000 rpm for 30 min at 4°C and further clearing by filtration through a 0.2 μm filter. All fractions eluted from the column were analysed by SDS-PAGE. Fractions containing purified protein were subject to dialysis at 4°C with storage buffer (20 mM Tris-HCL pH 7.5, 50% (v/v) glycerol, 1 mM DTT, 0.01% (v/v) NP40, 300 mM NaCl) and stored at -20°C in 50% (v/v) final concentration of glycerol.

2.27 Protein Mass Spectrometry Sequencing

Purified protein was sequenced by the Adelaide Proteomics Centre by Mass Spectrometry analysis using the following methods.

The protein band of interest was excised from an SDS-PAGE gel and chopped into 1 mm^3 pieces. The gel pieces were destained, reduced and alkylated with iodoacetamide, then digested with 100 ng of trypsin per sample. The resulting tryptic peptides were extracted from the gel pieces with 50% acetonitrile containing 1% formic acid (FA) followed by 100% acetonitrile. The volumes of the final samples were reduced from

approximately 120 mL to ~1 mL by vacuum centrifugation. The peptides were then diluted to ~5 mL FA30 (7 parts 0.1% FA to 3 parts acetonitrile).

Each sample was applied to a 600 mm AnchorChip (Bruker Daltonik GmbH, Bremen, Germany) in a matrix of a-cyano-4-hydroxycinnamic acid (HCCA) according to the manufacturer's thin layer method.

MALDI TOF mass spectra were acquired using a Bruker ultraflex II MALDI TOF/TOF mass spectrometer (Bruker Daltonik GmbH) operating in reflectron mode under the control of the flexControl software (Version 3.0, Bruker Daltonik GmbH). External calibration was performed using peptide standards (Bruker Daltonik GmbH) that were analysed under the same conditions. Spectra were obtained randomly over the surface of the matrix spot at a laser intensity determined by the operator. MALDI-TOF/TOF (MS/MS analysis) was performed in the LIFT mode using the same spot on the target.

The MS and MS/MS spectra were analysed in flexAnalysis (Version 3.0, Bruker Daltonik GmbH). Peak detection was performed using the SNAP algorithm with a signal to noise threshold of 3 and a quality factor threshold of 50. The processed spectra and mass lists were exported to BioTools (Version 3.0, Bruker Daltonik GmbH). Here, the MS and corresponding MS/MS spectra were combined and matched to the provided protein sequence using Sequence Editor. The following parameters were employed when exporting the theoretical peptide masses from Sequence Editor to BioTools: Global modifications: Carbamidomethyl (C), Variable modifications: Oxidation (M), Mass tol MS: 100 ppm, MS/MS tol: 0.5 Da, Missed cleavages: 1.

2.28 Determination of Protein Concentration

The concentration of purified protein preparations was estimated using the Coomassie$^{®}$ Plus Protein Assay Regent Kit (Pierce Biotechnology, Rockford, IL), as per the manufacturer's instructions. The absorbance of protein samples were measured spectrophotometrically at 562 nm in a BioRad Smart Spec^{TM} 3000 (BioRad Laboratories Pty Ltd., NSW, Australia).

2.29 Gel Shift Analysis

2.29.1 Labelling of DNA

Gel shift analysis was performed with Digoxigenin labelled DNA. PCR amplified DNA or double stranded oligonucleotide DNA (10 pmol) was labelled using digoxigenin (Dig)-11-dideoxy-UTP (Roche Diagnostics) and terminal transferase (Roche Diagnostics) according to the manufacturer's instructions using the supplied reagents. The labelling reaction consisted of $5 \times$ TdT reaction buffer, 5 mM $CoCl_2$, 10 pmol PCR product, 1 nmol DIG-11-ddUTP, 400 U Terminal Transferase in a total volume of 20 µL. The reaction mixture was incubated at 37°C for 15 min. The reaction was stopped by addition of 2 µL of 0.2 M EDTA (pH 8.0). Digoxigenin labelled double stranded oligonucleotides were prepared by mixing the two complementary sequences, heating to 80°C and then cooling to 30°C over a period of 1 h.

2.29.2 Binding Reaction

In a total volume of 20 µL, purified protein (see Section 2.26) was pre-incubated at RT with 100 µg of pre-sonicated salmon sperm DNA in $1 \times$ binding buffer [20 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 10% (v/v) glycerol, 1 mM DTT, 0.01% (v/v) NP-40, 50 mM NaCl] for 20 min. For assays analysing the effects of binding in the presence of copper or cadmium ions, $CuSO_4$ or $CdSO_4$ was added to the desired final concentration. Approximately 0.5 pmol of labelled target DNA (see Section 2.29.1) was then added and the mixture incubated at RT for an additional 20 min. The entire reaction mixture was mixed with $5 \times$ Loading buffer [250 mM Tris-HCL pH 7.5, 40% (v/v) glycerol, 0.2% (w/v) bromophenol blue] and electrophoresed on a 6% non-denaturing PAGE gel (Section 2.29.3).

2.29.3 Electrophoresis and Transfer

Non-denaturing PAGE gels (6% acrylamide) were assembled and electrophoresed in a Mini-protean II apparatus (BioRad). Stock solutions of acrylamide (40% w/v) and bis-acrylamide (2% w/v) (BioRad) were used to prepared a 6% polyacrylamide gel mix containing 29:1 acrylamide:bis, 10% glycerol (v/v), $1 \times$ TBE [90 mM Tris-borate, 2 mM

EDTA (pH 8.0)]. Following degassing, 100 µL ammonium persulphate (10% w/v) and 15 µL TEMED (Sigma) was added to the gel mix. Following pouring, gels were allowed to polymerise O/N at 4°C. Gels were then pre-electrophoresed in chilled $1 \times$ TBE at 100 V for 1 h at 4°C.

Reaction mixtures containing protein-DNA complexes (see Section 2.29.2) were mixed with $5 \times$ Loading buffer [250 mM Tris-HCL (pH 7.5), 40% (v/v) glycerol, 0.2% (w/v) bromophenol blue], loaded onto the gel and electrophoresed at 80 V for approximately 2 - 3 h at 4°C. The DNA-protein complexes were transferred onto positively charged nylon membranes (Roche Diagnostics) by electroblotting (400 mA for 15 min) in a Semi-Phor^{TM} semi dry transfer unit (Hoefer Scientific Instruments, San Francisco, USA). The membranes were placed in-between several sheets of Whatman paper pre-soaked in chilled $1 \times$ TBE during the transfer process. The membranes were then fixed in $10 \times$ SSC (5 min), followed by exposure to UV light (254 nm, 3 min).

2.29.4 Detection

DNA-protein complexes were detected according to Method 2, described in Section 2.12.4.

Table 2.1: *L. monocytogenes* **strains used in this study**

Bacterial strain	*ctpA* (+/-)[5]	Origin	Source
Dairy22	+	dairy	DTS[#]
Dairy25	+	dairy	DTS[#]
Dairy26	+	dairy	DTS[#]
DC1059	+	dairy	NSW DCL[#]
DC1060	+	dairy	NSW DCL[#]
DC1141	+	dairy	NSW DCL[#]
DRDC8	+	milk	NSW DCL[#]
DRDC8Sm*	+	spontaneous Sm[R] variant of DRDC8	This study
DSE201	-	DRDC8 containing *erm* insertion in *ctpA*, Sm[R], Em[R]	Francis and Thomas, (1997b)
DSE201PL	-	plasmid-cured variant of DSE201	This study
DSE951	+	DRDC8 containing *erm* insertion in ORF pCT0018, Em[R]	This study
DSE952	+	DRDC8 containing *erm* insertion in ORF pCT0019, Em[R]	This study
DSE955	+	DRDC8 containing *erm* insertion in ORF *cutR*, Em[R]	This study
DSE955PL	-	plasmid-cured variant of DSE955	This study
DTS47081	+	environmental	DTS[#]
EGD Kaufmann	-	laboratory isolate, serotype 1/2a	H. Hof[#]
EGD-Rif	-	spontaneous Rif[R] variant of EGD Kaufmann	This study
EGD-Sp	-	spontaneous Sp[R] variant of EGD Kaufmann	This study
ING2	+	poultry	IC[#]
KE391	-	clinical human	KEH[#]
KE795	-	clinical human	KEH[#]
KE1003	-	clinical human	KEH[#]

Table continued next page…..

---

[5] Wild type strains originally characterised by Bell, (2002).

Table 2.1 continued

NR28/1	+	poultry	APL[#]
pKS950 cointegrate	+	DRDC8 carrying the integrated plasmid pKS950	This study
pKS951 cointegrate	+	DRDC8 carrying the integrated plasmid pKS951	This study
pKS952 cointegrate	+	DRDC8 carrying the integrated plasmid pKS952	This study
pKS955 cointegrate	+	DRDC8 carrying the integrated plasmid pKS955	This study
TSANR31/5	+	poultry	APL[#]
0121	N/A	food	SARDI[#]
0124	N/A	food	SARDI[#]
0146	N/A	food	SARDI[#]
0177	N/A	food	SARDI[#]
0179	N/A	food	SARDI[#]
24/3	N/A	food	SARDI[#]

[#] DTS, Dairy Technical services; IC, Inghams Chickens; NSWDCL, New South Wales Dairy Corporation Laboratories; IMVS, Institute of Medical and Veterinary Science; KEH, King Edward Hospital; APL, Australian Poultry Ltd; H. Hof, University of Mannheim, Germany; SARDI, South Australian Research and Development Institute.

Table 2.2: *E. coli* strains used in this study

Bacterial strains	Genotype/Phenotype	Source/Reference[a]
DH5α	ϕ80dlacZΔM15, recA1, endA1, gyrA96, thi-1, hsdR17(r_K^-, m_K^+), supE44, relA1, deoR, Δ(lacZYA-argF)U169	BRL
XL10-Gold$^{\circledR}$	Tet^r Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lacHte [F' proAB lacI^{q}ZΔM15 Tn10 (Tet^r) Amy Cam^r]	Stratagene
BL21(DE3)	F$^-$ ompT gal [dcm] [lon] hsdS$_B$ (r_B^-, m_B^-; an *E. coli* B strain) with DE3, a λ prophage carrying the T7 RNA polymerase gene	Grodberg and Dunn, (1988)
MF185	DH5α harbouring plasmid pCT200	Francis, (1996)
MF18596HP	DH5α harbouring plasmid pCT217	Francis, (1996)

Table 2.3: Plasmids used in this study

Vector/Clone	Description/Marker	Source/Reference
pGI21	IS*10*. Em^R, Amp^R.	Mahillon and Kleckner, (1992)
pBR322	*rop*, *ori*. Tc^R, Amp^R.	GenBank Accession J01749
pGEM®-T Easy	Direct T-A cloning of PCR products, lacZ blue/white colour selection. Amp^R.	Promega
pUC19	*lacZα*, *ori*. Amp^R.	GenBank Accession L09137
pET-21a-d(+)	Expression vectors providing IPTG inducible T7 promoter and N-terminal 6×His-tag. Encodes LacI repressor. Amp^R.	Novagen
pKSV7	*lacZ*, *cat*, *ori*, Ts rep. Cm^R, Amp^R.	Smith and Youngman, (1992)
pFB186	pUC19 carrying a 5.73 kbp *Hind*III fragment of plasmid pCT200.	This study (see Section 3.3.5.1)
pFB190	pUC19 carrying a 1.61 kbp *Hind*III fragment of plasmid pCT200.	This study (see Section 3.3.5.1)
pCT100	*L. monocytogenes* strain DRDC8 large plasmid.	This study (see Section 3.3.1.1)
pCT200	pBR322 carrying a *ca*. 28.5 kbp *Bam*HI fragment of *L. monocytogenes* strain DRDC8 plasmid pCT100.	Francis, (1996)
pCT212	pGEM-7Zf(+) carrying a 5.3 kbp *Hind*III fragment of *L. monocytogenes* strain DRDC8 plasmid pCT100.	Francis, (1996)
pCT214	pGEM-7Zf(+) carrying a 3 kbp *Sac*I fragment of *L. monocytogenes* strain DRDC8 plasmid pCT100.	Francis, (1996)
pCT217	pGEM-7Zf(+) carrying a 2.3 kbp *Hpa*I fragment of plasmid pCT200.	Francis, (1996)
pCT750	pGEM®-T Easy carrying a 830 bp PCR product of ORF pCT0017 with an introduced *Bam*HI site.	This study (see Section 4.3.1)
pCT751	pGEM®-T Easy carrying a 1003 bp PCR product of ORF pCT0018 with an introduced *Bam*HI site.	This study (see Section 4.3.1)

Table continued next page …..

Table 2.3 continued

pCT752	pGEM®-T Easy carrying a 1003 bp PCR product of ORF pCT0019 with an introduced *Bam*HI site.	This study (see Section 4.3.1)
pCT755	pGEM®-T Easy carrying a 1250 bp PCR product of ORF pCT0018 with an introduced *Bam*HI site.	This study (see Section 4.3.1)
pCT800	pUC19 carrying a 1136 bp *Bam*HI *erm* cassette from pG121.	This study (see Section 4.3.1)
pCT850	pCT750 carrying a 1136 bp *Bam*HI *erm* fragment from pCT800.	This study (see Section 4.3.1)
pCT851	pCT751 carrying a 1136 bp *Bam*HI *erm* fragment from pCT800.	This study (see Section 4.3.1)
pCT852	pCT752 carrying a 1136 bp *Bam*HI *erm* fragment from pCT800.	This study (see Section 4.3.1)
pCT855	pCT755 carrying 1136 bp *Bam*HI *erm* fragment from pCT800.	This study (see Section 4.3.1)
pKS950	pKSV7 carrying a 1966 bp pCT0017::*erm Eco*RI fragment from pCT850.	This study (see Section 4.3.2.1)
pKS951	pKSV7 carrying a 2139 bp pCT0018::*erm Eco*RI fragment from pCT850.	This study (see Section 4.3.2.1)
pKS952	pKSV7 carrying a 2139 bp pCT0019::*erm Eco*RI fragment from pCT850.	This study (see Section 4.3.2.1)
pKS955	pKSV7 carrying a 2386 bp *cutR*::*erm Eco*RI fragment from pCT850.	This study (see Section 4.3.2.1)
pCTCF	pGEM®-T Easy carrying a 516 bp PCR product of ORF pCT0017 with an introduced *Nde*I (5' end) and *Xho*I site (3' end).	This study (see Section 5.3.1.1)
pETCF	491 bp *Nde*I/*Xho*I fragment from pCTCF cloned into pET21-c(+) with ORF pCT0017 in frame with the N-terminal 6×His-tag.	This study (see Section 5.3.1.1)

Table 2.4: DNA markers used in this study

Band	EcoRI/SPP1	2-Log DNA ladder	λ DNA-Mono Cut Mix
		Fragment Sizes (kbp)	
1	8.55	10	48.502
2	7.35	8	38.416
3	6.11	6	33.498
4	4.84	5	29.946
5	3.59	4	24.508
6	2.81	3	23.994
7	1.95	2	17.053
8	1.86	1.5	15.004
9	1.51	1.2	10.086
10	1.39	1	1.503
11	1.16	0.9	
12	0.98	0.8	
13	0.72	0.7	
14	0.48	0.6	
15	0.36	0.5	
16		0.4	
17		0.3	
18		0.2	
19		0.1	

Table 2.5: Protein markers used in this study

	Size (Daltons)
Band	**Prestained, broad range***
1	175,000 (MBP-β-galactosidase)
2	83,000 (MBP-paramyosin)
3	62,000 (glutamic dehydrogenase)
4	47,500 (aldolase)
5	32,500 (triosephosphate isomerase)
6	25,000 (β-lactoglobulin A)
7	16,500 (lysozyme)
8	6,500 (aprotinin)

* apparent molecular weight given for prestained markers. For source of protein see New England Biolabs; MBP = maltose binding protein

Table 2.6: Oligonucleotides used for sequencing plasmid pCT100

Oligonucleotide	Oligonucleotide sequence 5'→3'	Tm°C	Source
FB002	GAG GTT ACT CTC CTA GTA AA	48	This study
FB004	GAT ACA TCT AAG TCA GGC GA	50	This study
FB011	GTA GTG AGA CAC TTA TAG GT	48	This study
FB111	CTT GGA ACT AGC GAA GAA TA	48	This study
FB186F	GTT CGA TGT TCA GTG TTA GA	48	This study
FB1862F	GGC TAC CAT GAT GAT ATC CT	50	This study
FB1864F	GAT GTG GAT CAC ACG AAT AC	50	This study
FB190R	CAG ATG GTG TGA AGT TGT TG	50	This study
FB195F	GGT AAT ACG GTT ATC CAC AG	50	This study
FB211	CTG GAT AGT TCC GAT TAT CA	48	This study
FBTK1	TTA AGT ATG GCT GAA GTA CC	48	This study
p1037	TGA TCC GAT GAC ACC TAT GA	62	Francis, (1996)
p2037	CAA TAG TTG TGC CTA ATC CA	54	This study
p3037	TTA CTG GTT CTC AGG ATT GC	50	This study
pDR10R	CAG GTA TCG AAT TCA GAG TT	48	This study
pDR111F	CTC TGC TAA TAA TAC GAT CC	48	This study
pDR11R	CTC ACA ATA AGC TTA CTC CT	48	This study
pDR12	ACT CTG AAT TCG ATA CCT GT	48	This study
pDR121F	TAG GGA TAT AGG GAT ATA GG	48	This study
pDR13	AGA AGC GCT GCT GAC AGA AG	48	This study
pDR14	CAA GTT CAT TAG GAA GGA GG	50	This study
pDR16R	CCT CCT TCC TAA TGA ACT TG	50	This study
pDR18	CGA CAG TTT AAT GCT AAC AG	48	This study
pDR1F	GAT ATC ACC GTG TTG ACA TT	48	This study
pDR2F	GCT CCT ATC TAG TCC ATC AA	50	This study
pDR3	TAT AGC ATC CAT GGA TTG AG	48	This study
pDR4	GAA CAC CAT GGA TAT GGA AC	50	This study
pDR5	CCA CCC AGT TTT CTA ATA AG	48	This study
pDR6	GTC CTG CTT CTA TAA TCT GA	48	This study

Table continued next page….

Table 2.6 continued

pDR7	CTC ATA TCG TAA CTA CCA GG	50	This study
pDR8	CTG AGC AAT TCA TTC CAC TT	48	This study
pDR9	CAG CAA TAC TTT ACA GGA GA	48	This study
pUC/M13F	GTT TTC CCA GTC ACG AC	60	Promega
pUC/M13R	CAG GAA ACA GCT ATG AC	55	Promega
S12211	TTA TCG CAT GGA TGC TTA TC	48	This study
S12213	CTT GCA TTC CTA TAA AGA CC	48	This study
S12215	GCT GAA ATG GTC TTC CCG AC	62	This study
S1224	GCC TGA CTT AGA TGT ATC AA	48	This study
S1225	AAT TCG ATC ATC TTC CAC AC	48	This study
S1225	AAT TCG ATC ATC TTC CAC AC	48	This study
S1226	CAA TCA CAA GTT ACT TCG TC	48	This study
S1227	CTT CAC TGT TAT CCA CGC TA	50	This study
S1227	CTT CAC TGT TAT CCA CGC TA	50	This study
S1228	TTC ATC TGC TGT AGT ACT CC	50	This study
S1229	GCC TTA CAT AAC TGC TAT CA	48	This study

Table 2.7: Oligonucleotides used for *L. monocytogenes* mutant construction

*Bam*HI restriction sites are underlined.

Oligonucleotide	Oligonucleotide sequence (5'→ 3')	Tm°C	Source
catPF	AGA ACT GGT TAC AAT AGC GA	48	This study
catPR	CTG CAT GAT AAC CAT CAC AA	48	This study
CuBamF	GTG AAA AAG TTC CGA T<u>GG ATC CCG</u> AAA TTA TTT CTG G	63	This study
CuBamR	CCA GAA ATA ATT TCG <u>GGA TCC</u> ATC GGA ACT TTT TCA C	63	This study
CuMutR	GAT GCT CTG ATT GTT GTT CC	50	This study
CuTrM2F	GCG AAT GTC AAC TTA GTA AC	48	This study
CuTrM2R	CTT ACA AGT GTT ACA TCT CC	48	This study
CuTrF	GGA AGT TAC ATC GAC AGA AG	50	This study
CuTrR	ATG AAT CTA CTG GCA CTA ACC	50	This study
ermF	GTG TTG ATA GTG CAG TAT CT	48	This study
ermR	ATC TGG AAC ATC TGT GGT AT	48	This study
ermF2	TCG AGA ATA TCG TCA ACT GT	48	This study
ermR2	AGT GAC ACG AGT ATT CAG TT	48	This study
FB003	ACG AGT AAA AGA GAC TTC AG	48	This study
FB1863	TAC TAT TAG CAA TCG TCT CG	48	This study
FB1865	AAC TGA ACA CGT TGT TAC TG	48	This study
FBDBF	TAT AAA ATG TTC <u>GGA TCC</u> AAA ATT AAT AAT	51	This study
FBDBR	ATT ATT AAT TTT <u>GGA TCC</u> GAA CAT TTT ATA	51	This study
FBEBF	AGA TTT TGG AGT <u>GGA TCC</u> GGA TTT ACC ACT	60	This study
FBEBR	AGT GGT AAA TCC <u>GGA TCC</u> ACT CCA AAA TCT	60	This study
FBdef	TGC AAC TTG TTC TGA AGA TG	48	This study
MokFF	TCT ACA AGA GTA GAC TTT GG	50	Mok, (2003)
MokFR	AGT GGT AAA TCC GCA TCC AC	48	Mok, (2003)

Table continued next page…..

Table 2.7 continued

MokBF	AAG <u>GGA TCC</u> GCC AAA TGA CAC TAT C	48	Mok, (2003)
MokBR	GGC <u>GGA TCC</u> CTT CAT CTT CAG AAC A	48	Mok, (2003)
pUC/M13F	GTT TTC CCA GTC ACG AC	60	Promega
pUC/M13R	CAG GAA ACA GCT ATG AC	55	Promega
S12210	CCA AGC TGA CTA AGG TAT GT	50	This study

Table 2.8: Oligonucleotides used for protein overexpression and Gel shift analysis

*Nde*I restriction site is underlined and *Xho*I site shown in bold.

Oligonucleotide	Oligonucleotide sequence 5'→3'	Tm°C	Source
CBox1	ATT TAT CAT CTA CAA ATG TAA ATT TTT AGT TG	51	This study
CBox1C	CAA CTA AAA ATT TAC ATT TGT AGA TGA TAA AT	51	This study
CBox2	ACA AAT TAA ATC TAC AAG AGT AGA CTT TGG TCA T	57	This study
CBox2C	ATG ACC AAA GTC TAC TCT TGT AGA TTT AAT TTG T	57	This study
GS001F	GAA GTA GTT GTT GAC GGA GG	52	This study
GS002R	ACA TCC TGA TTG AGT GGT AA	48	This study
GS003F	AGA TGC CTC ATC TTG TCT TG	50	This study
GS004F	GAC GAA GTA ACT TGT GAT TG	48	This study
GS005R	CCA CCA TCC AAT AAA AGC AA	48	This study
GS006R	CTC GCG TCA ATC ATT TCT AA	48	This study
GS007F	TTG AGA GGA GAA ATC ATG GA	48	This study
GS007R	TCC ATG ATT TCT CCT CTC AA	48	This study
GS008R	TAG GAT CAG AAG GTA ATT GG	48	This study
GS009F	GTG TAG TAG CAT AGA TGA TT	46	This study
GS009R	AAT CAT CTA TGC TAC TAC AC	46	This study
GS010F	CTA ATA TCC TGC AAC AGC AA	48	This study
GS011F	AAT GAA TCT CAC CAT GTC TC	48	This study
GS012R	ACG CTT CTA GAT ATA CTC TC	48	This study
GSControlF	TGA CTA AAC GAA AAA AGG TTT TTT TCT CTA TTA A	55	This study
GSControlR	TTA ATA GAG AAA AAA ACC TTT TTT CGT TTA GTC A	55	This study
pETOrfFF	GAG AGG AGA <u>ACA TAT GGA</u> AAA TAA ACA TCA AAG	58	This study
pETOrfFR	CTC CTT TT**C TCG AGA** TTA TCA CAA CAA TTT AAT CTG	60	This study

Table 2.9: Oligonucleotides used in this study to amplify gene fragments.

Gene amplified	Oligonucleotide	Oligonucleotide sequence 5'→3'	Amplified product	Source
hly	p234	CAT CGA CGG CCA ACC TCG GAG A	417 bp fragment of the 3' region of the *hly* gene of *L. monocytogenes*	Fitter *et al.*, (1992)
	p319	ATC AAT TAC CGT TCT CCA CCA TTC		
ctpA	FB001	CAC ACT GTA ATG TTA ACT GG	396 bp fragment of the 3' region of the *ctpA* gene of *L. monocytogenes*	Bell, (2002)
	LM2004	TCC AAG TGC TGA GAA TAT AC		Webster, (2001)
cutR	CuTrF	GGA AGT TAC ATC GAC AGA AG	601 bp region of the 5' end of the ORF *cutR* of *L. monocytogenes*	This study
	CuTrR	ATG AAT CTA CTG GCA CTA ACC		This study
pCT0018, pCT0019	FB004	GAT ACA TCT AAG TCA GGC GA	469 bp region encompassing ORFs pCT0018 and pCT0019 of *L. monocytogenes*.	This study
	p1037	TGA TCC GAT GAC ACC TAT GA		Francis, (1996)
catP	catPF	AGA ACT GGT TAC AAT AGC GA	391 bp fragment of the 3' region of the *cat* gene of plasmid pKSV7	This study
	catPR	CTG CAT GAT AAC CAT CAC AA		This study
erm	ermF	GTG TTG ATA GTG CAG TAT CT	643 bp internal fragment of the *erm* gene of plasmid pG121	This study
	ermR	ATC TGG AAC ATC TGT GGT AT		This study

Chapter 3: Analysis of Plasmid DNA of *L. monocytogenes* strain DRDC8

3.1 Introduction

The key virulence determinants involved in the pathogenesis of *Listeria monocytogenes* have been studied extensively (see Vázquez-Boland *et al.*, (2001); Cossart, (2002); Dussurget *et al.*, (2004) for recent reviews). However, little is known about the role of plasmid DNA-associated genes for virulence of *L. monocytogenes*. The prevalence and genetic composition of plasmid DNA in *Listeria* species has been previously reported (Lebrun *et al.*, 1992; Perez-Diaz *et al.*, 1982; Peterkin *et al.*, 1992), but published reports describing plasmid-encoded proteins and their function are limited. To date (2009), only two *Listeria* plasmid DNA sequences have been completely sequenced and annotated[6].

Genes encoding proteins involved in cation transport are a common theme for *Listeria* plasmid DNA (refer to Section 1.7.2). This is apparently the case for plasmid DNA carried by strain DRDC8. For example, *ctpA* (a gene that encodes a copper transporting P-type ATPase carried by *L. monocytogenes* strain DRDC8 (Francis & Thomas, 1997a), is likely to be encoded by plasmid DNA carried strain DRDC8 (Bell, 2002). Indeed, the limited distribution of *ctpA* in Australian isolates of *L. monocytogenes* may be explained by carriage of this gene and accessory genes on plasmid DNA (Bell, 2002).

Previously described sequence data for DNA flanking the 5' end of *ctpA*, identified three open reading frames (ORFs) that encoded proteins similar to other copper transport accessory proteins (Bell, 2002) (described in more detail in Chapters 1 and 4). The protein encoded by one of these ORFs (pCT0017) is similar to the CopY negative transcriptional regulatory protein of the well-characterised *cop* operon of *E. hirae* (Strausak & Solioz, 1997). Other copper transport systems comprising CopY-like negative regulators in Gram-negative and Gram-positive bacteria are associated with other accessory proteins such as CopZ-like copper binding chaperone proteins (Reyes *et al.*, 2006). Based on this

---

[6] Since completion of the work described in this thesis, *Listeria monocytogenes* 08-5578 plasmid pLM5578 has been completely sequenced and annotated (GenBank Accession CP001603).

information, the hypothesis to be addressed by work presented in this chapter is that plasmid DNA from strain DRDC8 carries *ctpA* and flanking ORFs that encode proteins involved in copper homeostasis. The aims of work described are to:

1. Provide confirmatory evidence to demonstrate that the *ctpA* gene is encoded on plasmid DNA;

2. Identify additional accessory genes involved in copper homeostasis by extending the characterised sequence of plasmid DNA isolated from *L. monocytogenes* strain DRDC8.

3.2 Experimental design

1. To confirm that *ctpA* is located on plasmid DNA, Southern blot hybridisation analysis of plasmid DNA extracted from *L. monocytogenes* strain DRDC8 was conducted using a labelled *ctpA*-specific DNA probe. Mating experiments between a *ctpA*::*erm* insertion mutant strain of DRDC8 (DSE201) and an unrelated *ctpA* negative strain (EGD Kaufmann) of *L. monocytogenes* were conducted to determine whether the plasmid encoding *ctpA* was conjugative. DNA extracted from bacteriophage isolated from mitomycin induced DRDC8 cultures was also analysed by PCR to confirm that *ctpA* was not encoded on bacteriophage DNA. Transfection of *ctpA* negative *L. monocytogenes* strains with bacteriophage isolated from cultures of DSE201 was used to provide additional evidence to demonstrate that *ctpA* is not bacteriophage DNA-encoded.

2. Sequence data for plasmid DNA isolated from *L. monocytogenes* strain DRDC8 was extended by nucleotide sequence analysis of plasmid pCT200. This plasmid carries a *ca*. 28.5 kbp fragment of the *L. monocytogenes* plasmid pCT100. The sequence data was further extended by nucleotide sequence analysis of TempliPhi™ amplified DRDC8 plasmid DNA.

3.3 Results

3.3.1 *ctpA* is encoded by Plasmid DNA

3.3.1.1 Isolation of Plasmid pCT100 from *L. monocytogenes* strain DRDC8

The use of standard methods (described in Section 2.7.1.2, Method 1) to isolate large plasmid DNA from *L. monocytogenes* strain DRDC8 proved unsuccessful. Agarose gel electrophoresis analysis of DNA extracts showed isolation of plasmid DNA in very low concentrations with significant quantities of contaminating chromosomal DNA, irrespective of the culture volumes used (data not shown). Consequently, a more complex method of extraction (described in Section 2.7.1.2, Method 2) was employed to isolate large plasmid DNA. Gel electrophoresis of *Bam*HI digested preparations of CsCl gradient centrifugation purified plasmid DNA is shown in Figure 3.1. The plasmid (pCT100) comprised four large *Bam*HI fragments and was *ca*. 80 kbp in size.

3.3.1.2 Detection of *ctpA* in Plasmid DNA

Southern blot hybridisation analysis (methods described in Section 2.12) was used to provide evidence to confirm that *ctpA* is plasmid-encoded. Plasmid pCT100 was *Bam*HI digested and the DNA fragments separated by agarose gel electrophoresis (Figure 3.2 Panel A). The DNA was transferred from the gel to a nylon membrane and denatured, followed by hybridisation with a 396 bp digoxigenin-labelled DNA probe prepared from PCR amplified DRDC8 DNA using the *ctpA*-specific FB001/LM2004 oligonucleotide pair (refer to Table 2.9 for oligonucleotide description). Probe hybridisation was visualised using chemiluminescent DIG detection methods (refer to Section 2.12.4, Method 2). Preliminary experiments were conducted to optimise the detection system reaction conditions required to detect *ctpA*-specific probe-target DNA hybrids with minimal background signal.

The results obtained for the Southern hybridisation analysis are shown in Figure 3.2 Panel B. Labelled probe DNA hybridised specifically to the *ctpA* gene on a 28.5 kbp *Bam*HI fragment and a *ca*. 8.497 kbp *Hind*III fragment derived from the positive control plasmid pCT200 (refer to Table 2.3 for description). Probe DNA also hybridised to a *ca*.

28.5 kbp *Bam*HI fragment of plasmid pCT100. By contrast, probe DNA did not hybridise to *Bam*HI digested genomic DNA isolated from strain DRDC8 or strain DSE01PL (a plasmid cured variant of DRDC8, refer to Section 4.3.4 and Table 2.1). These results provided conclusive evidence to confirm that *ctpA* was carried by plasmid pCT100 isolated from *L. monocytogenes* strain DRDC8. Weak hybridisation of probe DNA to the pBR322 negative vector control is likely to be due to the high concentration of DNA present as no other non-specific probe hybridisation was observed.

3.3.2 c*tpA* is not encoded on Bacteriophage

3.3.2.1 Induction and Isolation of Bacteriophage from DRDC8

Bacteriophage was induced from cultures of *L. monocytogenes* strain DRDC8 and strain DSE201 (Sm^R, *ctpA*::*erm*, see Table 2.1) using Mitomycin C treatment as described in Section 2.22.1. Bacteriophage particles were isolated from the supernatants of induced cultures. This was followed by purification and confirmation of isolation by electron microscopy as described in Section 2.22.2.

3.3.2.2 Detection of *ctpA* by Differential PCR Analysis of Bacteriophage DNA

DNA was extracted from bacteriophage isolated from cultures of strains DRDC8 and DSE201 using methods described in Section 2.7.3. Following extraction, bacteriophage DNA was serially diluted (1:10, 1:20, 1:100 and 1:1000) and used as template DNA for differential PCR amplification of *ctpA* and the *L. monocytogenes* chromosomal marker gene, *hly*. The p234/p319 oligonucleotide pair (refer to Table 2.9 for oligonucleotide description) was used to amplify a 417 bp fragment from the 3' region of the *L. monocytogenes hly* gene and the FB001/LM2004 oligonucleotide pair to amplify a 396 bp fragment of the 3' region of the *ctpA* gene. The aim of this approach was firstly to detect *L. monocytogenes* DNA contaminants and secondly to dilute *L. monocytogenes* DNA contamination to extinction. As bacteriophage DNA was expected to be present in significantly higher concentrations than *L. monocytogenes* DNA contamination, in the event that *ctpA* is encoded by bacteriophage DNA, the amount of *hly* specific PCR product should have decreased disproportionately to that generated by *ctpA*. The results of this PCR analysis are shown in Figure 3.3. Amplification of *ctpA* and *hly* were not

independently affected by dilution. DNA preparations diluted to 1/1000 allowed amplification of both *ctpA* and *hly* with *ca*. equal efficiency. DNA amplicons representing *ctpA* and *hly* were amplified from DRDC8 DNA and were absent for the no DNA negative control reactions. These results suggested that contaminating *L. monocytogenes* DNA, rather than bacteriophage DNA, was likely to be the template for amplification of *ctpA*.

3.3.2.3 Transfection of *ctpA*⁻ strains with Bacteriophage isolated from DSE201

To further eliminate the possibility that *ctpA* was carried on bacteriophage DNA, *L. monocytogenes* strains KE501, KE795, KE1003 and EGD Kaufmann, which do not carry the *ctpA* gene, were used as host strains for transfection of bacteriophage (Section 2.22.3 for method description) purified from cultures of the DRDC8 derivative *ctpA*::*erm* strain DSE201 (refer to Table 2.1 for strain description). Transfectants were selected by plating on BHI agar containing Em (8 μg/ml). If *ctpA* was encoded by bacteriophage DNA, transfection of *ctpA*⁻ *L. monocytogenes* strains with bacteriophage isolated from strain DSE201 was expected to transfer the *ctpA*::*erm* mutation and confer erythromycin resistance to the recipient bacteria.

Following transfection, serial dilutions of bacterial cultures were subsequently plated onto BHI agar containing Em. Undiluted bacteriophage preparation was also plated onto the same media as a negative control. A second negative control, consisting of 100 μl of the transfection culture mix without the addition of bacteriophage for each of the four recipient strains used, was also included in the analysis. None of the transfected recipients yielded Em^R colonies when plated on BHI containing Em. These results indicated that *ctpA* was unlikely to be encoded by bacteriophage DNA.

3.3.3 Plasmid pCT100 is Non-conjugative

To determine whether plasmid pCT100 isolated from strain DRDC8 is conjugative, a series of filter mating experiments (refer to methods described in Section 2.11) were conducted to demonstrate transfer of pCT100 from the DRDC8 derivative DSE201 (Sm^R, *ctpA*::*erm*) to a rifampicin resistant variant of *L. monocytogenes* strain EGD Kaufmann

(EGD-Rif) (refer to Table 2.1)[7]. EGD-Rif transconjugants were selected by plating on BHI agar containing Em (8 μg/ml), Rif (50 μg/ml). Single Em^R, Rif^R resistant colonies were patched onto fresh BHI agar plates containing Em, Rif and Em, Sm (50 μg/ml) to identify transconjugants that were rifampicin and erythromycin resistant and streptomycin sensitive. Three of fifteen putative transconjugants were found to be Rif^R, Em^R, Sm^S and selected for further analysis. Control cultures of DSE201 and EGD-Rif did not grow on BHI containing Em, Rif or Em, Sm.

To confirm that the Rif^R, Em^R, Sm^S colonies were the result of conjugal transfer of plasmid pCT100 from DSE201 to strain EGD-Rif, PCR analysis using the *ctpA* gene specific LM2004/FB001 oligonucleotide pair showed that these strains were positive for *ctpA* (Figure 3.4). These preliminary observations suggested that the selected isolates were *L. monocytogenes* strain EGD-Rif that had acquired the *ctpA* gene as a result of conjugative transfer of plasmid pCT100 from the DRDC8 derivative strain DSE201.

To further confirm that the three Rif^R, Em^R, Sm^S isolates were the result of conjugal transfer of pCT100 from DSE201 to EGD-Rif, the nucleotide sequences of *cutR*[8] and the O serotype of each isolate were compared to those of the donor and recipient.

The nucleotide sequence of PCR product generated from DNA extracted from each isolate using the CuTrR/CuTrF oligonucleotide pair (see Table 2.9 for oligonucleotide description) was determined and the sequences aligned with ClustalX (refer to Section 2.16.3). The sequence data obtained for each of the three isolates was identical to the *cutR* sequence of DSE201, but significantly different to that of EGD-Rif (the nucleotide sequence alignment for one of the putative EGD-Rif transconjugants is shown in Appendix B, see page 326). This data indicated that the three Rif^R, Em^R, Sm^S isolates were likely to be spontaneous Rif^R resistant variants of DSE201 and not EGD-Rif transconjugants.

This data was confirmed by serotype analysis using *Listeria* O Antisera Type 1 and *Listeria* O Antisera Type Poly 1,4 (refer to Section 2.17.2) (data not shown). Strain EGD-Rif is serotype 1,4 and as expected reacted with both Type 1 and Poly 1,4 antisera. Strain

---

[7] Spontaneous rifampicin resistant strain of *L. monocytogenes* strain EGD Kaufmann isolated on BHI agar containing 100 μg mL^{-1} Rif.

[8] The nucleotide sequence of *cutR* from strain EGD-Rif differs from that of strain DSE201. See section 3.3.4 for details.

DSE201 belongs to serotype 4 and as expected was agglutinated by Poly 1,4 but not Type 1 antisera. By contrast, the Rif^R, Em^R, Sm^S isolates were agglutinated by Type Poly 1,4, but not Type 1 antisera and therefore have the same serotype as the DRDC8 derivative DSE201. These results, taken together with the *cutR* sequence data described above, confirmed that these strains were spontaneous Rif^R variants of DSE201.

A further thirty-six Rif^R, Em^R colonies were isolated from independent mating experiments using higher concentrations of rifampicin (200 μg/ml). Three additional putative Rif^R, Em^R, Sm^S transconjugants were identified. Subsequent *cutR* sequence and serotype analysis confirmed that these strains were also likely to be spontaneous Rif^R variants of DSE201 (data not shown). Similar results were obtained when a spectinomycin resistant variant[9] of EGD Kaufmann (EGD-Sp) was used as a recipient host. Overall, these results indicated that plasmid pCT100 carried by DRDC8/DSE201 is unlikely to be a conjugative plasmid.

3.3.4 Identification of the Putative ORF *cutR*

A BLASTX and BLASTN analysis (refer to Section 2.16.3) of five *Listeria* genome sequences (refer to Table 1.2), identified a chromosomal gene encoding a putative copper-translocating P-type ATPase in each genome. A ClustalX alignment of these genes was used to design oligonucleotide pairs CuTrF/CuTrR (refer to Table 2.9 for oligonucleotide description) and that would allow PCR amplification of a 601 bp region of the 5' end of the gene. Gel electrophoresis and sequence analysis confirmed amplification of a PCR product of the expected size (601 bp) from DNA extracted from *L. monocytogenes* strain DRDC8. BLAST analysis showed that the deduced polypeptide sequence of this product shared ≥ 97% identity to chromosomal copper-translocating P-type ATPases of several *Listeria* strains including *L. monocytogenes* EGD-e. Figure 3.5 presents an amino acid alignment of these results. The amino acid sequence of CtpA was also included in the alignment to show that no similarity was shared to the plasmid-encoded copper-transporting P-type ATPase (CtpA) of DRDC8.

---

[9] Spontaneous spectinomycin resistant strain of *L. monocytogenes* strain EGD Kaufmann isolated on BHI agar containing 100 μg mL^{-1} spectinomycin.

To further extend the DRDC8 sequence, oligonucleotides CuTrM2F and CuTrM2R (see Table 2.7 for oligonucleotide description) were designed to amplify a 1883 bp region of the gene. Gel electrophoresis and sequence analysis confirmed amplification of a PCR product of the expected size (1883 bp) (refer to Appendix C, page 327 for the nucleotide sequence). BLAST analysis of the deduced polypeptide extended sequence showed 100% sequence identity (627/627 amino acids) to copper-translocating P-type ATPases of the genomes of *Listeria monocytogenes* str. 4b F2365 (Accession number YP_014474), *Listeria monocytogenes* FSL R2-503 (Accession number ZP_05242105) and *Listeria monocytogenes* HPB2262 (Accession number ZP_05265336). Given the similarity shared with other chromosomally-encoded copper-translocating P-type ATPases from *Listeria* species, the protein product of this gene was implicated in copper transport. The identified gene was designated *cutR*.

3.3.5 Partial Nucleotide Sequence Analysis of Plasmid pCT100

Segments of plasmid pCT100 isolated from *L. monocytogenes* strain DRDC8 (see Section 3.3.1.1) were sequenced using two different approaches. DNA flanking the 5' end of the *ctpA* gene was sequenced from a plasmid clone (pCT200 (refer to Table 2.3)) prepared by Francis, (1996). *Hind*III subclones of pCT200 (see Section 3.3.5.1) were used as templates for nucleotide sequencing reactions. Sequence data obtained from pCT200 identified ORFs pCT0005, pCT0006, pCT0007, pCT0015 and pCT0016. TempliPhi™ amplification of plasmid pCT100 (Section 3.3.5.2) was used to derive DNA sequence data for ORFs pCT0001, pCT0024, pCT0025, pCT0026, pCT0027, pCT0028, pCT0029, pCT0030, pCT0031, and pCT0032.

3.3.5.1 Isolation and Subcloning of Plasmid pCT200

Plasmid pCT200 contains *ca.* 28.5 kbp *Bam*HI fragment of plasmid pCT100 inserted into the *Bam*HI site of pBR322 (Figure 3.6). This fragment carries a Tn917-*lacZ* fusion at the 3' end of *ctpA*. *Bam*HI digestion of plasmid pCT200 resulted in two bands, the expected 28.5 kbp insert fragment and the 4.361 kbp pBR322 vector fragment. *Hind*III digestion of pCT200 resulted in eight fragments *ca.* 11.627, 8.497, 5.473, 3.545, 1.613, 1.248, 0.858 and 0.355 kbp in size.

Southern hybridisation analysis was used to classify pCT200 *Hind*III fragments containing *ctpA*-encoding DNA and identify fragments of interest for sequence analysis. A 2.8 kbp *ctpA* specific DNA probe, constructed by PCR using the LM2004/FBEDF oligonucleotide pair, was used for the hybridisation. As expected, probe DNA hybridised to the 28.5 kbp *Bam*HI fragment of pCT200 (Figure 3.6). The labelled *ctpA* probe DNA also hybridised to the 8.497, 5.473 and 0.858 kbp *Hind*III fragments. This result was consistent with the presence of two *Hin*dIII restriction endonuclease recognition sites within *ctpA* (the *ctpA* (pCT0020) sequence is available from GenBank Accession U15554, presented in Appendix A, page 311).

All pCT200 *Hind*III fragments were subcloned by ligation with *Hind*III digested pUC19 (see Table 2.3). The resulting plasmid DNA was then used to chemically transform (refer to Section 2.9) competent *E. coli* XL10-Gold® (refer to Table 2.2) cells. Transformants were plated on LA containing ampicillin, IPTG and X-gal and incubated overnight at 37°C. White, *lac* negative colonies were selected, and plasmid DNA isolated (refer to Section 2.7.1.1, Method 1) and confirmed by *Hind*III restriction endonuclease digestion analysis (data not shown). Subclones carrying *Hind*III Fragment 3 (5.473 kbp), Fragment 4 (3.545 kbp), Fragment 5 (1.613 kbp), Fragment 6 (1.248 kbp), Fragment 7 (0.858 kbp) and Fragment 8 (0.355 kbp) were identified. Plasmid DNA was isolated from these subclones using methods described in Section 2.7.1.1, Method 2, and the *Hin*dIII inserts were partially sequenced from the pUC19 M13 forward and reverse sequence priming sites (refer to Table 2.6).

Subclones containing Fragment 3 (5.473 kbp) (plasmid pFB186) and Fragment 5 (1.613 kbp) (plasmid pFB190) (see Figure 3.7) were identified as containing fragments of interest and selected for further sequence analysis. All subclone nucleotide sequencing was performed by primer walking using the primers listed in Table 2.6. Analysis of the nucleotide sequence data for Fragment 3 carried on plasmid pFB186 identified seven putative ORFs (pCT0014, pCT0015, pCT0016, pCT0017, pCT0018, pCT0019, pCT0020). Fragment 5 carried on plasmid pFB190 was shown to encode three putative ORFs (pCT0006, pCT0007, pCT0008). Sequence analysis of these ORF is presented in Section 3.3.6.

3.3.5.2 TempliPhi^{TM} Amplification of Plasmid DNA

To extend the sequence data for plasmid pCT100 beyond that encompassed by the 28.5 kbp fragment carried on pCT200, TempliPhi^{TM} amplification (refer to Section 2.14.2 for details) of plasmid pCT100 was used to obtain template DNA for further sequence analysis. Briefly, plasmid DNA extracted from DRDC8 (Section 3.3.1.1) was used as template DNA for TempliPhi^{TM} amplification. Amplified product was confirmed by gel electrophoresis analysis (data not shown). Sequence analysis of TempliPhi^{TM} amplified plasmid DNA (refer to Section 2.16.2) was achieved by means of primer walking using the primers listed in Table 2.6.

3.3.6 Sequence Analysis and Annotation

A total of 37.279 kbp of contiguous nucleotide sequence data from plasmid pCT100 was constructed by a combination of sequence data obtained from this study and other studies (Bell, 2002; Akya, 2007; Webster, 2001; Francis and Thomas, 1996). This plasmid DNA sequence has been deposited in the GenBank database as Accession Number: U15554. The entire nucleotide sequence, complete with descriptions of all 32 identified ORFs and deduced peptides as well as relevant features, is shown in Appendix A (see page 300)[10]. This contiguous sequence was manually assembled following analysis using ClustalX 1.81 (Thompson *et al.*, 1994), BLASTN and BLASTX (Altschul *et al.*, 1990) alignment algorithms. All open reading frames (ORF) were identified using the Graphical Open Reading Frame analysis program hosted on the NCBI website[11]. Each identified ORF and respective polypeptide sequences were subjected to BLAST analysis to identify any similarity shared with sequence data in the NCBI GenBank Nucleotide and Protein databases (Altschul *et al.*, 1990). BLAST analysis and the Conserved Domain Database[12] search faculty was used to identify relevant conserved sequence domains and motifs (Marchler-Bauer *et al.*, 2005). Deduced polypeptide sequences of each ORF were also analysed for patterns, profiles and rules (motifs) stored in the PROSITE database[13]. The

---

[10] The naming scheme for these ORFs includes a "pCT" prefix. This is intended to indicate that they are located on plasmid pCT100 in *L. monocytogenes* strain DRDC8. This notation conforms with that used for *L. innocua* plasmid pLI100 (GenBank Accession NC_003383).
[11] www.ncbi.nlm.nih.gov/gorf/gorf.html
[12] www.ncbi.nlm.nih.gov/cdd
[13] www.expasy.org/tools/scanprosite/

Compute pI/Mw tool[14] was used to calculate the theoretical isoelectric point (pI) and molecular weight (Mw) of the translated protein sequence of each ORF. The SignalP 3.0 server[15] was used to predict the presence and location of signal peptide cleavage sites in the deduced polypeptide sequence of ORFs.

Only putative ORFs and their predicted functions characterised in this study are described in detail below. The genetic organisation, orientation and size of these ORFs with respect to all other ORFs characterised is illustrated in diagrammatic form in Figure 3.8. Regions of sequence data generated from subcloned fragments of pCT200 or TempliPhi™ amplified DNA are identified. In addition, this figure also shows the location of two regions of particular significance to this study: a copper gene cluster comprised of ORFs pCT0017, pCT0018, pCT0019, pCT0020, pCT0023, pCT0024, pCT0025, pCT0026 and pCT0027 and a putative plasmid replication region comprising ORFs pCT0030, pCT0031 and pCT0032. The significance of these regions is discussed in detail in Section 3.3.7.

3.3.6.1 pCT0001: An ORF encoding a DNA UV-damage Repair Protein.

ORF pCT0001 is comprised of 1281 nucleotides that span from base 3769 (ATG) to 2489 (TAA) on the complementary strand (shown in Figure 3.9 and Appendix A, page 302). This ORF encodes a 426 amino acid polypeptide sequence that is predicted to have a molecular weight of 48.8 kDa. A putative ribosome binding site is located at nucleotide 3787. BLAST analysis of the deduced polypeptide sequence of pCT0001 identified significant sequence similarity to the ImpB/MucB/SamB (pfam00817) family of proteins involved in UV protection that are typically 430 - 450 amino acids in size. Significantly, pCT0001 shares 96% (413/426 amino acids) sequence identity to a 426 amino acid protein similar to UV-damage repair proteins (Accession NP_569220) pli0067 of *L. innocua* plasmid pLI100 and 96% (412/426 amino acids) identity to a 426 amino acid ImpB/MucB/SamB family protein of *L. monocytogenes* strain 4b H7858 plasmid pLM80 (Accession ZP_00231649).

---

[14] www.expasy.org/tools/pi_tool.html
[15] www.cbs.dtu.dk/services/SignalP/

Typically the ImpB/MucB/SamB protein family is characterised by a conserved IMS region (pfam00817) and the Pol IV kappa region (cd03586) (Marchler-Bauer *et al.*, 2005). The Pol IV kappa domain is characteristic of members of the Y-family of DNA polymerases that transverse replication-blocking DNA lesions, such as cyclobutane pyrimidine dimers resulting from UV damage (Yang, 2003). Conserved IMS and Pol IV kappa regions were identified by BLAST analysis of pCT0001 spanning from amino acids 18 - 364 and 16 - 363 respectively.

In addition, analysis of the polypeptide sequence of ORF pCT0001 using the ScanProsite motif searcher identified a conserved UmuC domain (PROSITE entry PS50173) located at nucleotide 3155 and extending to base 3727. UmuC proteins are SOS-regulated proteins that help repair or prevent DNA damage caused by UV and chemicals and are often plasmid associated. The sequence similarity shared with the ImpB/MucB/SamB family of proteins involved in UV protection suggested that pCT0001 is likely to encode a protein involved in UV-damage DNA repair.

3.3.6.2 pCT0005: An ORF encoding a Transposase

pCT0005 is a 1434 nucleotide ORF spanning bases 7869 (ATG) to 6436 (TGA) on the complementary strand (refer to Figure 3.10 and Appendix A, page 303), and is predicted to encode a 56 kDa protein. A putative ribosome binding site is located at nucleotides 7882 - 7877 (AGGAGG). BLAST analysis of the deduced 477 amino acid polypeptide encoded by pCT0005 showed that it shares 95% (455/478 amino acids) similarity to a transposase-like protein Accession NP_569211) pli0058 of *L. innocua* plasmid pLI100 and a putative transposase (Accession ZP_00230422) of *L. monocytogenes* strain 4b H7858 plasmid pLM80. BLAST analysis also identified a conserved Mu-transposase C domain (pfam09299) and a Rve integrase core domain (pfam00665) spanning amino acid residues 350 - 412 and 154 - 334 respectively. The Mu-transposase C domain denotes the C-terminal end of the Mu family of transposases. The Rve integrase core domain is characteristic of integrase proteins that mediate integration of specific fragments of DNA into other region of DNA. Specifically, in the case of a virus, it is responsible for integration of the viral genome into the host chromosome. Integrases are composed of three domains, an amino-terminal zinc binding domain, the core domain and the carboxyl

terminal domain. The core domain is the central catalytic domain that acts as an endonuclease and catalyses the DNA transfer of the 3' ends of the insert DNA to the 5'ends of the integration site.

The high degree of sequence similarity pCT0005 shares with other ORFs encoding transposase-like proteins and the presence of relevant domains associated with transposable elements suggested that pCT0005 is likely to encode a transposase. Transposases are typically plasmid DNA associated determinants found in a range of bacteria.

3.3.6.3 pCT0006: An ORF encoding an Invertase/Recombinase

ORF pCT0006 encodes a 199 amino acid polypeptide with a predicted molecular weight of 22.7 kDa. This ORF begins at base 8028 and terminates at base 8627 (Figure 3.11 and Appendix A, page 303). A putative ribosome binding site is located at nucleotides 8015 - 8018 (AGGA). BLAST analysis indicated the deduced polypeptide sequence shared high similarity to proteins from the invertase/site-specific recombinase, resolvase protein family. Specifically, the deduced polypeptide sequence is identical to an invertase-like protein (Accession NP_569212) pli0059 *L. innocua* plasmid pLI100, and similar to a site-specific recombinase, resolvase family protein (Accession ZP_00230423) of *L. monocytogenes* strain 4b H7858 plasmid pLM80.

A conserved resolvase family signature pattern is located within of the polypeptide sequence of pCT0006. A site-specific recombinase active site (Y - [LIVAC] - R - [VA] - S - [ST] - x(2) - Q) (PROSITE entry PS00397) spans amino acids 5 - 13 and a Site-specific recombinase signature domain (G - [DE] - X(2) - [LIVM] – [E] - X – [V] - [LIVM] - [DT] - R - [LIVM] - [GSA]) (PROSITE entry PS00398) spans amino acids 56 – 68. The first consensus pattern represents a highly conserved N-terminal region that contains a serine residue that denotes an active site that is probably involved in covalent attachment to DNA. The second consensus pattern is based on a conserved region located approximately 50 residues upstream of the serine active site.

Site-specific recombination plays an important role in DNA rearrangement in prokaryotic organisms. Two families of recombinases exist (Argos *et al.*, 1986). The `phage integrase' family groups a number of bacterial, phage and yeast plasmid enzymes.

The `resolvase' enzyme family (Garnier *et al.*, 1987) contain an N-terminal catalytic and a dimerisation domain, that contains a conserved serine residue involved in the transient covalent attachment to DNA, and a C-terminal helix-turn-helix DNA-binding domain. In general, proteins from the resolvase family are 180 to 200 amino-acid residues in length.

3.3.6.4 pCT0007: An ORF encoding a Cadmium Resistance Regulator

pCT0007 is an ORF comprising 360 nucleotides that starts at base 8901 and terminates at base 9260 (shown in Figure 3.12 and Appendix A, page 304). This ORF is predicted to encode a 119 amino acid polypeptide with a molecular weight of 13.38 kDa. The deduced polypeptide sequence encoded by this ORF is identical to a number of proteins from the CadC Cadmium resistance regulator/accessory protein family found in Gram-positive bacteria. These include the cadmium resistance accessory-like protein (Accession NP_569213) pli0060 of *L. innocua* plasmid pLI100, CadC cadmium resistance regulator protein (Accession NP_862601) of *Lactococcus lactis* plasmid pAH82 and a cadmium efflux system accessory protein (Accession ZP_00231641) of *L. monocytogenes* strain 4b H7858 plasmid pLM80.

Further analysis of the polypeptide sequence using the ScanProsite motif searcher, identified an ArsR-type HTH domain (PS50987) and the bacterial regulatory protein ArsR family signature region (consensus sequence: C-X(2)-D-[LIVM]-X(6)-[ST]-X(4)-S-[HYR]-[HQ]) (PS00846)) that spans amino acids 23 - 118 and 57 - 75 respectively. The ArsR-type HTH domain is typically 90 - 100 amino acids in length and is characteristic of transcriptional regulators of the ArsR/SmtB family involved in stress-responses to heavy metal ions. These metal-sensing, transcriptional repressors regulate expression of metal resistance operons by specifically binding to the operator/promoter region. In the presence of metal ions the repressor dissociates from the DNA, thus permitting transcription of proteins involved in metal-ions efflux and/or detoxification. The ArsR/SmtB family includes transcription repressors responsive to Zn(II), As(III), Cd(II), Pb(II), Bi(III), Co(II), Ni(II), Cu(I), and Ag(I) (Busenlehner *et al.*, 2003; Liu *et al.*, 2004).

Immediately upstream of pCT0007, ORF pCT0008 encodes a 705 amino acid (refer to Appendix A, page 304) polypeptide that has significant sequence similarity to the cadmium-transporting ATPase family. The polypeptide sequence encoded by ORF

pCT0008 is identical to a cadmium-transporting ATPase-like protein (pli0061) of *L. innocua* plasmid pLI100 (Accession NP_569214), and similar to an E1-E2 cation-transporting ATPase (Accession ZP_00231642) of *L. monocytogenes* strain 4b H7858 plasmid pLM80. Analysis of DNA flanking pCT0007 and pCT0008 indicated that it is likely that these ORFs form a bi-cistronic operon involved in cadmium transport. A Sigma[70] type promoter is located upstream of pCT0007 with a -10 region located at bases 8857 - 8862 (TATAAT) and a -35 region at 8834 - 8839 (TTGAAT). A ribosome binding site (AGGA) is located at 8 bases upstream of the start codon for pCT0007, spanning from bases 8889 - 8892. The stop codon of pCT0007 and start codon of pCT0008 overlap (refer to Appendix A, page 304). A strong potential transcriptional terminator is located at bases 11385 - 11418 immediately following the pCT0008 stop codon (Akya, 2007).

Significantly, the sequence characteristics and genetic organisation of ORFs pCT0007 and pCT0008 is consistent with that of the cadmium efflux genes in *L. innocua* plasmid pLI100 and *L. monocytogenes* strain 4b H7858 plasmid pLM80, and the *cadAC* genes from plasmid pLm74 of *L. monocytogenes* strain Lm74 (Lebrun *et al.*, 1994b). It is likely that pCT0007 encodes a transcriptional regulator protein, responsible for regulation of expression of a cadmium transporting ATPase encoded by pCT0008.

3.3.6.5 pCT0015: An ORF encoding a Transposase identical to pCT0009 and pCT00011

ORF pCT0015 encodes a 226 amino acid polypeptide that is identical to ORFs pCT0009 and pCT0011 (refer to Appendix D, page 328 for polypeptide sequence alignment). The orientation of pCT0015 is identical to pCT0011 and opposite to pCT0009. ORF pCT0015 is 681 nucleotides in length, spanning nucleotides 22661 (ATG) to 21981 (TAA) (Figure 3.13 and Appendix A, page 309). A putative ribosome binding site is located at nucleotides 22673 - 22669 (AGGAG). The deduced polypeptide sequence has a predicted molecular weight of 26.7 kDa. ORF pCT0015 is flanked by ISSIs inverted repeats (see Figure 3.13), that are similar to the ISSIs elements of *Lactococcus lactis* (Ward *et al.*, 1996). Interestingly, the number of nucleotide bases between the ISSIs repeats and the start (54 bases) and stop codons (36 bases) for ORFs pCT0009, pCT0011 and pCT0015 is conserved with a one base variation in ORF pCT0009. These features are similar to those

described for plasmid pSK08 isolated from a strain of *Streptococcus lactis* (Polzin & Shimizu-Kadota, 1987).

BLAST analysis revealed the deduced polypeptide sequences of pCT0015 is identical to transposase-like proteins pli0024 (Accession NP_569178) (226/226 amino acids), pli0071 (Accession NP_569224) (226/226 amino acids) and pli0063 (Accession NP_569216) (225/226 amino acids) from *L. innocua* plasmid pLI100. ORF pCT0015 also shares 99% (225/226 amino acids) identity with transposase-like protein from *L. monocytogenes* strain 4b H7858 plasmid pLM80 (Accession ZP_00231645) and 71% (162/226 amino acids) with putative transposase (Accession NP_053043) of *Lactococcus lactis* plasmid pNZ4000. BLAST analysis of pCT0015 also identified an Rve Integrase core domain (pfam00665) that spans amino acids 65 - 211. The sequence similarity shared with ORFs encoding transposase-like proteins and identification of transposase related domains and ISSI elements associated with other transposable elements within the deduced peptide sequence of pCT0015 suggested this ORF is likely to encode a transposase.

3.3.6.6 pCT0016: An ORF encoding a Na^+-Driven Multidrug Efflux Pump

ORF pCT0016 begins at base 23209 and terminates at 24534 (shown in Figure 3.14 and Appendix A, page 309). This 1326 nucleotide ORF encodes a 441 amino acid polypeptide predicted to have a molecular weight of 47.2 kDa. BLAST analysis of this sequence showed significant similarity to the MatE (Multi antimicrobial extrusion protein) efflux pump protein family. This family of proteins is responsible for Na^+-driven multidrug efflux. Specifically, pCT0016 shares 72% (320/443 amino acids) identity to Na^+-driven multidrug efflux pump (Accession ZP_05648165) of *Enterococcus gallinarum* EG2 and 53% (228/423 amino acids) identity to MatE efflux family protein (Accession ZP_03625308) of *Streptococcus suis* 89/1591. BLAST analysis also identified three conserved domains within the deduced polypeptide sequence of pCT0016. These include two MatE domains (pfam01554 [amino acids 13 – 170] and pfam01554 [amino acids 233 – 393]) and a NorM domain (COG0534) characteristic of Na^+-driven multidrug efflux pumps (residues 1 – 428). The similarity shared with other Na^+-driven multidrug efflux pumps and identification of characteristic domains associated with this protein family,

provided substantial evidence to suggest that ORF pCT0016 encodes a Na$^+$-driven multidrug efflux pump.

3.3.6.7 pCT0024: An ORF encoding a Two-Component Response Regulator

pCT0024 encodes a 681 nucleotide ORF, spanning bases 33182 (ATG) to 32502 (TAA) on the complementary strand (shown in Figure 3.15 and Appendix A, page 313). The 226 amino acid polypeptide sequence encoded by pCT0024 is predicted to have a molecular weight of 25.6 kDa. A putative ribosome binding site was identified at nucleotides 33196 - 33192 (AAGGA). The deduced polypeptide sequence of pCT0024 has significant similarity to the prokaryotic cytoplasmic two-component, response regulator protein family of proteins. Specifically, pCT0024 shares 99% (224/226 amino acids) identity to the putative regulatory protein VanR of *Enterococcus faecalis* PC1.1 (Accession ZP_06623359), and 97% (220/226) to response regulators of *Enterococcus faecium* TX1330 (Accession ZP_03980126) and *Enterococcus faecium* 1,231,501 (Accession ZP_05665647).

As described in Section 1.4.2, response regulators function in unison with sensor histidine kinases to form two-component signal transduction systems that are responsible for detection of, and response to, changes in the environment. Response regulators typically contain a conserved N-terminal regulatory domain and a variable C-terminal effector domain. The regulatory domain, often called receiver domain, is typically comprised of approximately 125 residues and contains a conserved phosphorylatable aspartate residue that forms the active site. Phosphorylation of this residue activates the variable C-terminal effector domain of the response regulator, which triggers the cellular response. The C-terminal effector domain contains DNA and RNA polymerase binding sites.

The polypeptide encoded by ORF pCT0024 has an N-terminal response regulatory domain (PROSITE entry PS50110) that spans from amino acids 2 – 115 and a conserved active site (aspartate residue at amino acid 51) characteristic of response regulators. A consensus trans_reg_C (cd00383), response regulator C-terminal effector domain spans amino acid 133 – 233.

113

ORF pCT0024 is adjacent to pCT0023 (see Figure 3.8), a 1112 bp ORF that encodes a 370 amino acid polypeptide sequence similar to proteins of the histidine kinase family (Webster, 2001). Analysis of the deduced peptide sequence of this ORF identified four (H, N, D, G - box) of the five conserved motifs characteristic of histidine kinases (Wolanin *et al.*, 2002). The deduced polypeptide sequence is shorter than other members of the histidine kinase family and encodes only one transmembrane spanning domain. Typically these proteins encode two membrane spanning domains. The absence of one membrane spanning domain indicated that the polypeptide is truncated. It is possible that this apparent truncation is the result of a sequencing error[16].

The similarity of the polypeptide sequences encoded by pCT0024 and pCT0023 with other histidine kinases and response regulators suggested that together, these putative polypeptides comprise a two-component regulatory system.

3.3.6.8 pCT0025: An ORF encoding a Multi-Copper Oxidase

pCT0025 is an ORF comprising 1587 nucleotides that encodes a polypeptide with a predicted molecular weight of 59.95 kDa. This ORF spans bases 34796 (ATG) to 33210 (TAG) on the complement strand (shown in Figure 3.16 and Appendix A, page 313). BLAST analysis showed that the deduced 528 amino acid peptide sequence encoded by pCT0025 has significant similarity to multicopper oxidase proteins. Specifically, pCT0025 shares 97% (514/528 amino acids) identity to a putative multicopper oxidase (Accession ZP_ 03980127) of *Enterococcus faecium* TX1330, 95% (504/528) identity to a multicopper oxidase (Accession ZP_06623358) of *Enterococcus faecalis* PC1.1 and 95% (466/488 amino acids) identity to a blue copper oxidase CueO (Accession ZP_06679880) of *Enterococcus faecium* E1071.

Three domains characteristic of multicopper oxidases are located within the deduced polypeptide sequence of pCT0025. The Cu-oxidase_3 domain (pfam07732), the Cu-oxidase_2 domain (pfam07731) and the PRK10965 multicopper oxidase domain span from amino acids 128 - 235, 410 - 528 and 97 - 524 respectively. A multicopper oxidase

---

[16] Since completion of the work described in this thesis, this sequencing error has been resolved. Additional sequence encoding an N-terminal transmembrane spanning domain and putative sensor domain was identified (C. J. Thomas, personal communication).

signature 2 motif (PROSITE entry PS00080) is located within the Cu-oxidase_2 domain. This motif spans amino acids 510 - 521 and has the H-C-H-x(3)-H-x(3)-[AG]-[LM] consensus signature pattern. The first two histidines are copper type 3 binding residues and the C, the third H, and L or M are copper type 1 ligands.

In Gram-negative bacteria, multicopper oxidases share a significant degree of sequence and functional similarity, including an extensive methionine rich region located between amino acids 350 – 400 that is thought to function as a copper binding regulatory domain (Wernimont *et al.*, 2003). These proteins characteristically contain a signal peptide sequence, however, the deduced peptide sequence of pCT0025 does not encode a Type II signal peptide sequence nor does it contain a methionine rich domain. However, while not well described, it is clear from the available literature that these features may not be typical of multicopper oxidases of Gram-positive bacteria. For example, the LcoC multicopper oxidase of Gram-positive bacterium *Lactococcus lactis* also lacks conserved domains of Gram-negative multicopper oxidases (Liu *et al.*, 2002).

3.3.6.9 pCT0026: An ORF encoding a Truncated Copper Translocating P-type ATPase

pCT0026 is an ORF comprising 431 nucleotide bases and spans from base 35238 (ATG) to 34807 (TAA) on the complementary strand (shown in Figure 3.17 and Appendix A, page 314). This 143 amino acid ORF is predicted to have a molecular weight of 15.8 kDa. A putative ribosome binding site is located at nucleotides 35244 – 35247 (AGGA). This ORF overlaps the 3' end of pCT0011 by 16 nucleotide bases. BLAST analysis of the deduced peptide sequence of pCT0026 indicated similarity to the copper-translocating P-type ATPase protein family from Gram-positive bacteria. Specifically, pCT0026 shares 89% (87/97 amino acids) identity to a 816 amino acid copper translocating P-type ATPase (Accession ZP_06679878) of *Enterococcus faecium* E1071 and a 816 amino acid heavy metal translocating P-type ATPase (Accession ZP_05665643) of *Enterococcus faecium* 1,231,501. Detection of conserved domains by BLAST showed that the amino acid sequence encoded by this ORF contained two heavy metal associated (HMA) domains (cd00371) spanning amino acids 4 - 67 and 76 – 138. The HMA is a conserved domain of approximately 30 amino acid residues found in a number of proteins implicated in

transportation and/or detoxification of heavy metals. These domains include the highly conserved consensus GMXCXXC motif important for binding and transferring of copper ions. In the case of copper, stoichiometry of binding is one copper atom per binding domain.

While pCT0026 contains two consensus HMA domains that are characteristic of heavy metal transporting P-type ATPase proteins, a comparison of the length of this polypeptide with other members of this family suggested the ORF is truncated. The predicted polypeptide is only 143 amino acids in length. Heavy metal transporting P-type ATPases typically range from 750 - 1000 amino acids in length. Specifically, the most similar proteins identified by BLAST analysis, range from 800 - 980 amino acids in length. Furthermore, the polypeptide sequence of pCT0026 shares similarity to only the N-terminal region of these proteins. This observation suggested that the C-terminal portion of this ORF is truncated. While it is probable that the protein encoded by this ORF is a truncated P-type ATPase, it is also possible that it may function as a copper chaperone. Certainly, the predicted protein has two putative copper binding domains that might be important for this activity.

3.3.6.10 pCT0027: An ORF encoding a Transcriptional Regulator Protein

ORF pCT0027 consists of 366 nucleotides and spans from base 35587 (ATG) to 35222 (TAA) (Figure 3.18 and Appendix A, page 314). A putative ribosome binding site is located at nucleotide 35591 - 35595 (AAGGA). The deduced polypeptide sequence encoded by pCT0027 is 121 amino acids in length and is predicted to have a molecular weight of 14.27 kDa. The polypeptide is similar to negative transcriptional regulator/repressor proteins. Specifically, the pCT0027 polypeptide shares 35% (32/90 amino acids) similarity to the CopY transcriptional regulator protein of the *copYZAB* operon of *Lactobacillus acidophilus* (Accession YP_194787), 31% (29/91 amino acids) similarity with a Transcriptional regulator protein of *Clostridium kluyveri* (Accession YP_001393806) and 30% (35/114 amino acids) similarity with a transcriptional repressor of the CopY family of *Clostridium beijerinckii* (Accession YP_001310197). The size of each of these polypeptides (120 - 150 amino acids) is consistent with that encoded by pCT0027.

A consensus Pencillinase_R, Penicillinase repressor domain (pfam03965) spans amino acids 10 – 100 and a consensus $WX_3TX_2TX_3RLX_2K$ N-terminal DNA binding motif spans amino acids 42 - 57. The N-terminal region of the penicillinase repressor domain is involved in operator recognition, while the C-terminal end is responsible for dimerisation of proteins belonging to this family.

Although the predicted polypeptide sequence is similar to members of the CopY family of transcriptional repressors, pCT0027 does not encode a C-terminal heavy-metal binding motif ($CXCX_4CXC$) or the N-terminal motif ($IX_3EXEVMX_2W$) characteristic of CopY proteins (Vats & Lee, 2001). Furthermore, DNA sequence flanking the 5' end of pCT0027 does not encode '*cop* box' CopY binding site motifs (TACAXXTGTA). In addition, pCT0027 shares limited similarity to the putative CopY protein encoded by ORF pCT0017 (refer to amino acid alignment presented in Appendix E, page 329). Nevertheless, these finding indicated that ORF pCT0027 is likely to encode a polypeptide similar to negative regulators involved in repression of expression of other genes.

3.3.6.11 pCT0028: An ORF encoding an Integral Membrane Protein

ORF pCT0028 consists of 615 nucleotides and spans from base 36217 (ATG) to 35603 (TGA) (shown in Figure 3.19 and Appendix A, page 315). This ORF is predicted to encode a 204 amino acid polypeptide with a molecular weight of 24.3 kDa. The polypeptide sequence shares similarity with integral membrane proteins from other Gram-positive bacteria. Significantly, pCT0028 was 99% similar to Integral membrane proteins from *Enterococcus faecium* 1,231,501 (202/204 amino acids) (Accession ZP_05665641) and *Enterococcus faecium* E1071 (199/200 amino acids) (Accession ZP_06679877). A DUF1461 domain (pfam07314) spans amino acids 19 – 175. This domain is characteristic of a family of proteins with unknown function, typically 200 residues in size, that are likely to be integral membrane proteins. A putative signal peptide cleavage site was identified between amino acid position 18 and 19 of the deduced peptide sequence of pCT0028.

3.3.6.12 pCT0029: An ORF encoding a Truncated or Inactivated Pro-lipoprotein Diacylglyceryl Transferase

pCT0029 is a truncated or inactivated ORF comprising 390 nucleotides, beginning at base 36787 and terminating at base 37176 (shown in Figure 3.20 and Appendix A, page 315). BLAST analysis of the 129 amino acid deduced polypeptide sequence of pCT0029 indicated significant similarity to the C-terminal region (amino acids 142 - 262) of proteins from Gram-positive bacteria, particularly the *Enterococcus* species. The sequence shares 99% (128/129 amino acids) identity with 267 amino acid prolipoprotein diacylglyceryl transferases of *Enterococcus faecalis* PC1.1 (Accession ZP_06623354), *Enterococcus faecium* 1,231,501 (Accession ZP_05665640) and *Enterococcus faecium* E1071 (Accession ZP_06679876). The C-terminal region (amino acids 1 – 99) comprises a conserved PRK12437 prolipoprotein diacylglyceryl domain (PRK12437). This signature domain is typically 240 amino acids in length. Members of the prolipoprotein diacylglyceryl transferase protein family are typically 270 - 274 amino acids in length. This indicated that pCT0029 is either truncated or inactivated at the N-terminal region.

However, further examination of the sequence data 5' upstream of the putative start codon of ORF pCT0029 identified a 143 amino acid stretch of coding region that is in frame with the described ORF (refer to Figure 3.20). BLAST analysis of this 143 amino acid peptide sequence revealed that it shared 99% (137/138 amino acids) to amino acids 1 - 138 (N-terminal region) of the prolipoprotein diacylglyceryl transferase of *Enterococcus faecium* 1,231,501 (Accession ZP_05665640) and *Enterococcus faecium* E1071 (Accession ZP_06679876). Furthermore the remaining N-terminal portion (approximately 140 amino acids) of the PRK12437 prolipoprotein diacylglyceryl conserved domain was identified within this additional coding region. This data suggested that the start codon (ATG) identified at position 36787 is an internal methionine. Detailed sequence analysis of both strands the 5' end of this extra coding region stretch failed to identify any sequencing errors in the region of sequence encoding the annotated ATG start codon. Consequently, it is possible that this ORF has been inactivated by point mutations between bases 36356 - 36366.

3.3.6.13 pCT0030: An ORF encoding a Plasmid Replication Protein

pCT0030 is an incomplete ORF that is 450 nucleotides in length, spanning from base 1 to 450 on the complement strand (shown in Figure 3.21 and Appendix A, page 301). A putative ribosome binding site (AGGA) was identified at base 455 to 458. This ORF encodes for a 163 amino acid polypeptide that shares significant similarity with a number of *L. monocytogenes* plasmid replication genes that are typically 500 - 550 amino acids in size. This includes; 100% identity (150/150 amino acids) to 544 amino acid plasmid replication proteins of *Listeria monocytogenes* FSL J1-194 (Accession ZP_05231473) and *Listeria monocytogenes* FSL R2-503 (Accession ZP_05243604), and 72% identity (108/150 amino acids) to 549 amino acid plasmid replication proteins of *L. innocua* plasmid pLI100 (Accession NP_569223) and *L. monocytogenes* strain 4b H7858 plasmid pLM80 (Accession ZP_0231652). In addition, this region has 35% identity (45/127 amino acids) to the Rep63A replication protein of *Bacillus thuringiensis* plasmid pAW63 (Accession CAB43191). Rep63A is a 513 amino acid protein that essential for plasmid replication and has been previously characterised by (Wilcks *et al.*, 1999).

This family of plasmid replication proteins are typically 549 amino acids in length and contain the characteristic C-terminal alpha helical Primase C terminal 1 domain (PriCT_1 domain (pfam08708)). This domain is not present in the deduced polypeptide sequence for pCT0030.

Additional plasmid pCT100 sequence data is required to complete the sequence for this ORF. Nevertheless, the data available indicated that a complete pCT0030 sequence would likely encode a polypeptide similar to plasmid replication proteins involved in initiation of plasmid replication.

3.3.6.14 pCT0031: An ORF encoding a Copy Control Protein Essential for Plasmid Replication

pCT0031 is an ORF comprising 1047 nucleotides that begins at base 1118 and terminates at base 2164 (shown in Figure 3.22 and Appendix A, page 301). A putative ribosome binding site is located at nucleotide 1105 - 1109 (AAGGA). The coding region of this ORF has a 22 base overlap with the 3' end of ORF pCT0032. The deduced 348 amino acid sequence encoded by this ORF is predicted to have a molecular weight of 40 kDa.

119

This sequence is similar to proteins expressed by Gram-positive bacteria that are involved in plasmid replication and partitioning. Most significantly, pCT0031 shares 100% identity (324/324 amino acids) to conserved hypothetical proteins of *Listeria monocytogenes* FSL J1-194 (Accession ZP_05231475) and *Listeria monocytogenes* FSL R2-503 (Accession ZP_05243603). This polypeptide sequence also shares 57% identity (166/291 amino acids) to the Rep63B protein of *Bacillus thuringiensis* plasmid pAW63 (Accession CAB43193). This 308 amino acid protein is a putative copy control protein essential for replication of plasmid pAW63 (Wilcks *et al.*, 1999).

Several conserved motifs consistent with a function in plasmid replication and partition in Gram-positive bacteria are present. These include two ParA domains (cd02042) and a Soj domain (COG1192) located at amino acid positions 49 - 102, 169 - 236 and 48 - 306 respectively. In addition, a P-loop site (CDD:73302) was also identified within first ParA domain, located at amino acids 56 - 62. The P-loop site functions as an ADP binding site and is characterised by a KGGVGKT consensus motif. A magnesium ion binding site (CDD:73302) characteristic of the P-loop is located at amino acid 62.

ParA and ParB of *Caulobacter crescentus* belong to a family of conserved bacterial proteins that are implicated in chromosome segregation. ParA ATPase activity is regulated by ParB through promoting nucleotide exchange. ADP-bound ParA binds single-stranded DNA and ATP-bound ParA dissociates ParB from its DNA binding sites. ParA shares sequence similarity to a widespread and conserved family of ATPases involved in plasmid replication and partition. The Soj domain is characteristic of ATPases involved in chromosome partitioning and cell division. Significantly, these motifs and other sequence characteristics are consistent with that described for rep63B from *Bacillus thuringiensis* plasmid pAW63 (Wilcks *et al.*, 1999). Collectively, this data suggested that pCT0031 is likely to encode a protein involved in copy control of replication of plasmid pCT100.

3.3.6.15 pCT0032: An ORF encoding a Plasmid Replication and Copy Control-Associated Protein

ORF pCT0032 comprises 297 nucleotides and begins at base 2142 and terminates at 2438 (shown in Figure 3.23 and Appendix A, page 302). A putative ribosome binding site (AAGGAG) was identified at nucleotides 2126 - 2131. The coding region of this ORF

shares 22 bp with ORF pCT0031, overlapping at the 3' end. The deduced polypeptide sequence of ORF pCT0032 is 98 amino acids in length and has an estimated molecular weight of 11.57 kDa. Most significantly, pCT0032 shares 100% identity (98/98 amino acids) to conserved hypothetical proteins of *Listeria monocytogenes* FSL J1-194 (Accession ZP_05231476) and *Listeria monocytogenes* FSL R2-503 (Accession ZP_05243602). This polypeptide sequence also shares 33% (34/102 amino acids) identity to the hypothetical protein Rep63C found in *Bacillus thuringiensis* conjugative plasmid pAW63 (Accession AAZ06620). Rep63C is a 102 amino acid putative plasmid replication and copy control-associated protein that shares analogous characteristics with RepC from *Enterococcus faecalis* conjugative plasmid pAD1 (Wilcks *et al.*, 1999). RepC is implicated in regulating stable inheritance of pAD1 (Weaver *et al.*, 1993). This suggested that pCT0032 is likely to encode a small protein that functions in plasmid replication and copy control of plasmid pCT100.

3.3.7 Significant Features of Plasmid pCT100

3.3.7.1 Conserved DNA Regions

Of the 32 ORFs identified and characterised in the 37.279 kbp contiguous DNA fragment from plasmid pCT100, 13 (pCT0030, pCT0001, pCT0002, pCT0004, pCT0005, pCT0006, pCT0007, pCT0008, pCT0009, pCT0011, pCT0013, pCT0014, pCT0015) share significant similarity to genes encoded on the *L. innocua* plasmid pLI100 (GenBank Accession NC_003383) (Glaser *et al.*, 2001) and *L. monocytogenes* strain 4b H7858 plasmid pLM80 (GenBank Accession AADR01000010/AADR01000058) (Nelson *et al.*, 2004). Furthermore the arrangement and organisation of these ORFs is consistent with that of plasmids pLI100 and pLM80. To map and visualise the regions of conserved sequence between plasmids pCT100, pLM80 and pLI100, Mauve[17] was used to align the nucleotide sequences and map regions of conservation. The results of this analysis are presented in Figure 3.24. Four distinct regions of conservation were identified. Significantly, the figure shows that these conserved regions are commonly flanked by transposable DNA elements.

3.3.7.2 Copper Gene Cluster unique to *ctpA* positive strains

BLAST analysis and identification of characteristic conserved domains has implicated nine ORFs (pCT0017, pCT0018, pCT0019, pCT0020, pCT0023, pCT0024, pCT0025, pCT0026 and pCT0027) as putative proteins involved in copper ion homeostasis. These ORFs represent a putative copper gene cluster that may be responsible for maintenance of intracellular copper ion levels (refer to Figure 3.8 for location).

Sequence analysis, conducted in this study and other studies (Bell, 2002; Webster, 2001; Francis and Thomas, 1996), showed that the ORFs comprising this putative gene cluster have not been previously described in any of the published *Listeria* sequences to date (refer to Table 1.2). To determine if ORFs associated with the copper gene cluster were unique to strain DRDC8 or characteristic of *ctpA* positive *L. monocytogenes* strains, PCR and sequence analysis was used to identify DNA characteristic of the cluster in other isolates. Preliminary PCR analysis involved testing strains Dairy 25, Dairy 26, NR28/1, and TSANR44/1 (refer Table 2.1 for strain description) previously characterised as *ctpA* positive (Bell, 2002) for DNA encoding ORFs pCT0018 and pCT0019 using FB004/p1037 oligonucleotide pair (refer to Table 2.6 for oligonucleotide description). The expected 469 bp amplicon, identical to that from DRDC8, was produced for the four isolates tested. Sequence analysis was used to confirm that the nucleotide sequence of the amplified products was identical to that of DRDC8 (data not shown).

A more comprehensive PCR analysis was conducted for an additional 10 strains characterised as *ctpA* positive (Bell, 2002) (DC1141, Dairy 22, Dairy 25, Dairy 26, NR28/1, DTS5708, ING2, TSANR31/5, DC1059, and DC1060), one *ctpA* negative strain (KE391), and six recently acquired uncharacterised strains (0121, 0124, 0146, 0177, 0179 and 24/3) (refer to Table 2.1 for strain description). DNA isolated from each strain was used as template to amplify a 396 bp fragment of *ctpA* using the FB001/LM2004 primer pair and a 1339 bp fragment of ORF pCT0025 using the pDR6/pDR10R primer pair (refer to Table 2.6 for oligonucleotide description). Amplicons indicative of *ctpA* (396 bp) were obtained for all previously characterised *ctpA* positive strains, but no DNA fragment was amplified from DNA from the six uncharacterised strains (0121, 0124, 0146, 0177, 0179

---

[17] Constructed using the Mauve 2 (Progressive Mauve) genome alignment system software from

and 24/3) (Figure 3.25). As previously observed by Bell, (2002), qualitative estimates of the amount of *ctpA* specific PCR product generated by some isolates (DRDC8, DSE201, DC1141, Dairy 22, Dairy 25, Dairy 26 and NR28/1) was greater than amplified by other strains (DTS5708, ING2, TSANR 31/5, DC1059 and DC1060) (see Section 1.9.3 and Figure 1.11). The data presented in Figure 3.26 shows that amplicons indicative of ORF pCT0025 (1339 bp) were only observed from *ctpA* positive isolates that consistently yielded greater amounts of *ctpA* product (refer to Figure 3.25). Amplicons indicative of *ctpA* and pCT0025 were amplified from the positive control strains DRDC8 and DSE201, but not from the negative control strain DSE201PL (see Section 4.3.4 and Table 2.1 for strain description), and strains EGD Kaufmann and KE391 (refer to Figure 3.25 and Figure 3.26). Sequence analysis was used to confirm that the nucleotide sequence of the amplified products was identical to that of DRDC8 (data not shown). These results indicated the *ctpA* positive strains DRDC8, DSE201, DC1141, Dairy 22, Dairy 25, Dairy 26 and NR28/1 may also carry the putative copper gene cluster.

3.3.7.3 Putative Plasmid Replication Region

ORFs pCT0030, pCT0031 and pCT0032 represent a putative plasmid replication region for pCT100. These ORFs share significant similarity with proteins involved in plasmid replication (refer to Section 3.3.6.13, Section 3.3.6.14, and Section 3.3.6.15). The genetic arrangement of pCT0030, pCT0031 and pCT0032 is consistent with other Gram-positive plasmid replicons, including the *rep63A, rep63B*, *rep63C* replicon of *Bacillus thuringiensis* conjugative plasmid pAW63 (Wilcks *et al.*, 1999) and the *repA*, *repB*, *repC* replicon of *Enterococcus faecalis* plasmid pAD1 (Weaver *et al.*, 1993). Figure 3.27 diagrammatically shows the similarity shared between the putative replicon of plasmid pCT100, and the pAW63 and pAD1 replicons. A putative region of repeated DNA motives called iterons, was identified between ORFs pCT0030 and pCT0031. This is consistent with the arrangement of the pAW63 and pAD1 replicons. The iteron-like region of pCT100 consists of seven direct and one inverted repeat of a 10 bp repeat (TCTTTTTATC). Four CTATATA repeats and two AT-rich (80 – 85%) DNA regions were also identified. In addition, a 14 bp region comprising two 7 bp repeats (TCCTCTCTCCTCTC) and an 18 bp

perfect inverted repeat (AAAGATAATTAGATAAAA) were identified within the putative promoter region of pCT0030 and downstream of pCT0032 respectively. The location of the downstream inverted repeat is consistent with that described for the pAD1 replicon (Weaver *et al.*, 1993). Relevant features of the putative replicon of pCT100 are shown in Figure 3.28.

3.4 Discussion

The purpose of work described in this chapter was to establish whether *ctpA* was located on plasmid DNA or other mobile genetic elements. In addition, this work used sequence analysis to identify and characterise genes which flank *ctpA* in *L. monocytogenes* strain DRDC8. The following discussion addresses each of these aims.

3.4.1 *L. monocytogenes* strain DRDC8 carries *ctpA* on Plasmid DNA

Studies by Bell, (2002) first provided evidence that suggested that the *ctpA* gene, which is significant for the persistence of *L. monocytogenes* strain DRDC8 during infection of the liver and spleen of mice (Francis & Thomas, 1997b), was likely to be encoded by plasmid DNA in DRDC8. Experimental data presented in this thesis provided direct evidence to support this hypothesis. Southern hybridisation showed that *ctpA* is located on a distinct *Bam*HI fragment of plasmid pCT100 from DRDC8. No evidence was obtained to suggest that *ctpA* is encoded by bacteriophage DNA. Furthermore, the DNA sequence that flanks *ctpA* encodes proteins involved in plasmid replication and other plasmid-associated functions. For example:

- ORF pCT0001 encodes a putative UV-damage repair protein responsible for prevention or repair of plasmid DNA damage.
- ORFs encoding putative cation-transporting ATPases (eg. Sodium, cadmium) typical of those also encoded on other *Listeria* plasmids. This is consistent with reports that most heavy metal ion transporters are carried on plasmids in bacteria (Dabbs & Sole, 1988; Silver & Ji, 1994; Silver, 1996a; Silver, 1996b; Silver & Phung, 1996).
- ORFs encoding transposase/integrase-like proteins which flank sequence data that encodes large, conserved sequence domains that are present in other *Listeria* plasmids.

Collectively, this data provided compelling evidence to demonstrate that the *ctpA* gene is encoded by plasmid pCT100 in *L. monocytogenes* strain DRDC8. Indeed, genes involved in cadmium transport/resistance in *L. monocytogenes* have been described as more frequently plasmid-borne than chromosomal (Lebrun *et al.*, 1994b). This may be the first report of a plasmid-encoded copper transport gene that is also important for virulence in the *L. monocytogenes* genus. In addition, the data presented in this chapter indicated that plasmid pCT100 is non-conjugative.

3.4.2 Features of Plasmid pCT100 from Sequence Analysis

Strain DRDC8 carries a large plasmid (pCT100) with an estimated size of 80 kbp. The size of this plasmid was consistent with plasmid pLI100 (Accession NC_003383) from *L. innocua* CLIP11262 and the estimated size of plasmid pLM80 (Accession AADR01000010, AADR01000058) from *L. monocytogenes* strain 4b H7858. This thesis presents the nucleotide sequence data for several fragments of plasmid pCT100 from DRDC8. In conjunction with other co-workers, 37.279 kbp of contiguous nucleotide sequence for plasmid pCT100 has been assembled (presented in Appendix A, see page 300). Sequence analysis has identified ORFs that encode proteins involved in heavy metal transport, DNA transcription regulation, plasmid DNA replication, DNA transposition/insertion, and other plasmid DNA associated functions. Blocks of homologous and colinear DNA has been identified between the nucleotide sequence data for pCT100, and *L. innocua* plasmid pLI100 and *L. monocytogenes* strain 4b H7858 plasmid pLM80. Typically these blocks included genes implicated in plasmid DNA replication, DNA transposition/insertion and heavy metal transport, including cadmium. Six putative ORFs encoding transposase/integrase-like proteins, several of which are duplicated and identical to proteins encoded by genes on pLI100 and pLM80, were identified. In addition, a putative cadmium efflux operon, encoded by ORFs pCT0007 and pCT0008, similar to that of pLI100 and pLM80 was identified.

Significantly, ORFs pCT0030, pCT0031 and pCT0032 were similar to plasmid replication proteins encoded by plasmids pLI100 of *L. innocua* CLIP11262 and pLM80 *L. monocytogenes* strain 4b H7858. These ORFs represent a putative plasmid DNA replication region that may be involved in replication and copy number control of plasmid

pCT100. The genetic composition and arrangement of this putative replicon is consistent with that of the pAMβ1 family of Gram-positive theta-replicating plasmids, including the Rep63A, Rep63B and Rep63C replicon of the *Bacillus thuringiensis* plasmid pAW63 (Wilcks *et al.*, 1999) and the *repA*, *repB*, *repC* replicon of *Enterococcus faecalis* plasmid pAD1 (Weaver et al., 1993). Replication proteins of the pAMβ1 family are rate limiting factors for plasmid replication involved in plasmid maintenance (Brantl & Behnke, 1992). Replication is initiated at the origin (*ori*), a cis-functioning locus which is typically located downstream of the first replication protein. Additional plasmid pCT100 sequence data is required to identify the *ori* region. Other features of theta-replicating plasmid replicons include repeated DNA motives called iterons that are involved in replication and regulation of stability and/or copy control (del Solar *et al.*, 1998). Iteron repeats of the pAW63 replicon are described as 10 direct and 6 inverted repeats of an 8 bp sequence (AAAGATAC). The region also contained several abbreviated repeats of the consensus sequence, a direct repeat of 23 bp and three AT-rich (83 - 87%) DNA regions. The iteron region of the pAD1 replicon consists of 25 direct TAGTAXXX repeats, and a 21 bp inverted repeat downstream of *repC*. The intergenic DNA regions of the putative replicon of pCT100 contain similar iteron features. These similarities suggested that pCT100 may be a member of the pAMβ1 plasmid family.

Although plasmid pCT100 contained colinear blocks of homologous sequence shared with other *Listeria* plasmids, pCT100 is unique by virtue of the fact that it contains a segment of DNA that encodes a series of polypeptide sequences that are similar to proteins from the genomes of other Gram-positive bacteria, particularly strains of the *Enterococcus* and *Lactobacillus* species. These proteins are dedicated to functions that include; copper transport, regulation of copper transport genes, copper binding or contain copper reactive centres. Significantly, this region shared no similarity with other *Listeria* sequences. The discrete location of these ORFs suggested they represent a collection or cluster of genes implicated in copper homeostasis/tolerance. Collation of sequence data presented in this thesis as well as that prepared by Bell (2002), Francis and Thomas, 1997(a) and Webster (2001), has identified nine discrete ORFs that form part of this gene cluster (ORFs pCT0017, pCT0018, pCT0019 and pCT0020 (*ctpA*), pCT0023, pCT0024, pCT0025, pCT0026, pCT0027). While the putative function and genetic organisation of the ORFs

within this cluster have been characterised, the organisation as transcriptional units is not yet known. ORFs pCT0017, pCT0023, pCT0024 and pCT0027 are predicted to encode polypeptides that are similar to transcriptional regulators and may be involved in regulation of expression of this cluster or parts there of.

Given that CtpA of *L. monocytogenes* DRDC8 has been proposed to function in copper influx (Francis and Thomas, 1997a), it is intriguing that the copper gene cluster of DRDC8 does not encode a CopB-like efflux protein. This observation is not unusual as *copB*-like genes are not typically found associated with CopY genes, except for *E. hirae* and some *Streptococcus* strains (Reyes *et al.*, 2006). It is possible however, that a *copB* homolog may exist elsewhere in the genome. Alternatively, there may be another unidentified gene that encodes a protein with CopB-like function. It is therefore interesting to speculate that the chromosomally-encoded putative gene (*cutR*) in strain DRDC8, which encodes a polypeptide similar to chromosomally-encoded copper-translocating P-type ATPases from other *Listeria* species, may be implicated in copper ion transport and could encode a CopB-like protein. This contention is supported by the lack of sequence similarity between the polypeptide encoded by *cutR* and CtpA. Thus CutR may provide an alternative mechanism for copper ion transport. However, it is unknown if CutR-like proteins function in efflux of copper ions in other *Listeria* species. CtpA and CutR may act in synergy or autonomously to maintain copper homeostasis.

In addition to a CopY-like negative regulator (encoded by putative ORF pCT0017), the copper gene cluster of DRDC8 also contains a putative two-component regulatory system comprising ORFs pCT0023 and pCT0024 that encode proteins similar to histidine kinases and response regulators, respectively. Regulation of copper homeostasis by two-component regulatory systems in other bacteria has been well described (Grass & Rensing, 2001a; Khunajakr *et al.*, 1999; Mills *et al.*, 1993; Munson *et al.*, 2000; Quaranta *et al.*, 2009). For example, in *Pseudomonas syringae* the plasmid-encoded *cop* operon is regulated by the CopRS two-component regulatory system in response to high levels of copper (see Section 1.8.2.2 for details) (Bender & Cooksey, 1987; Mellano & Cooksey, 1988b; Mills *et al.*, 1993). In the Gram-positive bacterium *Lactococcus lactis*, the plasmid-encoded LcoRS two-component regulatory system regulates expression of the copper resistance *locABC* operon in response to copper (Liu *et al.*, 2002). Intriguingly,

locA is predicted to encode a prolipoprotein diacylglycerol transferase and *lcoC* a multicopper oxidase. ORFs (pCT0025 and pCT0029) encoding similar putative proteins (or parts of) were identified within close proximity to pCT0023 and pCT0024. On this basis, ORFs pCT0023 and pCT0024 may encode a two-component regulatory system that acts in manner similar to the LcoRS system to regulate expression of genes involved in copper resistance/tolerance eg. pCT0025.

pCT0025 encodes a polypeptide similar to the multicopper oxidase proteins from other Gram-positive bacteria. Multicopper oxidases play a critical role in copper homeostasis by protecting periplasmic enzymes from copper-mediated damage (Roberts *et al.*, 2002). These proteins are dependant upon oxygen for their activity and typically possess three spectroscopically different copper centers; type 1 (or blue), type 2 (or normal) and type 3 (or coupled binuclear) (Solomon *et al.*, 1996). Characterised examples of multicopper oxidases involved in copper resistance in bacteria are generally limited to Gram-negative species. These include the CueO and PcoA multicopper oxidases of *E. coli* (refer to Section 1.8.2.3). These proteins are thought to have similar functions to detoxify copper by oxidising Cu(I) to less toxic Cu(II) and/or by oxidising catechol siderophores such as enterobactin, which can then sequester copper (Magnani & Solioz, 2007). A related multicopper oxidase (LcoC) is located on a plasmid conferring copper-resistance in the Gram-positive bacterium *L. lactis* (Liu *et al.*, 2002). In Gram-negative bacteria, multicopper oxidases are located in the periplasm. Since there is no periplasmic space in Gram-positive bacteria, it is thought that LcoC is anchored to the outer leaflet of the cytoplasmic membrane as a lipoprotein (Liu *et al.*, 2002). The protein encoded by pCT0025 may function in a LcoC-like manner to protect the cell from excess copper.

Interestingly, ORF pCT0027 encodes a putative negative transcriptional regulator that shares similarity with CopY-like proteins. Whilst the polypeptide sequence lacks the consensus motifs characteristic of CopY proteins, it does contain domains that are typical of DNA binding proteins, and operator recognition and protein dimerisation. This suggested that it is likely that the protein encoded by this ORF is a transcriptional repressor that may bind to DNA. However it may function differently to CopY proteins. This is supported by the fact that pCT0027 shared limited similarity to the putative CopY protein encoded by ORF pCT0017. Due to its close proximity, the protein encoded by ORF

pCT0027 may be involved in regulation of expression of ORF pCT0026. ORF pCT0026 has been identified as a truncated ORF that encodes the N-terminal region of a copper translocating P-type ATPase. Whilst it is unknown whether the protein encoded by this ORF is functional, it is significant that it contained the two consensus domains (HMA) important for copper binding and transfer.

Whilst examples of copper homeostasis mechanisms have been described in other Gram-positive and Gram-negative bacteria, the organisation of this putative copper gene cluster is unique. Although fundamentally unrelated, the gene cluster of plasmid pCT100 is perhaps most similar to the copper resistance *pco* operon (*pcoABCDRSE*) of *E. coli* plasmid pRJ1004 (refer to Section 1.8.2.3), primarily because of its complexity and the type of genes involved.

3.4.3 Distribution of the Copper Gene Cluster in isolates of
L. monocytogenes

A survey of Australian clinical and environmental isolates of *L. monocytogenes* highlighted the fact that prevalence of *ctpA* is limited to environmental isolates (Bell, 2002). This observation was confirmed and extended to other isolates by work described in this thesis. If all *ctpA* positive strains of *L. monocytogenes* carried this gene on non-conjugative, large plasmids, as is the case for strain DRDC8, then this observation may explain the limited distribution of *ctpA*. The fact that carriage of plasmid DNA is limited in clinical *L. monocytogenes* isolates and more prevalent in environmental isolates (Akya, 2007; Kolstad *et al.*, 1992; Perez-Diaz *et al.*, 1982; Peterkin *et al.*, 1992) supports this contention. If this is confirmed, lateral transfer of plasmid DNA between cohabitating organisms may be limited and dependent on the local environment. Indeed, a high incidence of genes involved in cadmium transport/resistance has been described in *L. monocytogenes* strains isolated from the environment and food (Lebrun *et al.*, 1994a).

Intriguingly, when *ctpA* amplicons are produced from *ctpA* positive Australian isolates of *L. monocytogenes*, two groups of strains can be identified by qualitative analysis of the amount of amplicon produced (Section 3.3.7.2). This data could be interpreted as a gene dosage effect. This may be achieved if transposition events result in transfer of copies of *ctpA* from plasmid DNA to chromosomal DNA and *vice versa*. Assuming plasmids such

as pCT100 are present in cells as multiple copies, the gene dosage for plasmid-encoded *ctpA* would be higher than that for *ctpA* encoded as a single copy on chromosomal DNA. Alternatively, different isolates may carry *ctpA* on plasmids that are maintained at different copy numbers. In addition, it is also intriguing that while carriage of the putative copper gene cluster is not confined to strain DRDC8, it may be limited to only some *ctpA* positive isolates, mainly from dairy origin. Although the proper controls required to quantitate the amount of each amplicon were not included, only strains that appeared to yield greater amounts of the *ctpA* product, were found to carry another ORF (pCT0025) associated with the copper gene cluster. There are several possible explanations for this finding. For example, some strains may carry a single copy of the *ctpA* gene independently of the copper gene cluster, while other strains may carry multiple copies of *ctpA*, with one or more copies associated with the copper gene cluster. Multiple copies of the gene could be carried on separate elements of DNA ie. plasmid and chromosome, or *ctpA* may be carried by multiple plasmids. Alternatively, an insertion event that resulted in acquisition of all DNA associated with the copper gene cluster may have taken place in only some *ctpA* positive strains. Conversely, it is also possible that some *ctpA* positive strains may have undergone a series of deletions that resulted in loss of copper gene cluster-associated DNA. These observations require further investigation.

3.5 Conclusions

The results described in this chapter provided compelling evidence to show that the *ctpA* gene is encoded by non-conjugative plasmid pCT100 in *L. monocytogenes* strain DRDC8. This may explain the limited distribution of the *ctpA* gene in *L. monocytogenes* isolates. Sequence data analysis has showed that plasmid pCT100 shares characteristics with plasmid pLI100 of *L. innocua* and plasmid pLM80 of *L. monocytogenes* strain 4b H7858. Conserved blocks of homologous and colinear DNA have been identified between these *Listeria* plasmids. A putative replication region was identified for plasmid pCT100 that is similar to that of plasmids pLI100 and pLM80, and other Gram-positive plasmids. In addition, a putative copper gene cluster carrying additional ORFs implicated in copper ion homeostasis/tolerance in strain DRDC8 and some other *ctpA* positive isolates was identified. Furthermore, a putative copper-translocating P-type ATPase encoded by the

chromosomal ORF *cutR* was also identified. *cutR* and *ctpA* may function collaboratively or autonomously to maintain copper ion homeostasis.
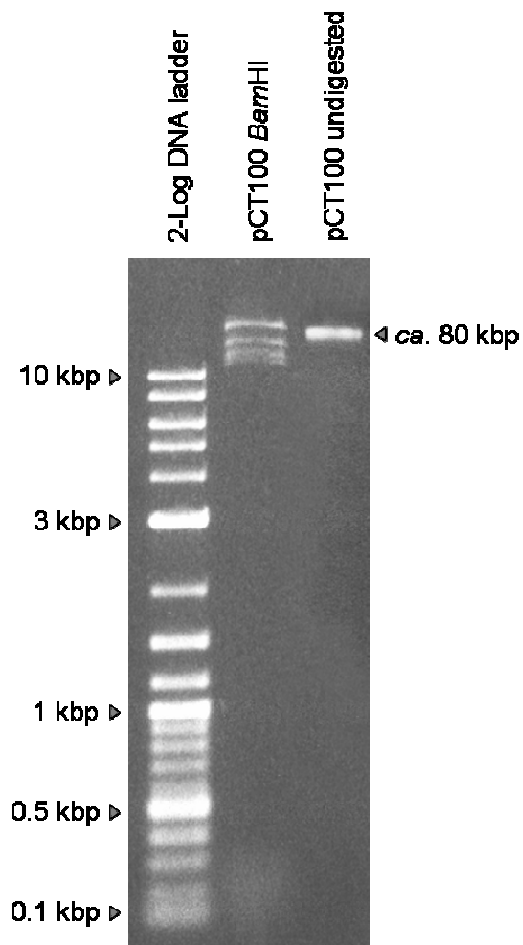
Figure 3.1: Restriction enzyme digestion of plasmid pCT100

*Bam*HI digestion of plasmid pCT100 yielded four large DNA fragments. Undigested plasmid pCT100 was *ca.* 80 kbp in size.
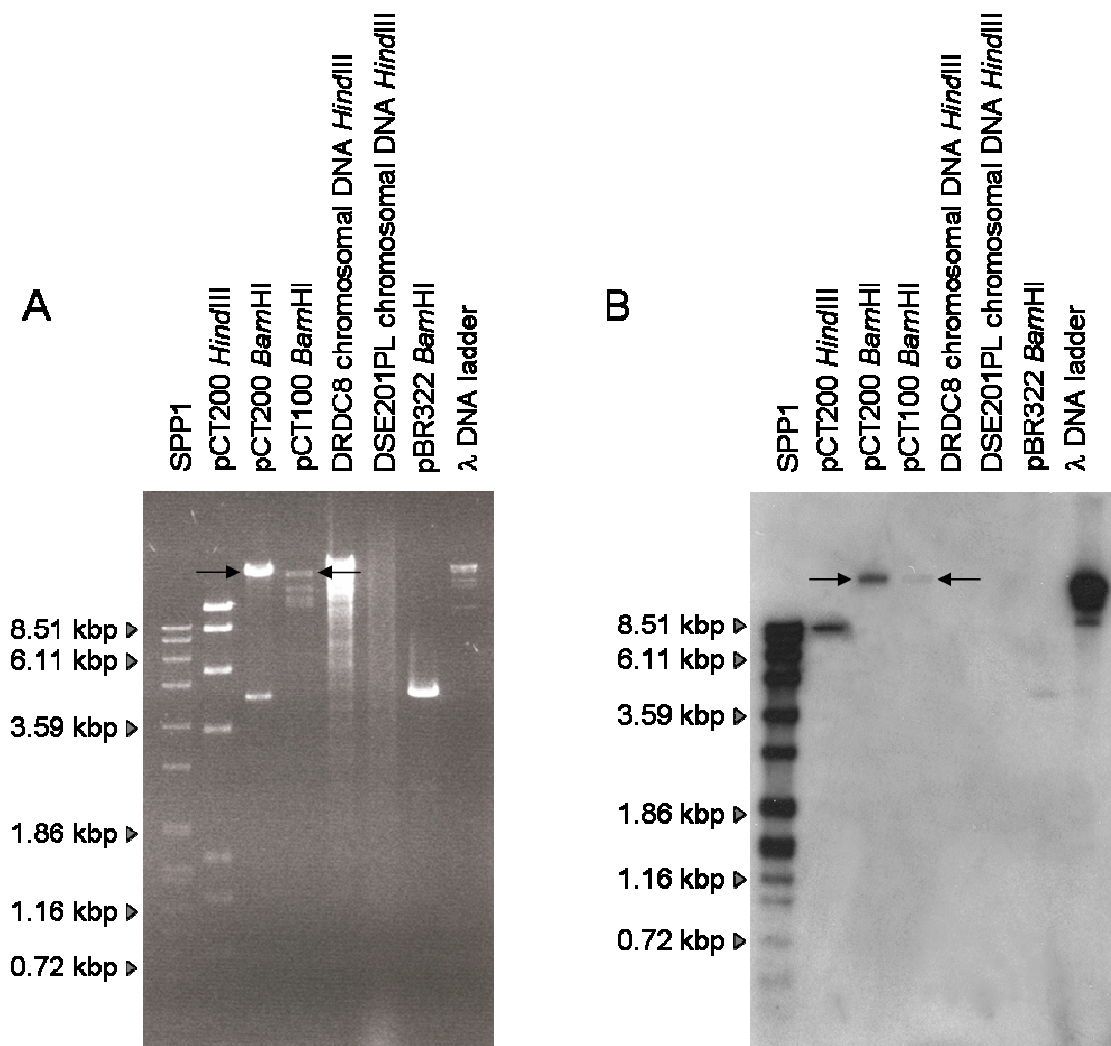
Figure 3.2: Gel electrophoresis and Southern hybridisation analysis of pCT100

Panel A. Agarose gel electrophoresis analysis of *Bam*HI digested plasmid pCT100. The left arrow indicates a 28.5 kbp *Bam*HI fragment of plasmid pCT200 that encodes *ctpA*. *Bam*HI digestion of plasmid pCT100 yields a similarly sized DNA fragment, as indicated by the right arrow.

Panel B. Southern blot hybridisation analysis of Panel A. A labelled 396 bp *ctpA*-specific DNA probe hybridised to a *ca*. 8.497 kbp *Hind*III fragment and a 28.5 kbp *Bam*HI fragment (indicated by the left arrow) derived from the positive control plasmid pCT200. Probe DNA also hybridised to a *Bam*HI fragment of plasmid pCT100 *ca*. 28.5 kbp in size (indicated by right arrow). Probe DNA did not hybridise to the negative control strain DSE201PL (refer to Table 2.1) DNA, but minor non-specific probe hybridisation to the pBR322 negative control was detected.
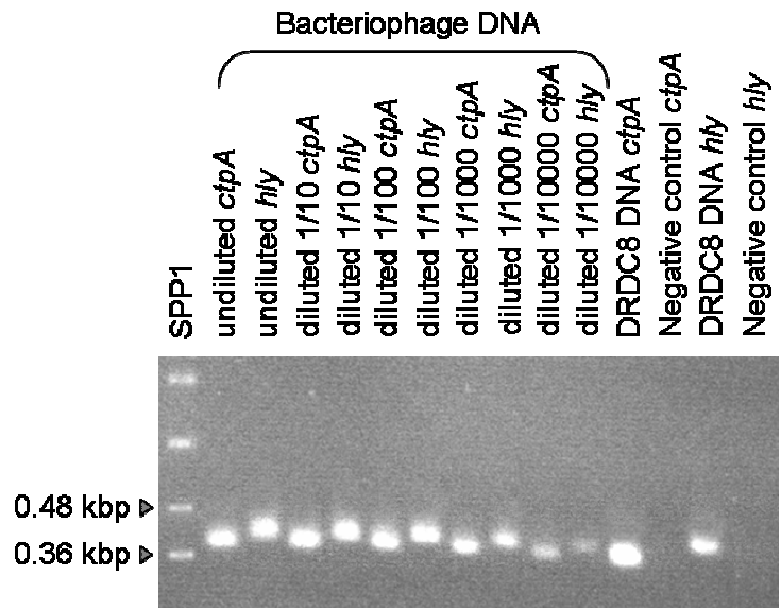
Figure 3.3: PCR analysis of bacteriophage DNA extracts.

The FB001/LM2004 and p234/p319 oligonucleotide pairs were used to amplify *ctpA* (396 bp) and *hly* fragments (417 bp), respectively. Amplification reactions were carried out on serial dilutions of bacteriophage DNA extracted from bacteriophage isolated from cultures of *L. monocytogenes* strain DRDC8. DNA amplicons representing the *ctpA* (396 bp) and *hly* (417 bp) fragments were amplified from undiluted samples of bacteriophage DNA and in proportionately diminishing amounts for successive dilutions of bacteriophage DNA. 396 bp (*ctpA*) and 417 bp (*hly*) fragments were amplified from positive control DRDC8 DNA. No amplicons were produced for the no DNA negative controls.
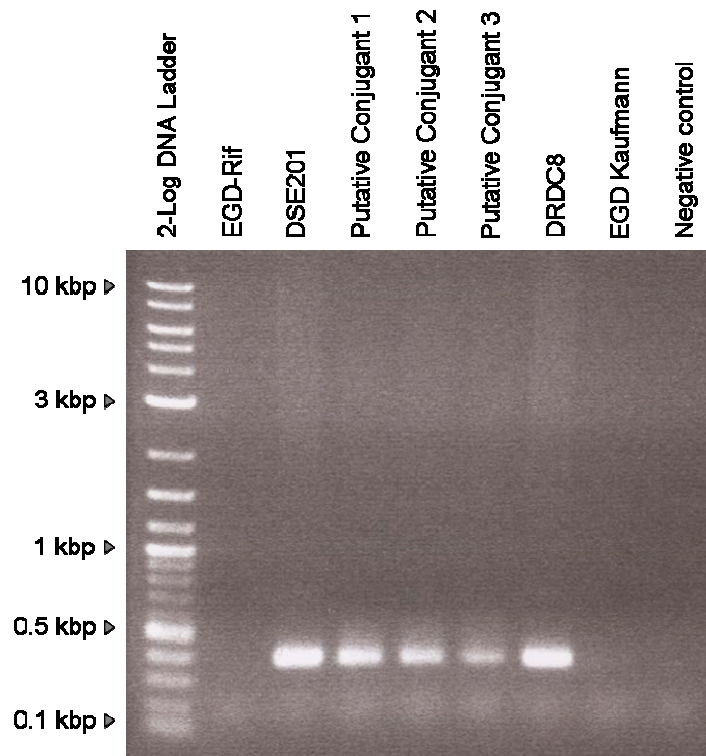
Figure 3.4: PCR amplification of *ctpA* from putative transconjugants DNA.

A 396 bp DNA fragment of *ctpA* was amplified using the FB001/LM2004 oligonucleotide pair from DNA extracted from putative conjugant 1, 2 and 3. DNA from DSE201 and DRDC8 was used for positive control amplifications of *ctpA*. PCR products were not amplified from DNA extracted from negative control strains EGD-Rif and EGD Kaufmann or the no DNA negative control.

```
a   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSAVLSLPLLLTMVTHIPYIHEMV
b   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSALLSLPLLLTMVTHIPYIHEMA
c   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSALLSLPLLLTMVTHIPYIHEMA
d   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSAVLSLPLLLTMVTHIPYIHEMV
e   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSAVLSLPLLLTMVTHIPYIHEMA
f   IKVVKHAGYDAAEKMSKEEKDAVLEKNFKKEVRRFILSAVLSLPLLLTMVTHIPYIHEMA
g   -------------MSIKN-------------RFIIGVIGSVPLLINMFMSLG--GSML
                 ** ::            ***:..: *:***:.*.  :    .*


a   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
b   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
c   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
d   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
e   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
f   FAETIGNWINPTIQLVLATIVQFYIGWRFYDGAYKALRGKSANMDVLVALGTSAAYFYSV
g   GGDKYGVWI----LFAFGSLVYWFSGLPFLRTAVASFKNHHANMDTLVGLGTTIAYVYSL
     .:. * **       :.:.::* :: *   *    * :::.: ****.**.***: **.**:


a   VEYIRHMIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
b   VEYVRHIIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
c   VEYVRHIIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
d   VEYIRHMIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
e   VEYIRHMIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
f   VEYIRHMIDPSVMPHYYFETSAVLITLILLGKLLESYATSRTTESIAGLLELQAKEATVI
g   YAMFARPNET------YFEAVAVVITLILLGSYFEERMKASASSAVDKLMGLQAKDAEVL
     . :  :.       ***: **:*******. :*.  .: ::.::  *: ****:* *:
```

Figure 3.5: Multiple amino acid sequence alignment.

Alignment of the translated polypeptide sequence of 601 bp PCR product amplified from *L. monocytogenes* strain DRDC8 using the CuTrF/CuTrR oligonucleotide pair **(a)** with the appropriate region of amino acid sequence of copper-translocating P-type ATPases from five *Listeria* strains: **(b)**, *L. monocytogenes* EGD (Accession CAC99931); **(c)**, *L. monocytogenes* str. 4b F2365 (Accession AAT04651); **(d)**, *L. monocytogenes* 4b H7858 (Accession EAL09129); **(e)** *L. monocytogenes* 1/2a F6854 (Accession ZP_00234473); **(f)** *L. innocua* Clip11262 (Accession CAC97197), and the plasmid-encoded copper-transporting P-type ATPase CtpA **(g)**.

'*' indicates positions which have a single, fully conserved residue shared between all sequences, ':' indicates amino acids that are in a 'strong' similarity group are fully conserved and '.' indicates that the amino acids of a 'weaker' similarity group are fully conserved.

Figure 3.6: Restriction digestion and Southern hybridisation analysis of pCT200.

Plasmid pCT200 is a *ca.* 32.861 kbp pBR322 based vector containing a 28.5 kbp *Bam*HI fragment of plasmid pCT100. This fragment has a Tn*917* insertion at the 3' end of *ctpA*. The genetic map of plasmid pCT200 shows that it is comprised of eight *Hind*III fragments and two *Bam*HI fragments. Gel electrophoresis of *Hind*III digested pCT200 showed fragments *ca.* 11.627, 8.497, 5.473, 3.545, 1.613, 1.248, 0.858 and 0.355 kbp in size. *Bam*HI digestion showed fragments *ca.* 28.5 and 4.361 (pBR322) kbp in size. Southern blot hybridisation analysis of *Hind*III and *Bam*HI digested pCT200 using a 2.8 kbp *ctpA*-specific labelled probe shows hybridisation to three *Hind*III fragments (*ca.* 8.497, 5.473 and 0.858 kbp) and 28.5 kbp *Bam*HI fragment.

Abbreviations: Amp^R, ampicillin resistance cartridge; Tc^R, tetracycline resistance cartridge; *rop*, *rop* gene responsible for regulation of plasmid replication and stability; *ori*, pBR322 origin of replication; *lacZ*, β-galactosidase gene from Tn*917-lacZ-cat86*; LIR$_{Tn917}$, left inverted repeat region from Tn*917-lacZ-cat8*.

The agarose gel electrophoresis photo presented in this figure has been modified for clarity purposes.
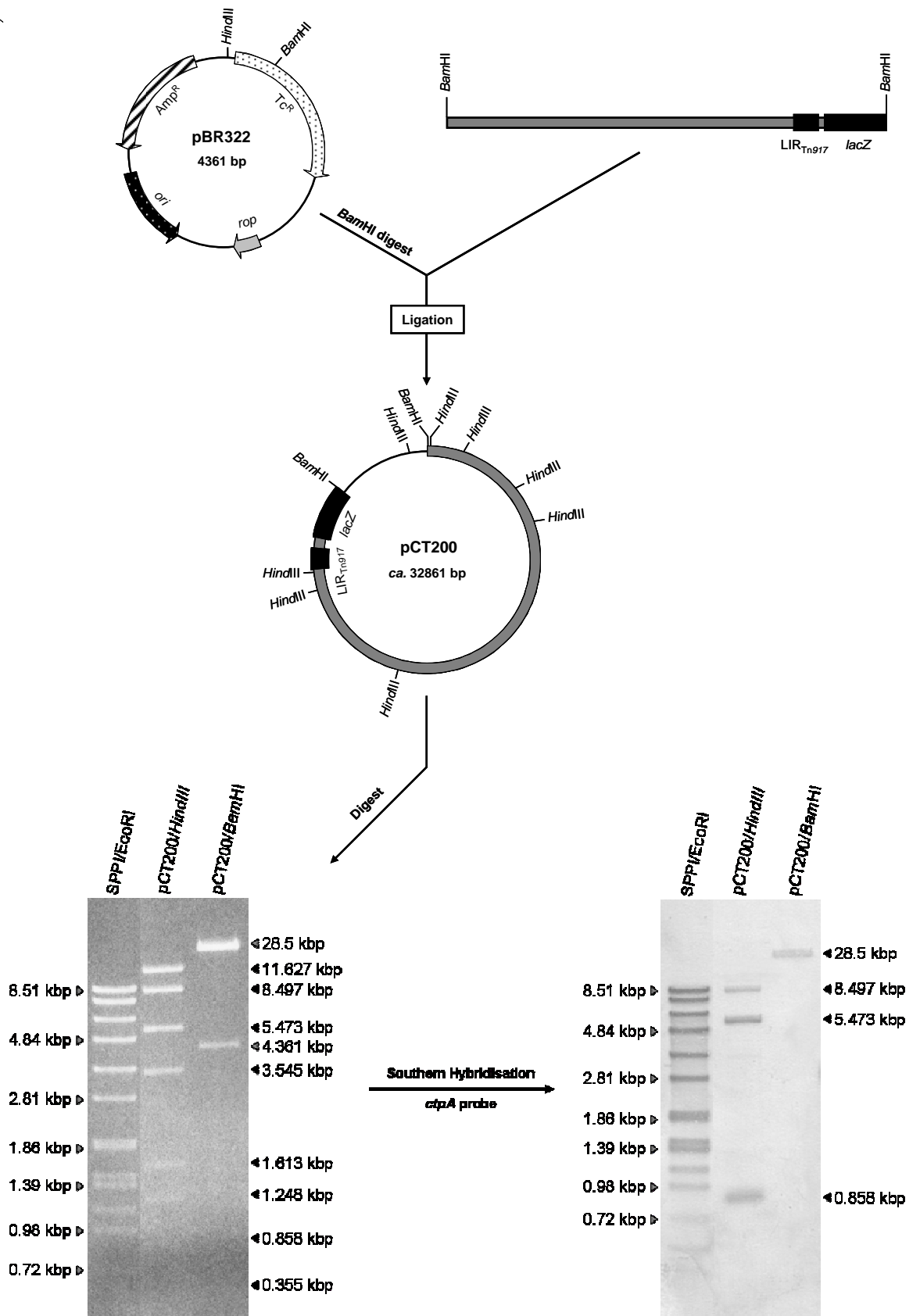
Figure 3.7: Construction of plasmids pFB186 and pFB190.

Following digestion of plasmid pCT200, two *Hind*III fragments, 5.473 kbp and 1.613 kbp in size, were isolated, then purified and subcloned into *Hind*III digested cloning vector pUC19 to create plasmids pFB186 and pFB190, respectively.

Abbreviations: Amp^R, ampicillin resistance cartridge; *ori*, origin of replication; *lacZα*: β-galactosidase gene.
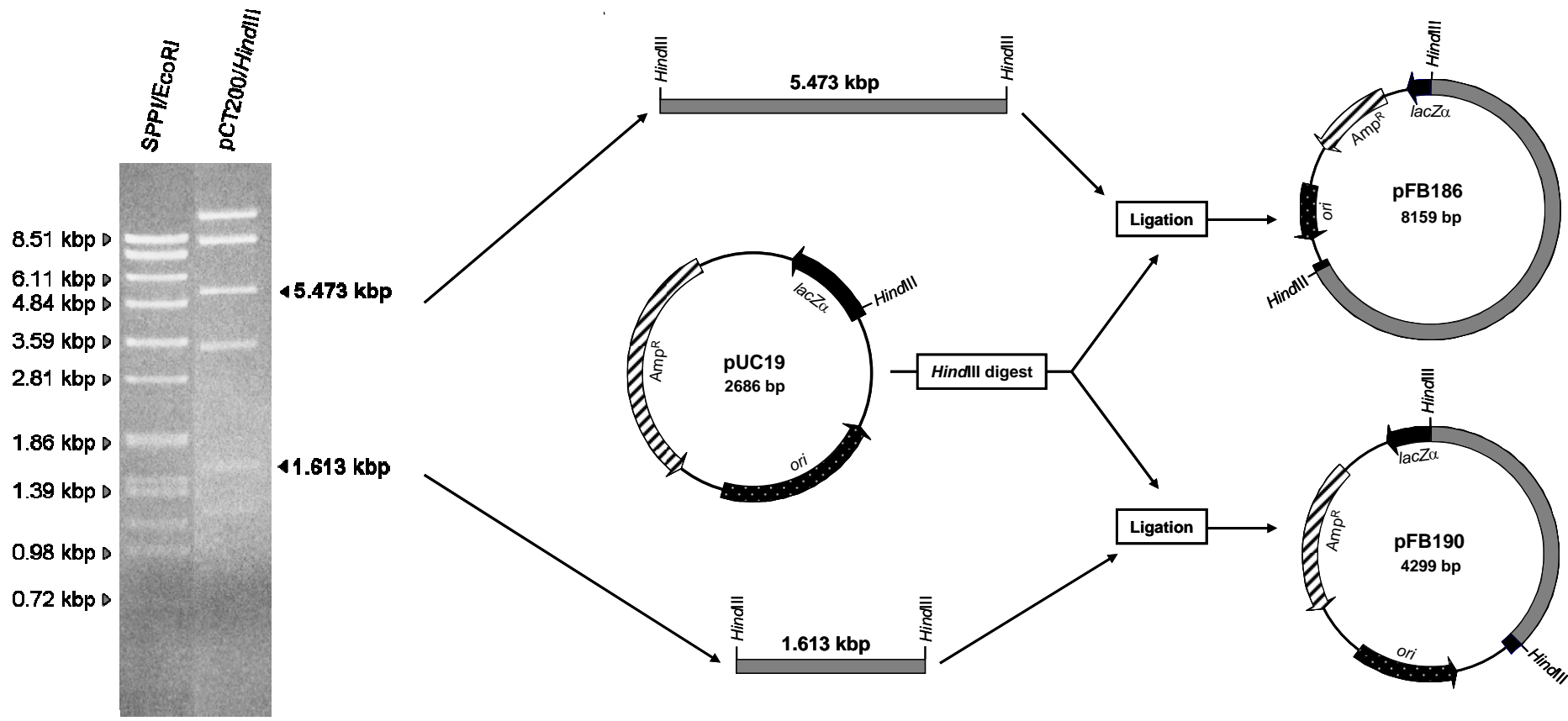
Figure 3.8: Genetic map of 37.279 kbp of plasmid pCT100.

A total of 37.279 kbp of DNA from plasmid pCT100 has been sequenced. From this data, 32 open reading frames have been identified and characterised. ORFs sequenced and characterised in this study are shown in colour, with their respective deduced putative functions indicated in the legend. ORFs shown in grey have been characterised in other studies. Regions of DNA sequenced from *Hind*III fragment 3 or 5 of plasmid pCT200 or from TempliPhi amplified DNA, are shown relative to their position on the map. The location of a putative plasmid replication region and a putative copper gene cluster is also shown.

TempliPhi Amplified Region

pCT0030
pCT0031
pCT0032
pCT0001
pCT0002
pCT0003
pCT0004
pCT0005
pCT0006
pCT0007
pCT0008
pCT0009
pCT0010
pCT0011
pCT0012
pCT0013
pCT0014
pCT0015
pCT0016
pCT0017
pCT0018
pCT0019
pCT0020
pCT0021
pCT0022
pCT0023
pCT0024
pCT0025
pCT0026
pCT0027
pCT0028
pCT0029

HindIII Fragment 5

Replication Region

pCT100

~ 80 kbp

Copper Gene Cluster

HindIII Fragment 3

TempliPhi Amplified Region

Plasmid DNA Replication Protein	Two-component Response Regulator
UV Damage Repair Protein	Multi-Copper Oxidase
Transposase	Truncated Copper Translocating P-type ATPase
Invertase/Recombinase	CopY Transcriptional Regulator
Cadmium Resistance Regulator	Integral Membrane Protein
Sodium Driven Multidrug Efflux Pump	Truncated Prolipoprotein Diacylglyceryl Transferase

Figure 3.9: Nucleotide and deduced amino acid sequence of ORF pCT0001.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 302). Relevant sequence features are identified and discussed in Section 3.3.6.1.

```
2401 -  AACAAAAGATACTTAAGGATTTATTGAAAAACAGATAGTTTTATCTAAAGATAATTAGAT

2461 -  AAAAAGATAGTTTTATCTAATTATCTTTTTATCCTTTATGCCCACCGACCAAATCAGATC
                                    *  G  K  H  G  G  V  L  D  S  R
                                    Stop
2521 -  GTTTCAATCCTGTAGCCCCGGCTACTAAGCTATTGGCATGCATCAATGCTTTAAATCCAT
         K  L  G  T  A  G  A  V  L  S  N  A  H  M  L  A  K  F  G  Y

2581 -  AACGATCACGAATCTTATCAATCGTACGCTCTAGCTGTTCTTGCTGGAGCGTTTTTGTAT
         R  D  R  I  K  D  I  T  R  E  L  Q  E  Q  Q  L  T  K  T  Y

2641 -  AGTCTTCAAATACGCTCAGCTGCATCGATGTTTTTCTTTTGATGCCACCACAACTAACCG
         D  E  F  V  S  L  Q  M  S  T  K  R  K  I  G  G  C  S  V  A

2701 -  CAACTGATCGCACCGCATCATTTTCCTGATACTTCCAAAACATTTCTAAGAAGTAGGGGA
         V  S  R  V  A  D  N  E  Q  Y  K  W  F  M  E  L  F  Y  P  V
                         IMS and Pol IV kappa regions
2761 -  CAAGCTCCTTGCTACTGTCGGTTGCCATTATCTTTTTTTGGTGGCTAAAACCGTTCCGAG
         L  E  K  S  S  D  T  A  M  I  K  K  Q  H  S  F  G  N  R  T

2821 -  TGGAATATCGGCTGTAGCCGGCGCTTAGATGAATGACGCTGGTATCAACATGATGATTGC
         S  Y  R  S  Y  G  A  S  L  H  I  V  S  T  D  V  H  H  N  R

2881 -  GTAGGCGCATCGCTACTTCTTCCACCATTTCTTGAATAATCACGGCGATCTGTTCCGGAT
         L  R  M  A  V  E  E  V  M  E  Q  I  I  V  A  I  Q  E  P  D

2941 -  CATGGTAATCCTTTTCTAGGATTTGGCTCTTTCCATACGACTTCTCAAGTGGCTTATACT
         H  Y  D  K  E  L  I  Q  S  K  G  Y  S  K  E  L  P  K  Y

3001 -  TTTCTGACAGACGGCTATAATCGATCCCATGCGCATGATAATACAATTGTTCCCCAATCA
         E  S  L  R  S  Y  D  I  G  H  A  H  Y  Y  L  Q  E  G  I  V

3061 -  CACCTAAACGTCTTTTTTAAGTAGGAGGGGAGCATCTGGCTCAATCCATGGATGCTATAGA
         G  L  R  R  K  L  Y  S  P  L  M  Q  S  L  G  H  I  S  Y  I

3129 -  TCCCCATTCGCTCCAGCTGAAAAGCAGTCCTTTTTTCCAATTCCCCAAAAATCAGTGATCG
         G  M  R  E  L  Q  F  A  T  R  K  G  I  G  W  F  D  T  I  P

3181 -  GTCGGATCTTCCACACTGTTTCCGGTACATCTTTATATCGCCATTCCGCAATACCAGATT
         R  I  K  W  V  T  E  P  V  D  K  Y  R  W  E  A  I  G  S  K

3241 -  TTTGATGTTTGGCTTCATTATCCAATGCTAATTTTTGCAAGAAGGGGATTGTCTCCAATAC
         Q  H  K  A  E  N  D  L  A  L  K  A  L  L  P  N  D  G  I  G
                         UmuC domain
3301 -  CAACCGTGACATACAGCTTTAACTCCTTTAAAATATCTGCTTGTATCTTTTTTGCGATGA
         V  T  V  Y  L  K  L  E  K  L  I  D  A  Q  I  K  K  A  I  V

3361 -  CTTCAGCGGATCCAAAAAGCTTTTTTACTGTGCGTGACATCCAAAAAGGCTTCGTCGATAG
         E  A  S  G  F  L  K  K  S  H  T  V  D  L  F  A  E  D  I  S

3421 -  AATAGACATGAAAATCTTCCCTGGATACATATCGAAGGAAGATTTTCTGGATGAGCTCAT
         Y  V  H  F  D  E  R  S  V  Y  R  L  F  I  K  Q  I  L  E  N
```

143

```
3481 -  TCACTTTTAAGTATAGTTGCATTCTAGGTGGAGCAATAATGATGCGAGGATCGTCCTCTG
         V  K  L  Y  L  Q  M  R  P  P  A  I  I  I  R  P  D  D  E  P

3541 -  GCACTTCAAATTTTCTAGATCCTGTCTTAATTCGGAAAACTTTTTTCATTTGAGGACTTG
         V  E  F  K  R  S  G  T  K  I  R  F  V  K  K  M  Q  P  S  A

3601 -  CTGCAAGCACCAATCCTCCAGCGCGATCTGCATTACTCATGACAACAAGAAAAGCCTGGA
         A  L  V  L  G  G  A  R  D  A  N  S  M  V  V  L  F  A  Q  L

3661 -  GTGGATCTAACCCACGCTTAACGCATTCCACACTGGCGAAAAACGACTTCACATCGACAC
         P  D  L  G  R  K  V  C  E  V  S  A  F  F  S  K  V  D  V  C

3721 -  AAAGAATATCGCGAGAGGGTACGTGGGTGTAATCTTCTATCTGTACCATTTTTCTATTGT
         L  I  D  R  S  P  V  H  T  Y  D  E  I  Q  V  M
                                                      Start
3781 -  CCCTTCTGAAATATGATTAATATCGTCAAATGTCAACACGGTGATATCTCCGGCCATATC
           RBS

3841 -  TCTAATATAAATTCGATCATCTTCCACACCAGCTATTCTTCCCATGACAAAGGGGAGGGG
```

144

Figure 3.10: Nucleotide and deduced amino acid sequence of ORF pCT0005.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 303). Relevant sequence features are identified and discussed in Section 3.3.6.2.

```
6361 – ATATATTTATATTTCATGCATGCTCTACAAAATTCGGCAAATTTCTTTTATTCTTTAGTT

6421 – TCAATAAAGTTGAGTTCATTCATTTTCATACCTCCTCAAACTTGATACATTATTTTTAAC
                      *  E  N  E  Y  R  R  L  S  S  V  N  N  K  V
                   Stop
6481 – AGTACTACTATTCTCTTTTTGTCTCCCAACCAATTTTTCTATCACTTTATTAGTGGAGAC
         T  S  S  N  E  K  Q  R  G  V  L  K  E  I  V  K  N  T  S  V

6541 – GATTAGTTCCCTTTTTAACTCTCTTTTCAGTTTGCTCCTTTCTTTTTCTATTTCTCTAAT
         I  L  E  R  K  L  E  R  K  L  K  S  R  E  K  E  I  E  R  I

6601 – TCCGATAGCGTATTGTTCAAAAGATGTAGACACTGCGGTACAAAGGAATACATTTTTATA
         G  I  A  Y  Q  E  F  S  T  S  V  A  T  C  L  F  V  N  K  Y

6661 – AAAGACTCTAACATCACTAATATCTCTAGGATCATATCTAATCACCACAGGTTCACTAAC
         F  V  R  V  D  S  I  D  R  P  D  Y  R  I  V  V  P  E  S  V
                              Mu-transpos C domain
6721 – AAAAGCTGAAAGAGTTGGGTGGACATACTTCATCCCAAATAAGTGAATACCGTCAGAGTG
         F  A  S  L  T  P  H  V  Y  K  M  G  F  L  H  I  G  D  S  H

6781 – TACCACCCGTTCTTTCTTAACTTTTATTAAAAGAAGATCTAAATCCTCCAAAGAGTTTGG
         V  V  R  E  K  K  V  K  I  L  L  L  D  L  D  E  L  S  N  P

6841 – CATATTCGGTAGAAAATCATAGTTATTCCACATAGCTATAGGTTTTTCTTTAGTTGTACT
         M  N  P  L  F  D  Y  N  N  W  M  A  I  P  K  E  K  T  T  S

6901 – ATGCTGTTTTTGATGGTATATTCTTAACAACCAATTTCTGAAAATATCGGAAAACTCCTC
         H  Q  K  Q  H  Y  I  R  L  L  W  N  R  F  I  D  S  F  E  E

6961 – CAATGTTAGCATTGACTTTGCTTTCCGATTCTTAATATAACCTGGAAGATCTTGTAGTAA
         L  T  L  M  S  K  A  K  R  N  K  I  Y  G  P  L  D  Q  L  L

7021 – CATAGAATTTACAGTATGAAAAAACCGTTCGATTTTCCCTCTACCTCTAGGCACCCCAAC
         M  S  N  V  T  H  F  F  R  E  I  K  G  R  G  R  P  V  G  V

7081 – CTTAGAAAACTTTATTCTATTTTCAATCAGCAGAGACCTGTTCAAGATGTTGAGAAGTAA
         K  S  F  K  I  R  N  E  I  L  L  S  R  N  L  I  N  L  L  L
                              Rve integrase core domain
7141 – AATCGCTACCGTGATCTGTAAAAAAATTCTAGGGATACCACAAATAGGCCAGTCAGCCTG
         I  A  V  T  I  Q  L  F  I  R  P  I  G  C  I  P  W  D  A  Q

7201 – ATCTTTTTCTCCATATTGCTTGTCTTAAAGTTAGAGCTGTTCGTAGTGAGTTTGGTGATTC
         D  K  R  W  I  A  Q  R  L  T  L  A  T  R  L  S  N  P  S  E

7261 – AAACGCAATAAAGTACCCAGTAATTGCTCGACTGTAGTCATCCATTATTATTGTTAGCCA
         F  A  I  F  Y  G  T  I  A  R  S  Y  D  D  M  I  I  T  L  W

7321 – AGGCCTTTCAGGATAACCTTTCGCATTCAATACCTCAATGTCTAATAATGTATGATCAGC
         P  R  E  P  Y  G  K  A  N  L  V  E  I  D  L  L  T  H  D  A

7381 – TTGCCATATGATATTGGGAGCTGTAGTTTCCCTAATATGAACTAATTCGTAATGATTCAT
         Q  W  I  I  N  P  A  T  T  E  R  I  H  V  L  E  Y  H  N  M
```

145

```
7441 - ATATATTTTAGATCCCTGATGCGCTAGTACTCTCAATTTATCGGGAACTAATTGAATAAT
         Y  I  K  S  G  Q  H  A  L  V  R  L  K  D  P  V  L  Q  I  I

7501 - TTTCCTTACCTGATAGTAAGAAGGCTCCTGAATATTATTTTCTTTACACCAAGACACAGT
         K  R  V  Q  Y  Y  S  P  E  Q  I  N  N  E  K  C  W  S  V  T

7561 - TCTCCTATAAATGGAAGTTATAGAAACATTTTTATTTTCTAAGTAAATGTGTTTAATGAT
         R  R  Y  I  S  T  I  S  V  N  K  N  E  L  Y  I  H  K  I  I

7621 - TTTACAAAGGTCTTCAGAAATTTTCATTGAATTTTTATCCGACCTTTCTTTTGTAATTAA
         K  C  L  D  E  S  I  K  M  S  N  K  D  S  R  E  K  T  I  L

7681 - TCCCACCAATCCCGATTTTTTATAATTCGAAATCCACTTTCTAAGTGTACGTGTTGAAAT
         G  V  L  G  S  K  K  Y  N  S  I  W  K  R  L  T  R  T  S  I

7741 - ATCATTTTCTTCACTTACTTTGGAAAGTGGAATTTCTTTTTTTAAGTATGGCTGAAGTAC
         D  N  E  E  S  V  K  S  L  P  I  E  K  K  L  Y  P  Q  L  V

7801 - CCGAAATTTTTTTAATGCTTTTAATCTCTGCTCCTCTGAGTAATTAGCTAAAATTTTTTT
         R  F  K  K  L  A  K  L  R  Q  E  E  S  Y  N  A  L  I  K  K

7861 - ATTACTCATTTCCTCACCTCCTGATATGAACAAATTATTTCATATATCAGATTATATGTC
         N  S  M
              Start        RBS

7921 - TAATTAGGGTCTACCCTAATTAGACATATAAAACAGATTGAAAACAAACGATTTTTAGTG

8041 - TATGTTTAGGGTGTGTGATAAACAGTCAGAAAATAGGAAGTGGGAAAATGAAAATTGCTT
```

Figure 3.11: Nucleotide and deduced amino acid sequence of ORF pCT0006.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 303). Relevant sequence features are identified and discussed in Section 3.3.6.3.

```
7921 -  TAATTAGGGTCTACCCTAATTAGACATATAAAACAGATTGAAAACAAACGATTTTTAGTG

7981 -  TATGTTTAGGGTGTGTGATAAACAGTCAGAAAATAGGAAGTGGGAAAATGAAAATTGCTT
                                                        M   K   I   A   Y
                                          RBS           Start
8041 -  ATGCTCGTGTTTCAACTTTAGAACAAAATTTAACACGACAAATCGAAACACTCAAAAGTT
         A   R   V   S   T   L   E   Q   N   L   T   R   Q   I   E   T   L   K   S   F
        Site-specific recombinase active site

8101 -  TTGGTGCTGAAAAAATTTTTACGGAACAAAAGTCAGGTGCTTCTATTAAAAATAGAAAAG
          G   A   E   K   I   F   T   E   Q   K   S   G   A   S   I   K   N   R   K   A

8161 -  CATTTCAAGAAGCTTTGCAATTTGTTCGAAGTGGTGACACTTTTATGGTAGAAGCGATTG
          F   Q   E   A   L   Q   F   V   R   S   G   D   T   F   M   V   E   A   I   D
                                              Site-specific recombinase signature domain

8221 -  ATCGATTGGGTCGTAATTATGATGAAATCATCCAAACCGTTTATTTTCTAAAAGAAAAAG
          R   L   G   R   N   Y   D   E   I   I   Q   T   V   Y   F   L   K   E   E

8281 -  AGGTCAAACTGGTTATTACGAGTCTGCCTATTATGGCTGAAGCAATAGGAAATCCATTAT
          V   K   L   V   I   T   S   L   P   I   M   A   E   A   I   G   N   P   L   L

8341 -  TGGATAAGTTTATTAAAGACTTGATTATCCAGATATTAGCCATGATTGCCGAGCAAGAAC
          D   K   F   I   K   D   L   I   I   Q   I   L   A   M   I   A   E   Q   E   R

8401 -  GAGCAGAGAGTAAGAGACGGCAGGCACAAGGCATTCACAGTGCCAAGCAAAAAGGAATTT
          A   E   S   K   R   R   Q   A   Q   G   I   H   S   A   K   Q   K   G   I   Y

8461 -  ACAAAGGAAGACCTATTCTATATTCGCCTACTTCAAAAGATCCACAGAAACGAGTTATCT
          K   G   R   P   I   L   Y   S   P   T   S   K   D   P   Q   K   R   V   I   Y

8521 -  ATCAGAGAGTCATTGACATGTTAAGTAAAGGAGAGTCTATTTCAAAAATAGCTGAAGAAA
          Q   R   V   I   D   M   L   S   K   G   E   S   I   S   K   I   A   E   E   S

8581 -  GTGGTATTACAAGACAAACCGTATATCGAATTAAAAATTCGAACTAATAAACTAAAAGGA
          G   I   T   R   Q   T   V   Y   R   I   K   N   S   N   *
                                                           Stop
8641 -  TCTCTTCAGCCAATAAAAAGTAGCCCTTTTGATCTAACAGCAAAAGGGCTACTTATATTA

8701 -  TTATACGTTATTATATTTAGATGAAAATTATCAATAAAAAGCGGTGAAATTTATTAATGG
```

147

Figure 3.12: Nucleotide and deduced amino acid sequence of ORF pCT0007.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 304). Relevant sequence features are identified and discussed in Section 3.3.6.4.

```
8821 -  TTCAAACATTCACTTGAATATAGATGGATGTTCTGATATAATACATTCAAACAAACATTT
                        -35                           -10
8881 -  GAATGTAAAGGAATGATCGTATGAATAATGAAATTTGTGAAATAACTTGTATTCATGAAG
                                   M  N  N  E  I  C  E  I  T  C  I  H  E  D
                 RBS              Start
8941 -  ACAAAGTTAATCGTGCTAAATCTAAACTAGCAAACTTTGATACTCCGTCAGTAAGTGGTT
           K  V  N  R  A  K  S  K  L  A  N  F  D  T  P  S  V  S  G  F
                                      ArsR-type HTH domain
9001 -  TCTTTAAAATATTGTCAGATGAAAATCGACTTAAAATTGTACATGCTTTAGTTCATGAAG
              F  K  I  L  S  D  E  N  R  L  K  I  V  H  A  L  V  H  E  D
9061 -  ATGAGCTATGTGTTTGTGATATTGCTAATATCATTGATGCTTCGGTCGCAACAACTTCAC
              E  L  C  V  C  D  I  A  N  I  I  D  A  S  V  A  T  T  S  H
                                  ArsR family signature region
9121 -  ACCATCTGAACTCCTTAAAAAAATTGGGAGTAGTTGATAGTCACAAGGATGGTAAGTTAG
              H  L  N  S  L  K  K  L  G  V  V  D  S  H  K  D  G  K  L  V
9181 -  TCTACTATTTTATTAAAAATATCAAAATCTTAAATCTTATGGAGCTTGGCGTAAATTTTA
              Y  Y  F  I  K  N  I  K  I  L  N  L  M  E  L  G  V  N  F  K
9241 -  AAGAAGAGGTGCTAGCATGAGTGAGAAAATGACCGAAAAAACCTATCGAATAGAAGGATT
              E  E  V  L  A  *
                              Stop
9301 -  AAGCTGTACAAATTGTGCAGGTAAATTTGAAAAAAATGTAAAGCAGTTGCCAGGAGTTAC
```

148

Figure 3.13: Nucleotide and deduced amino acid sequence of ORF pCT0015.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 309). Relevant sequence features are identified and discussed in Section 3.3.6.5.

```
21901 - TGACGCAAAAGAATAAAAGATTAAAAGGTTCTGTTGCAAAGTTCAAAGAAGTAAAATGAT
                                   Inverted Repeat Unit
21961 - ATCTCAACTATTTTCTAATTTTAAGCTGGAACACCCATTAGATTATTTATTTCATGAATC
                                *  A  P  V  G  M  L  N  N  I  E  H  I  V
                               Stop
22021 - ACTGAAAAGCCAAAGAGAGAAAAGTCTCTTTGGCTAGTTTTGTATAAAGCATGTATAGTT
           S  F  G  F  L  S  F  D  R  Q  S  T  K  Y  L  A  H  I  T  E

22081 - TCAATTCCTTTGATTGTGGATGAGGCAGTCCTTATGGACTGATATAATTTGTGCCTTTTC
            I  G  K  I  T  S  S  A  T  R  I  S  Q  Y  L  K  H  R  K  K

22141 - TTGATAGGCCTATGATCTTGTTCAATGATATTATTCAAATATTTACTCGTTCGATGTTCA
            I  P  R  H  D  Q  E  I  I  N  N  L  Y  K  S  T  R  H  E  T

22201 - GTGTTAGAAAATAAACCTTCAGATTTAAGTCGATTGAACGCACTGCTTAGCGAAGGAGCT
            N  S  F  L  G  E  S  K  L  R  N  F  A  S  S  L  S  P  A  K
                           Rve Integrase core domain
22261 - TTGTCCGTTACAAAAACTCTAGGTTCTCCAAATTGCTTGATTAATCGTTTAAAGAAAGCA
            D  T  V  F  V  R  P  E  G  F  Q  K  I  L  R  K  F  F  A  Y

22321 - TAGGCCGCCTGTTTATCCCGATTCTTTCGCAACCAAATATCTAATGTCAAACCTGTAGAA
            A  A  Q  K  D  R  N  K  R  L  W  I  D  L  T  L  G  T  S  D

22381 - TCGATGGCACGATACAAATAATGCCATTTTCCTTTTACTTTAATATATGTTTCATCCATT
            I  A  R  Y  L  Y  H  W  K  G  K  V  K  I  Y  T  E  D  M  R

22441 - CTCCAAGAATCGCTAGCAGACTTATTTCTTTTCTTCCATAGGCAATAAATAATCGAACCA
            W  S  D  S  A  S  K  N  R  K  K  W  L  C  Y  I  I  S  G  Y

22501 - TATTCTTGTACCCATCGATAAACAGTTGTGTGACAAACCTTGATTCCACGCTCATTAAGG
            E  Q  V  W  R  Y  V  T  T  H  C  V  K  I  G  R  E  N  L  I

22561 - ATTTCTGAAACATCACGATAGCTCAAATTGTATCTTAAGTAGTAGCCCACGGCTACCATG
            E  S  V  D  R  Y  S  L  N  Y  R  L  Y  Y  G  V  A  V  M  I

22621 - ATGATATCCTGTTTGAATTGTTTACCTTTAAAATGGGACATTGTTATTCTCCTCGTGTAG
            I  D  Q  K  F  Q  K  G  K  F  H  S  M
                                              Start       RBS
22681 - TTTCTTCCATTCTACACTAAATAATCAGTTAAATAAAACTTTGCAACAGAACCCATAATT
                                           Inverted Repeat Unit
22741 - AACCTTATATGAATACCAAAAAAATAATATATCTTTCTCTATTCAACTTATCATTTACAT
```

149

Figure 3.14: Nucleotide and deduced amino acid sequence of ORF pCT0016.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 309). Relevant sequence features are identified and discussed in Section 3.3.6.6.

```
23101 -  GTGATGGTTCTGATTTAATGAGAGAGATAGAAATGCGAAAACGTGTAATCAATATAATAA

23161 -  ATTGAATTGAGTGGTTGTATATGAAGACAATGAAAATCGAATCTAACGATGACCAGAATG
                                                             M   T   R   M
                                                                 Start
23221 -  TTAAAAAACCTGATTAAGTTAGCACTTCCTGCTACGGTCGAAAATATATTAGAAACATCT
          L   K   N   L   I   K   L   A   L   P   A   T   V   E   N   I   L   E   T   S
           NorM domain
23281 -  GTAGGATTTGTTGATTCGTTGATGATTTCTCAGATTGGATTATTTGCTGTAGCAGGGATT
          V   G   F   V   D   S   L   M   I   S   Q   I   G   L   F   A   V   A   G   I

23341 -  GGGATAGCAAATGCAATACTAAATGTTTATATTGCTGTATTTATTGCTTTAGGAATCGGA
          G   I   A   N   A   I   L   N   V   Y   I   A   V   F   I   A   L   G   I   G

23401 -  ACATCTTCTCTGATTTCTAGAAATATTGGAGCAAAAAATAATGAAAAGGCGAAGTTATTG
          T   S   S   L   I   S   R   N   I   G   A   K   N   N   E   K   A   K   L   L

23461 -  CTAACAATCTTTACTATTAGCAATCGTCTCGGGTTATTCTGGGTTGGTGAGTCTTCTAAT
          L   T   I   F   T   I   S   N   R   L   G   L   F   W   V   G   E   S   S   N
                                  MatE domain
23521 -  AGGTCCACAACTCTTAAACTTATGGGTGCAACAAAACAAACACTTAACTATTCCTTACAA
          R   S   T   T   L   K   L   M   G   A   T   K   Q   T   L   N   Y   S   L   Q

23581 -  TTTTTTAGTATAGTAGGTGGCGGATCGGTGGCTATTGCCACAATGATTATATTAGGAAGT
          F   F   S   I   V   G   G   G   S   V   A   I   A   T   M   I   I   L   G   S

23641 -  ATGTTAAGGGCGATAGGTGATACGAAAACTCCTATGAAAATTGGACTTATCACGAATATT
          M   L   R   A   I   G   D   T   K   T   P   M   K   I   G   L   I   T   N   I

23701 -  CTAAATATTTCCCTAGATTATATTCTGATTTTTGGTTTAGGACCTATACCAGCTTTGGGT
          L   N   I   S   L   D   Y   I   L   I   F   G   L   G   P   I   P   A   L   G

23761 -  GTAGTAGGAACAGCAATAGGAACGTTAGTTGCGAGATTACTAGGTACTATTTTGTTATAT
          V   V   G   T   A   I   G   T   L   V   A   R   L   L   G   T   I   L   L   Y

23821 -  CGAAAAGTACATTTATCTGTTCTTTCTTTCCCTTTACTCTCAATCTTTCAACGAAGTAAT
          R   K   V   H   L   S   V   L   S   F   P   L   L   S   I   F   Q   R   S   N

23881 -  TATCAAGAATTATTAAAACTATCGTTTCCTGCTGCATTAGAACGTTTAGTGATGCGGTTA
          Y   Q   E   L   L   K   L   S   F   P   A   A   L   E   R   L   V   M   R   L

23941 -  GGGCAAGTTTTGTATTTTGGAGTGATCGTAGCAATTGGCGCAAAAACTTATTCAGCTCAC
          G   Q   V   L   Y   F   G   V   I   V   A   I   G   A   K   T   Y   S   A   H
                                                  MatE domain
24001 -  TCCATCGCAGGCAGCATTGAAAGTTTTGTCTATATGCCAGCCTATGGGTTAGCTACTGCT
          S   I   A   G   S   I   E   S   F   V   Y   M   P   A   Y   G   L   A   T   A

24061 -  GCAGCAACTTTAGCAGGAAATAGTATTGGAAAAAAAGACTATGCTGAAACAAAACGTGTT
          A   A   T   L   A   G   N   S   I   G   K   K   D   Y   A   E   T   K   R   V

24121 -  GCACTTTATGCGGTTAAATACGGTGTTATTGTTTTAAGTATTCTAGGTATAGGATTGTTT
          A   L   Y   A   V   K   Y   G   V   I   V   L   S   I   L   G   I   G   L   F
```

150

```
24181 - TTTGGTACGCCATATATTGCACCTTGGTTTACCATAGATGCTGAGGCAATCAATCAAATT
         F  G  T  P  Y  I  A  P  W  F  T  I  D  A  E  A  I  N  Q  I

24241 - GTCATTGCACTCAAAATTGACGCTTTTAATCAACCAGGATTAGCTATCAGTCTAATATTA
         V  I  A  L  K  I  D  A  F  N  Q  P  G  L  A  I  S  L  I  L

24301 - GCAGGAGTACTCCAAGGAATGGGAGACACTAAAACCCCATTATATAGCACCGCATTTGGG
         A  G  V  L  Q  G  M  G  D  T  K  T  P  L  Y  S  T  A  F  G

24361 - ATGTGGATCACACGAATACTAGGGTGTATACTTTTAGGGAGTATAATGGGTTTAGGGATT
         M  W  I  T  R  I  L  G  C  I  L  L  G  S  I  M  G  L  G  I

24421 - GCTGGAGTGTGGTTGGCAATCGGCTTGGATCTATATGTGCGTTCTCTTTTTTTAACCTAT
         A  G  V  W  L  A  I  G  L  D  L  Y  V  R  S  L  F  L  T  Y

24481 - CATTTCAAACGAAATCTTCAAATTTTAAACAGTAATAACGAACTTTCTTCTTAAGCAAGT
         H  F  K  R  N  L  Q  I  L  N  S  N  N  E  L  S  S  *
                                                           Stop
24541 - GATTTTTATCACTTGTTTTTTTCGTTTAACGATAAATTTCTTTGCCATTAGTAGGTAGA
```

Figure 3.15: Nucleotide and deduced amino acid sequence of ORF pCT0024.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 313). Relevant sequence features are identified and discussed in Section 3.3.6.7.

```
32401 - TTCGTCAAACTTAATGCGACTAAACTAATCGATCCTATTATTAATGTTGCGATTGCTAAA

32461 - AAGGAAACAGCCTAATTGAGTTTTTATCGTTCTTTTTCTCATTAGTGGCCACCAAATCTG
                                                 *  H  G  G  F  R  Y
                                                Stop
32521 - TATCCTGTACCATGGACTGTCAAAATATACATAGGTGATTTAGCATCATTCTCAATTTTA
          G  T  G  H  V  T  L  I  Y  M  P  S  K  A  D  N  E  I  K  H

32581 - TGTCTTAAGTTTTTTTATATGGGAATCGATGCTCCTATCTAGTCCATCAAATTCTATTCCC
          R  L  N  K  I  H  S  D  I  S  R  D  L  G  D  F  E  I  G  K
                            trans_reg_C domain
32641 - TTAACTTTCTCTATCAAATCACTACGAGAAAAGACTTTATTAGGTGTGGACATTAGAGCG
          V  K  E  I  L  D  S  R  S  F  V  K  N  P  T  S  M  L  A  H

32701 - TGTAGAATATCAAATTCAGTTGTTGTAAGGTTTAACTCCACTTGATGTAGATACACTTGT
          L  I  D  F  E  T  T  T  L  N  L  E  V  Q  H  L  Y  V  Q  R

32761 - CTACTATCAGGGTATACTATCAATTCTTTATTATTTGAAGAAAGCTTATTAGGATTTGAT
          S  D  P  Y  V  I  L  E  K  N  N  S  S  L  K  N  P  N  S  L

32821 - AAGGTCTCAGTTCTGCGCAATACAGTTTGTATTCTCGCAACTAATTCTTTGGGACTAAAT
          T  E  T  R  R  L  V  T  Q  I  R  A  V  L  E  K  P  S  F  P

32881 - GGTTTAAGCATATAGTCATCTGCACCTAACTTGAGACCAGTAAGAATATCATCCTCATCA
          K  L  M  Y  D  D  A  G  L  K  L  G  T  L  I  D  D  E  D  S

32941 - GACTTAGCTGTTAACATTATAATTGGTACATTTGATATTTCTCTGATTTTTCTACATACA
          K  A  T  L  M  I  I  P  V  N  S  I  E  R  I  K  R  C  V  T
                    N-terminal response regulatory domain
33001 - GTCAATCCATCAGTATCCGGAAGCATTAGATCTAGCACAACTAAATCAGGATTATTTTTC
          L  G  D  T  D  P  L  M  L  D  L  V  V  L  D  P  N  N  K  E
        Conserved phosphorylatable aspartate residue
33061 - TCGAATTTTTCCAATGCCATTGCGCCACTTATTGCTTTATAAACTGAATAATGATTAGCA
          F  K  E  L  A  M  A  G  S  I  A  K  Y  V  S  Y  H  N  A  V

33121 - ACTAGATAGGCATCTATAATTTCTAATATTTTTTCTTCATCGTCCACAATAAGAATTTTC
          L  Y  A  D  I  I  E  L  I  K  E  E  D  D  V  I  L  I  K  M

33181 - ATAGACAATCATCCTTCTTAAATTATTTCCTATTTAATTTCCATTTGTCCCATCATTCCA
          Start          RBS
33241 - TACTCTTCATGTTCTAATATATGACAATGATACATGAAAATACCTTCACGGTCGAATTTC
```

152

Figure 3.16: Nucleotide and deduced amino acid sequence of ORF pCT0025.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 313). Relevant sequence features are identified and discussed in Section 3.3.6.8.

```
33121 - ACTAGATAGGCATCTATAATTTCTAATATTTTTTCTTCATCGTCCACAATAAGAATTTTC

33181 - ATAGACAATCATCCTTCTTAAATTATTTCCTATTTAATTTCCATTTGTCCCATCATTCCA
                                      *  K  I  E  M  Q  G  M  M  G  Y
                                      Stop
33241 - TACTCTTCATGTTCTAATATATGACAATGATACATGAAAATACCTTCACGGTCGAATTTC
         E  E  H  E  L  I  H  C  H  Y  M  F  I  G  E  R  D  F  K  V
        Multicopper oxidases signature 2
33301 - ACAAGAAGTTCTACTGTTTCATCCGGATTTACCAGCACAGTATCTTTCCATCCTTGCTCG
         L  L  E  V  T  E  D  P  N  V  L  V  T  D  K  W  G  Q  E  N

33361 - TTTAGTGCGGGCTGATTTCCATCTCTAGATAAAATTTGAAACTGAACACCATGGATATGG
         L  A  P  Q  N  G  D  R  S  L  I  Q  F  Q  V  G  H  I  H  F
                            Cu-oxidase_2 domain
33421 - AACGGATGAATCATGCCTCCCATCATACTACTAATATTGTTAACTTCCCAAATCTCTTGA
         P  H  I  M  G  G  M  M  S  S  I  N  N  V  E  W  I  E  Q  T

33481 - GTACCTAATTTTTTATATAAATCTATTCTTTCCATATCAAATTGCTTATTGTTGATGTTT
         G  L  K  K  Y  L  D  I  R  E  M  D  F  Q  K  N  N  I  N  V

33541 - ACCATGTGAGACATTCCACTTAAATTAATACTCTGTCTTGTTAAATCTTCCAATTTCTTT
         M  H  S  M  G  S  L  N  I  S  Q  R  T  L  D  E  L  K  K  E

33601 - TCATCTAAAGTACTAATAGTATTTAGACTATCAGTGGGATTAAACTCTTTGTTGTCTATC
         D  L  T  S  I  T  N  L  S  D  T  P  N  F  E  K  N  D  I  T

33661 - GTATTCTCAATTCTCATTGTCAGTGCAACCAAATTATTTGCTAGTAAATGAATCACCTTA
         N  E  I  R  M  T  L  A  V  L  N  N  A  L  L  H  I  V  K  G

33721 - CCTTTTTTATAATTCTGAGTATCCACTAATATTTCTGCCCTTTCCCCGGGCGCTAAAAGT
         K  K  Y  N  Q  T  D  V  L  I  E  A  R  E  G  P  A  L  L  L

33781 - AGTTTACTAAGTTTAACAGAAGTATTCAGAAAACCACCATCTGTAGCAATTTGGTAGAAT
         K  S  L  K  V  S  T  N  L  F  G  G  D  T  A  I  Q  Y  F  S
                         PRK10965 multicopper oxidase domain
33841 - GATTCGTCATTATCCAAATTAAAAGTGAAGTTTCTTGCATTGGATCCATTAACAATTCTA
         E  D  N  D  L  N  F  T  F  N  R  A  N  S  G  N  V  I  R  Y

33901 - TACCTCATCCATCGGCTTTTTATCTCAACGTATGGATTGATTGTGCCATTAGTAAGCAAC
         R  M  W  R  S  K  I  E  V  Y  P  N  I  T  G  N  T  L  L  T

33961 - GTCTCCCCTTTTGTTCCATCAGAGTTAAAATCATTTTCATAGTTAATTTGATTAGTAGAA
         E  G  K  T  G  D  S  N  F  D  N  E  Y  N  I  Q  N  T  S  S

34021 - CTGAAAGATTTATCTTGCACAATCAAGGGAATGTCGTCGACACCATATTTGGAAGGTAAA
         F  S  K  D  Q  V  I  L  P  I  D  D  V  G  Y  K  S  P  L  D

34081 - TCTAATAATTTTGAGTTGCCGTCATCAATATACATCAAACCTGCTAACCCCTTATAAACT
         L  L  K  S  N  G  D  D  I  Y  M  L  G  A  L  G  K  Y  V  Q

34141 - TGAGATGCTGTTTCTCCTTCCGGATGGGGATGAAACCAGAGTGTAGAAGCTTCTTGATCA
         S  A  T  E  G  E  P  H  P  H  F  W  L  T  S  A  E  Q  D  V
```

153

```
34201 - ACCTCGAATGTTACACTTTTTTTCTGTCCTGCTTCTATAATCTGATGTGGACCTCCATCT
                E  F  T  V  S  K  K  Q  G  A  E  I  I  Q  H  P  G  G  D  A
                                   Cu-oxidase_3 domain
34261 - GCATCAGAAGCTACTTTCAAACCGTGCCAATGAAAAGATGTGCTAGCATCAAGATTATTG
                D  S  A  V  K  L  G  H  W  H  F  S  T  S  A  D  L  N  N  T

34321 - GTTGTATTGATAGTAACTTTTTTGACCTTTTTTCAGTCTAATAACTGGCCCTAAAAAATCC
                T  N  I  T  V  K  Q  G  K  K  L  R  I  V  P  G  L  F  D  G

34381 - CCGTTATAGCCTAATGTTTCAGTTTTTTTCTCCATCCATGATTTGAACTTCCCCGTTTTGA
                N  Y  G  L  T  E  T  K  E  G  D  M  I  Q  V  E  G  N  Q  T

34441 - GTAACTATATCATATGTCACATTTTCTTTGCTTTTTGAAGATGGTTCAAGAACAGGGGGA
                V  I  D  Y  T  V  N  E  K  S  K  S  S  P  E  L  V  P  P  I

34501 - ATCAATAGTTTTCTATCTGTTTGGTGTACAACCTTCAAATCAGCTTCTTGATTATCCCCC
                L  L  K  R  D  T  Q  H  V  V  K  L  D  A  E  Q  N  D  G  M

34561 - ATCATATTGGACATTCTACTTGAGATTTGTCGTTGTTCTTCTTGATAGATAACTTCAGCA
                M  N  S  M  R  S  S  I  Q  R  Q  E  E  Q  Y  I  V  E  A  H

34621 - TGCTCCTCCTTGTTCTGCTGTTGAATCAATATGAAAATACCGCCTAGGGTAAATATTGAT
                E  E  K  N  Q  Q  Q  I  L  I  F  I  G  G  L  T  F  I  S  I

34681 - ATTGTTGTTGCTGCTAACCAAATATATTTATTATTCTTCAACCCATCATCTCCCATATCT
                T  T  A  A  L  W  I  Y  K  N  N  K  L  G  D  D  G  M  D  I

34741 - ATTATGTATTTACCATAAAAAAAATATGTGGAGATTTCGTGTCGATCCCCCCACATATAT
                I  Y  K  G  Y  F  F  Y  T  S  I  E  H  R  D  G  W  M
                                                                  Start
34801 - TTATATTTATATAGATATATCCCTAGTTACAGTATACTTTGTATTTTTTAGTAGTTGATT

34861 - TAAGAATGTTAATGTTAATTTTGAATCAGCATCGAAAGTAACTGTACCTCTAGCCAAGTC
```

Figure 3.17: Nucleotide and deduced amino acid sequence of ORF pCT0026.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 314). Relevant sequence features are identified and discussed in Section 3.3.6.9.

```
34741 - ATTATGTATTTACCATAAAAAAAATATGTGGAGATTTCGTGTCGATCCCCCCACATATAT

34801 - TTATATTTATATAGATATATCCCTAGTTACAGTATACTTTGTATTTTTTAGTAGTTGATT
                 *   I   S   I   D   R   T   V   T   Y   K   T   N   K   L   L   Q   N
              Stop
34861 - TAAGAATGTTAATGTTAATTTTGAATCAGCATCGAAAGTAACTGTACCTCTAGCCAAGTC
              L   F   T   L   T   L   K   S   D   A   D   F   T   V   T   G   R   A   L   D
                              HMA domain
34921 - AACTATTTTTACACTGACTGTTGGCTCTGCTTCCAAAGCATTTCTTACTTTTGTTGCACA
              V   I   K   V   S   V   T   P   E   A   E   L   A   N   R   V   K   T   A   C

34981 - CCCTGAGCAATTCATTCCACTTATTCCAAATACCTTTATCTCAGTTTTTTTTCGAATCTC
              G   S   C   N   M   G   S   I   G   F   V   K   I   E   T   K   K   R   I   E
              Copper binding motif
35041 - AGCGTCATAACCAATGTCTTGAATAGCTGTTATAACTGTATGATTAGACAAAACACCTTC
              A   D   Y   G   I   D   Q   I   A   T   I   V   T   H   N   S   L   V   G   E

35101 - ATCATATTCCAATTTTAGAACCTCAGTTGTTAAGTTAACTTTAGCGAAATATACTCCATC
              D   Y   E   L   K   L   V   E   T   T   L   N   V   K   A   F   Y   V   G   D
                              HMA domain
35161 - CAACTCATGAATTTTTTTCTCTATTGTTGTTGCGCAAACTGCACAAGTCATTCCATAAAG
              L   E   H   I   K   K   E   I   T   T   A   C   V   A   C   T   M   G   Y   L
                                                      Copper binding motif
35221 - ATTATATTTTGAAATCATTGAATTCCTCCGTGTCAGAACGTTTAACCAATCCGTTTAAAT
              N   Y   K   S   I   M
                              Start       RBS
35281 - AGTAGCTTACTTCTGAGATACAACCTTCATCCATTCCAGAAAGTAACCTTTGTATTTCGG

35341 - TTTCTAAAAATTCTTGCTTCTTAATCAGTGAAGTGTAGTATCTTTGGTTTTTTGAAGTGT
```

Figure 3.18: Nucleotide and deduced amino acid sequence of ORF pCT0027.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 314). Relevant sequence features are identified and discussed in Section 3.3.6.10.

```
35161 - CAACTCATGAATTTTTTTCTCTATTGTTGTTGCGCAAACTGCACAAGTCATTCCATAAAG

35221 - ATTATATTTTGAAATCATTGAATTCCTCCGTGTCAGAACGTTTAACCAATCCGTTTAAAT
          *   I   K   F   D   N   F   E   E   T   D   S   R   K   V   L   G   N   L   Y
         Stop
35281 - AGTAGCTTACTTCTGAGATACAACCTTCATCCATTCCAGAAAGTAACCTTTGTATTTCGG
          Y   S   V   E   S   I   C   G   E   D   M   G   S   L   L   R   Q   I   E   T

35341 - TTTCTAAAAATTCTTGCTTCTTAATCAGTGAAGTGTAGTATCTTTGGTTTTTTGAAGTGT
          E   L   F   E   Q   K   K   I   L   S   T   Y   Y   R   Q   N   K   S   T   N
                         Penicillinase repressor domain
35401 - TAAACGTGATCACATTTTTATCTAACAACCTTCTTATTAACGTTTTTGTTGTGGAGGGAT
          F   T   I   V   N   K   D   L   L   R   R   I   L   T   K   T   T   S   P   H
                             DNA binding motif
35461 - GCCAACTATTACGCTCTTGCATTCCATGTGTTATTAAACCACATGTTACCGGTTCATATC
          W   S   N   R   E   Q   M   G   H   T   I   L   G   C   T   V   P   E   Y   R

35521 - TCCAGATAAACTTTAAAACATCTAACTCTGAATTCGATACCTGTATTTTGTGTTCGACCA
          W   I   F   K   L   V   D   L   E   S   N   S   V   Q   I   K   H   E   V   L

35581 - AGACCATCCCTCCTTAGGCTATTCATAAAGTACGTTCTTCTTTATTAAACAAATGATGAA
          V   M
              Start    RBS
35641 - TTGTTGTCAAAGTAGAAACACTTATAGTTACAATTAATAAAAAGCATAAAAAGAAAAATG

35701 - TTTCCGGTAAAAAACTGATAATTGGATCTGTACGATAGTCAAACATCCAATCATCATTGT
```

Figure 3.19: Nucleotide and deduced amino acid sequence of ORF pCT0028.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 315). Relevant sequence features are identified and discussed in Section 3.3.6.11.

```
35521 - TCCAGATAAACTTTAAAACATCTAACTCTGAATTCGATACCTGTATTTTGTGTTCGACCA

35581 - AGACCATCCCTCCTTAGGCTATTCATAAAGTACGTTCTTCTTTATTAAACAAATGATGAA
                                  *  L  T  R  E  E  K  N  F  L  H  H  I
                                  Stop
35641 - TTGTTGTCAAAGTAGAAACACTTATAGTTACAATTAATAAAAAGCATAAAAAGAAAAATG
           T  T  L  T  S  V  S  I  T  V  I  L  L  F  C  L  F  F  F  T

35701 - TTTCCGGTAAAAAACTGATAATTGGATCTGTACGATAGTCAAACATCCAATCATCATTGT
           E  P  L  F  S  I  I  P  D  T  R  Y  D  F  M  W  D  D  N  D

35761 - CAAAAAATACTTGATGAAAAAGCACAAATGCTTTATCAAAAAACAGAAGTGCCATTACAC
           F  F  V  Q  H  F  L  V  F  A  K  D  F  F  L  L  A  M  V  C

35821 - AGACTGATATTATCATCCACTTGATTAAACTAGTGATTTCTCTGACTTGATCACGCGTTA
           V  S  I  I  M  W  K  I  L  S  T  I  E  R  V  Q  D  R  T  V
                                  DUF1461 domain
35881 - CTAATTTTTTCCTGTTTAGGTATAAAAGAATGCTACTCGTAGAAAGAAGCAATGAATAGT
           L  K  K  R  N  L  Y  L  L  I  S  S  T  S  L  L  L  S  Y  N

35941 - TAAGCTGGAACAAAAATTTAACATCTTTAAAATGGTTTATACCGCGATCTGAGACGTATA
           L  Q  F  L  F  K  V  D  K  F  H  N  I  G  R  D  S  V  Y  L

36001 - GATTATCAAAAACCAATTCTTTTTGAAAAGGATTATTTAAATAGCCCATTAAATTTCTAA
           N  D  F  V  L  E  K  Q  F  P  N  N  L  Y  G  M  L  N  R  F

36061 - AATCTTGATTAACTTGAGATCTACTCCAGTTTTGAAAATTTATATCGCTTTGAAAGAATA
           D  Q  N  V  Q  S  R  S  W  N  Q  F  N  I  D  S  Q  F  F  L

36121 - AATTGATTTTGAGCAAACCCGGGAAGTTGATTGTCAATGTAATCGCGAGGGATAAAGCAA
           N  I  K  L  L  G  P  F  N  I  T  L  T  **I  A** L  S  L  A  F
                                                signal peptide cleavage site
36181 - AGACAAAAAATAATAGAAATAGCATGCTTTTTTTTCATTAAATTAATATTTATATTAGATC
           V  F  F  L  L  F  L  M  S  K  K  M
                                          Start
36241 - ACTCCTTTTAATATAATGATAAAAAAACTTGTGTAGAAAATGTGTACCAATGTATCTGTC

36301 - AAAAAGTCTACATAAAATTTCCATAATAGCCTGCTATAGTCAAAACAAGTTCATTAGGAA
```

Figure 3.20: Nucleotide and deduced amino acid sequence of ORF pCT0029.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 315). Relevant sequence features are identified and discussed in Section 3.3.6.12.

```
36301 - AAAAAGTCTACATAAAATTTCCATAATAGCCTGCTATAGTCAAAACAAGTTCATTAGGAA
         K  K  S  T  *  N  F  H  N  S  L  L  *  S  K  Q  V  H  *  E

36361 - GGAGGGAAGAATTTGAGAGAACCCTACGACAGAGTATTTATTTCGTTTGGTCCATTTACT
         G  G  K  N  L  R  E  P  Y  D  R  V  F  I  S  F  G  P  F  T

36421 - ATTTATTGGTATGCCATTTTCATAGTGTTTGGGATCATAATCGGATATTTTGTTGCTAAT
         I  Y  W  Y  A  I  F  I  V  F  G  I  I  I  G  Y  F  V  A  N

36481 - AGAAGAGCAAAGCGTGCGGCATTGCCTGAGGATACAATTGTTGATCTGTTGTTATATGGA
         R  R  A  K  R  A  A  L  P  E  D  T  I  V  D  L  L  L  Y  G

36541 - TTGCCAGTTTCTATTATATCGGCCAGATTATACTATGTACTTTTTGAATTACCTTTGTAC
         L  P  V  S  I  I  S  A  R  L  Y  Y  V  L  F  E  L  P  L  Y

36601 - ATTGATGATCCTTTAAGTATCTTAAAAATTTGGGAAGGTGGTCTAGCAATACATGGTGGA
         I  D  D  P  L  S  I  L  K  I  W  E  G  G  L  A  I  H  G  G

36661 - TTAATTGGTGCTGTATTGACAGGGGGTTATTTATTGCAAGAAGAAAAATCTTTCTTTTTGG
         L  I  G  A  V  L  T  G  V  I  Y  C  K  K  K  N  L  S  F  W

36721 - AATGTTGCTGACGTAGTTGCTCCCAGTATAGTAATTGGACAAATCATTGGACGATGGGGA
         N  V  A  D  V  V  A  P  S  I  V  I  G  Q  I  I  G  R  W  G

36781 - AATTTTATGAATCAAGAGGCCTATGGAGAAATTGTAGACATAGAGTTCCTGCAATCTCTG
         N  F  M  N  Q  E  A  Y  G  E  I  V  D  I  E  F  L  Q  S  L
                  Start

36841 - AAACTCCCATCTTTCATAATTGATCAAATGTATATTGATGGGGCCTACCATCATCCGACT
         K  L  P  S  F  I  I  D  Q  M  Y  I  D  G  A  Y  H  H  P  T

36901 - TTTTTATATGAGTCTTTATGGAATATAGGAGTACTCATTATTTTGCTTCTGGTGTCTAGA
         F  L  Y  E  S  L  W  N  I  G  V  L  I  I  L  L  L  V  S  R
                  PRK12437 prolipoprotein diacylglyceryl domain

36961 - AATAGAATGTTCTTTGGGCAAATATTTTTAATTTATCTTTCACTATATTCCGTCGGGAGA
         N  R  M  F  F  G  Q  I  F  L  I  Y  L  S  L  Y  S  V  G  R

37021 - TTCTGGATTGAAGGATTACGGACCGACAGTTTAATGCTAACAGCAAACTTGAGAATGGCT
         F  W  I  E  G  L  R  T  D  S  L  M  L  T  A  N  L  R  M  A

37081 - CAAGTCCTGTCCATAGTCCTATTAATTGGCTCTATATTAACTTACATCTATCTCAAAAAA
         Q  V  L  S  I  V  L  L  I  G  S  I  L  T  Y  I  Y  L  K  K

37141 - TCTAAGGAGGAAGATTTACATGGAAGCATTACTTAATCTCTTACCNCTACTATTACTACT
         S  K  E  E  D  L  H  G  S  I  T  *
                                            Stop
```

Figure 3.21: Nucleotide and deduced amino acid sequence of ORF pCT0030.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 301). Relevant sequence features are identified and discussed in Section 3.3.6.13.

```
  1 - AAAAGTCGGCGTTTGATCCAGCACTTTCATNGATGCCGTCAGCATCTTCGCAACATCGAC
      F  T  P  T  Q  D  L  V  K  M  S  A  T  L  M  K  A  V  D  V

 61 - TTGTTGTGTATCGACATCGACCACGAATGTATTGATTTGTTGCAAGTTTTGTTCGTCAAA
      Q  Q  T  D  V  D  V  V  F  T  N  I  Q  Q  L  N  Q  E  D  F

121 - ACCTTTTATGTATTGCCGTTTATCGTCCGTATATGTTCCCCATCTAAAAACGTTTGGTGT
      G  K  I  Y  Q  R  K  D  D  T  Y  T  G  W  R  F  V  N  P  T

181 - CCAATGCGTGATTTTATCTGCTTCTGTCAGCAGCGCTTCTTGGCTCGTAACAATAAACCC
      W  H  T  I  K  D  A  E  T  L  L  A  E  Q  S  T  V  I  F  G

241 - TTTGACGCCATGCTCCATCTGCAGATGACCTTTATCGCGAACGACAAATTGCACGCCTTT
      K  V  G  H  E  M  Q  L  H  G  K  D  R  V  V  F  Q  V  G  K

301 - ACGATTTTCGCTTGCTGCTTCGACAGCTGCATAGTGATCTCGAATAGCCAAAACGCCTTT
      R  N  E  S  A  A  E  V  A  A  Y  H  D  R  I  A  L  V  G  K

361 - ACTTCCCTTTTTCCGAAAAGCAGTGATTCCTTCATGCACGACCGCATCGATCACGCTGCT
      S  G  K  K  R  F  A  T  I  G  E  H  V  V  A  D  I  V  S  S

421 - ATATTGATACATCGCTTCGTTGTCAAACATCTGATCCTCTCTCCTCTCTTCGTTGATTCT
      Y  Q  Y  M  T  E  N  D  F  M
                                   Start      RBS

481 - CACTTTTAGCAAACAAAAAAGGACAAGATCCCTCCCTGCAGAGAAATCTTGTCCTTTTAA

541 - ACAAAAATCTCGCGCAGGAAAAGGAAACTTTTTCACATGCACGAAACGTTCAAAAACCTT
```

159

Figure 3.22: Nucleotide and deduced amino acid sequence of ORF pCT0031.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 301). Relevant sequence features are identified and discussed in Section 3.3.6.14.

```
1021 -  TTTATAAGTTACTATGTTACTTTATACTACAAATAATTTATATAACCTCCGAATATTTAT

1081 -  CTTTTTTCGCTTTACATAAAATTCAAGGAATGTTATCATGAGTTTATCAGATAAATTTAT
                                                M  S  L  S  D  K  F  I
                        RBS              Start
1141 -  CTTTTTATCCTTTTATCTTTATTACAAATATATTACAGGAGGCGAGAGAATGTTTAATAG
         F  L  S  F  Y  L  Y  Y  K  Y  I  T  G  G  E  R  M  F  N  S

1201 -  TCCGGCTTATATAGAAGAATATAAAAAAGTGCAGAAAGTACTAAGTGATAATCGATGCAG
          P  A  Y  I  E  E  Y  K  K  V  Q  K  V  L  S  D  N  R  C  R

1261 -  AGTTATTATTAATGCAAATCAAAAAGGTGGAGTAGGGAAAACTACCAACACCACTATGGA
          V  I  I  N  A  N  Q  K  G  G  V  G  K  T  T  N  T  T  M  E
                              P-loop site
1321 -  GGCAATTATTCTTGCGCATGTCTTTAATAAAAAAGTTTTATTGATTGATGAAGATATGCA
          A  I  I  L  A  H  V  F  N  K  K  V  L  L  I  D  E  D  M  Q
                      ParA domain
1381 -  AGGTAATGAAACTAGTTTCATGAGTAAAACTTATGATGTTATAGAGTTCCCCATGTCACT
          G  N  E  T  S  F  M  S  K  T  Y  D  V  I  E  F  P  M  S  L

1441 -  AATGAAAGCAATTGAAGAAGGAGACTTAACAAAAGCGATTGTTAATTTAGACGAAAATAT
          M  K  A  I  E  E  G  D  L  T  K  A  I  V  N  L  D  E  N  I

1501 -  TGATATTATCCCTGGTAGTTACGATATGAGAAAAATGGTTAATTTTTTAATTGGTAAATT
          D  I  I  P  G  S  Y  D  M  R  K  M  V  N  F  L  I  G  K  F

1561 -  CAAAACTGAAGAAGCACAAACTTTTTATCTCTCTTCACTTATAGACAAAATTCGTGATGA
          K  T  E  E  A  Q  T  F  Y  L  S  S  L  I  D  K  I  R  D  D

1621 -  CTATGATTTTATTTTTATTGATGTTCCACCTTCTACAGATTTAAAAGTAGATAATGCTGT
          Y  D  F  I  F  I  D  V  P  P  S  T  D  L  K  V  D  N  A  V
                              ParA domain
1681 -  AGTTGCTGCTGATTACTTAGTTGTGGTTCAAGAAACTCAACAATTTGCATTAGAAGGCTC
          V  A  A  D  Y  L  V  V  V  Q  E  T  Q  Q  F  A  L  E  G  S

1741 -  AAACACCTTTATTTCAACTTATCTTAATACCTTAGTAGATGATTTTGGCAGTCGCTTCAA
          N  T  F  I  S  T  Y  L  N  T  L  V  D  D  F  G  S  R  F  K

1801 -  AACAGAAATAATTGGTATTTTACCAGTTTTACTACAGAAAAAAAGAAAATTGCATGAAAA
          T  E  I  I  G  I  L  P  V  L  L  Q  K  K  R  K  L  H  E  K

1861 -  AATACTAGAAACTACAAAAGAAAAATTTGGTAAAGAAAATATATTTGGAACTACAATAAA
          I  L  E  T  T  K  E  K  F  G  K  E  N  I  F  G  T  T  I  N
                              Soj domain
1921 -  CAATCACGCAAGACTAGAGTTATATGGTAAATTTGGCGTTCAATTTGAAGATCATTGGGA
          N  H  A  R  L  E  L  Y  G  K  F  G  V  Q  F  E  D  H  W  D

1981 -  TCGAACAATGTTTGCACTCTATACTGATATAGTACAAGAGATGCTTGAACGTATGCAGCT
          R  T  M  F  A  L  Y  T  D  I  V  Q  E  M  L  E  R  M  Q  L

2041 -  TATTGAAAACGGCAGTGATTTAGATAACTATCAATATAAACCAATTTTTTACAATCCAAA
          I  E  N  G  S  D  L  D  N  Y  Q  Y  K  P  I  F  Y  N  P  K
```

160

```
2101 -  AATTGGTAAAGTAACTGCTTTGGGAAAGGAGATTGTTGTTAATGGCTCGATTAATAGATC
          I  G  K  V  T  A  L  G  K  E  I  V  V  N  G  S  I  N  R  S

2161 -  GTAAAGACAAGGAACAAGCACAAGTAAAAACTCAAGCAGATACATCGAGAAAAGTTAATA
         *
        Stop
2221 -  CAAAGTCCAACTCCGAAAAAAAAGTTAATGCAAAAGACAGGAAAAATATTAAGGTTAACC
```

Figure 3.23: Nucleotide and deduced amino acid sequence of ORF pCT0032.

Nucleotide sequence numbering is relevant to sections of the GenBank entry U15553 shown in Appendix A (page 302). Relevant sequence features are identified and discussed in Section 3.3.6.15.

```
2041 - TATTGAAAACGGCAGTGATTTAGATAACTATCAATATAAACCAATTTTTTACAATCCAAA

2101 - AATTGGTAAAGTAACTGCTTTGGGAAAGGAGATTGTTGTTAATGGCTCGATTAATAGATC
                                                  M   A   R   L   I   D   R
                               RBS                    Start
2161 - GTAAAGACAAGGAACAAGCACAAGTAAAAACTCAAGCAGATACATCGAGAAAAGTTAATA
          K   D   K   E   Q   A   Q   V   K   T   Q   A   D   T   S   R   K   V   N   T

2221 - CAAAGTCCAACTCCGAAAAAAAAGTTAATGCAAAAGACAGGAAAAATATTAAGGTTAACC
          K   S   N   S   E   K   K   V   N   A   K   D   R   K   N   I   K   V   N   P

2281 - CTGACGATATGTTATTAATTGATACAATAGCTAATAATTTCTTTGAAACAAAATTAAAAC
          D   D   M   L   L   I   D   T   I   A   N   N   F   F   E   T   K   L   K   H

2341 - ATTATGAATTGGTAAATCTCTTATTAGAAAACTGGGTGGATACTAAGCTAGAACCAAGAC
          Y   E   L   V   N   L   L   L   E   N   W   V   D   T   K   L   E   P   R   Q

2401 - AACAAAAGATACTTAAGGATTTATTGAAAAACAGATAGTTTTATCTAAAGATAATTAGAT
          Q   K   I   L   K   D   L   L   K   N   R   *
                                                   Stop
2461 - AAAAAGATAGTTTTATCTAATTATCTTTTTATCCTTTATGCCCACCGACCAAATCAGATC
```

Figure 3.24: DNA regions conserved between plasmids pCT100, pLI100 and pLM80.

The genetic organisation of all characterised ORFs or genes encoded by plasmids pCT100 of *L. monocytogenes* strain DRDC8, pLI100 of *L. innocua* and contiguous sequence 1 and 2 of plasmid pLM80 from *L. monocytogenes* strain 4b H7858 is aligned using the Mauve 2 (Progressive Mauve) genome alignment algorithm. Mauve identifies and aligns regions of local collinearity called locally collinear blocks (LCBs). Each locally collinear block is a homologous region of sequence shared by two or more of the plasmid sequences and does not contain any rearrangements of homologous sequence. Each sequence of identically coloured blocks represents a collinear set of matching regions. These are shown below their respective locations. Uncoloured regions within the blocks represent areas of nucleotide sequence that it non-homologous to the corresponding LCBs of other plasmid sequence. ORFs or genes that encode transposase/invertase elements are shown in red.

pCT100
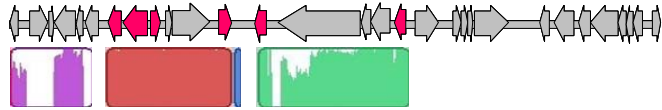
pLI100
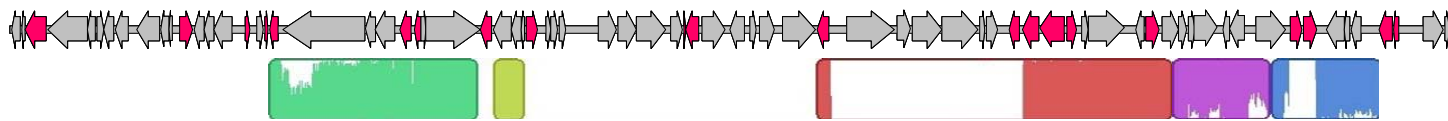

pLM80 Cont 1

pLM80 Cont 2

164
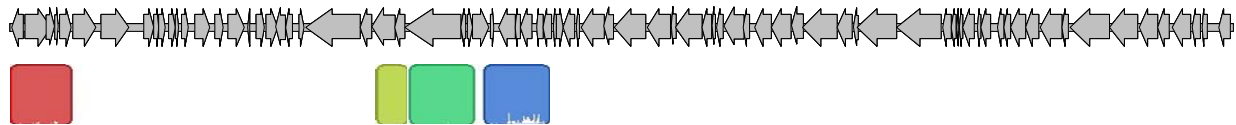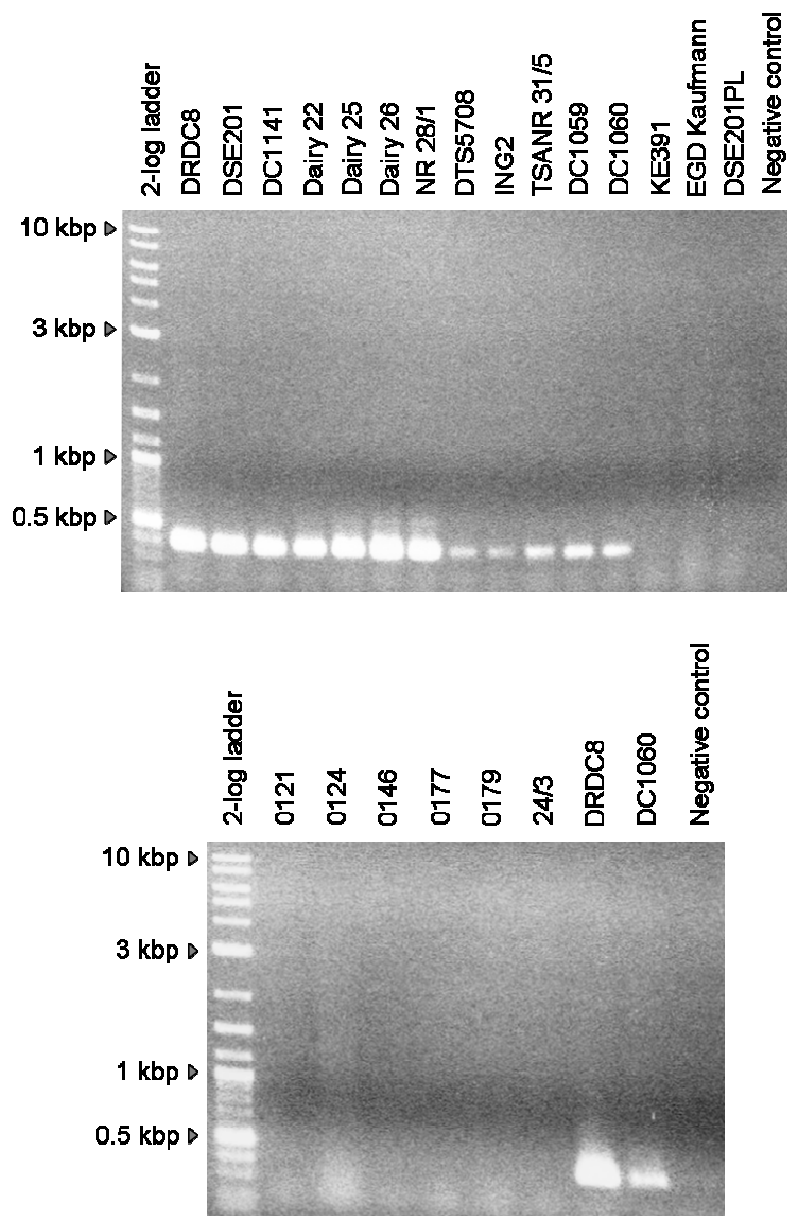
Figure 3.25: PCR amplification of the *ctpA* gene from *L. monocytogenes* isolates.

A 396 bp *ctpA* fragment was amplified using the FB001/LM2004 oligonucleotide pair from *L. monocytogenes* strains DRDC8, DSE201, DC1141, Dairy 22, Dairy 25, Dairy 26, NR28/1, DTS5708, ING2, TSANR 31/5, DC1059 and DC1060 DNA. Note the different amounts of amplicon produced for DRDC8, DSE201, DC1141, Dairy 22, Dairy 25, Dairy 26 and NR28/1 compared with DTS5708, ING2, TSANR 31/5, and DC1059 and DC1060. As expected, no amplicons were obtained for the no DNA negative control or from DNA isolated from the *ctpA* negative strains KE391, EGD Kaufmann and DSE201PL. No amplicons were produced from DNA of strains 0121, 0124, 0146, 0177, 0179 and 24/3.
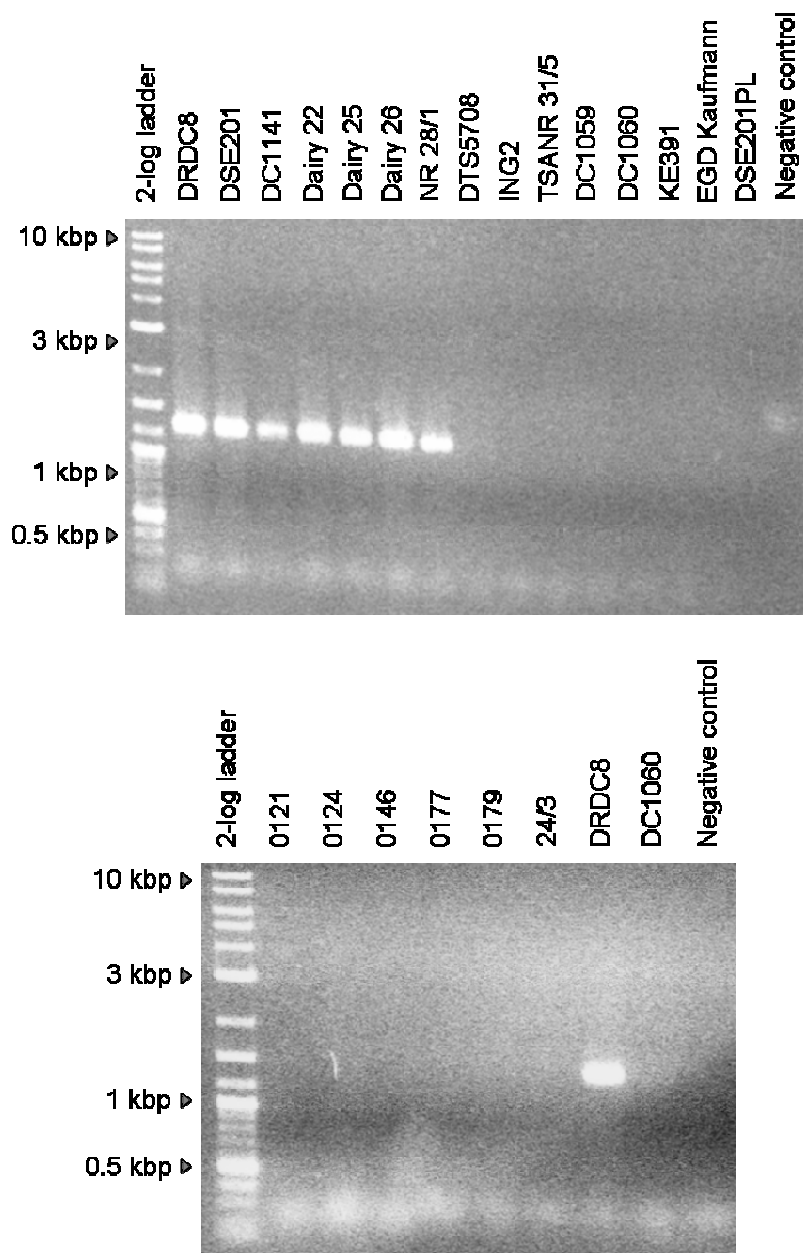
Figure 3.26: PCR amplification of ORF pCT0025 from *L. monocytogenes* isolates.

A 1339 bp fragment of ORF pCT0025 was amplified using the pDR6/pDR10R oligonucleotide pair from *L. monocytogenes* strains DRDC8, DSE201, DC1141, Dairy 22, Dairy 25, Dairy 26 and NR28/1 DNA. Amplicons were not produced from DNA of *ctpA* positive isolates that produced low levels of the *ctpA* amplicon (DTS5708, ING2, TSANR 31/5, DC1059 and DC1060) as shown in Figure 3.25. No amplicon was produced from the no DNA negative control or from DNA of *ctpA* negative isolates (KE391, EGD Kaufmann, DSE201PL, 0121, 0124, 0146, 0179 and 24/3).
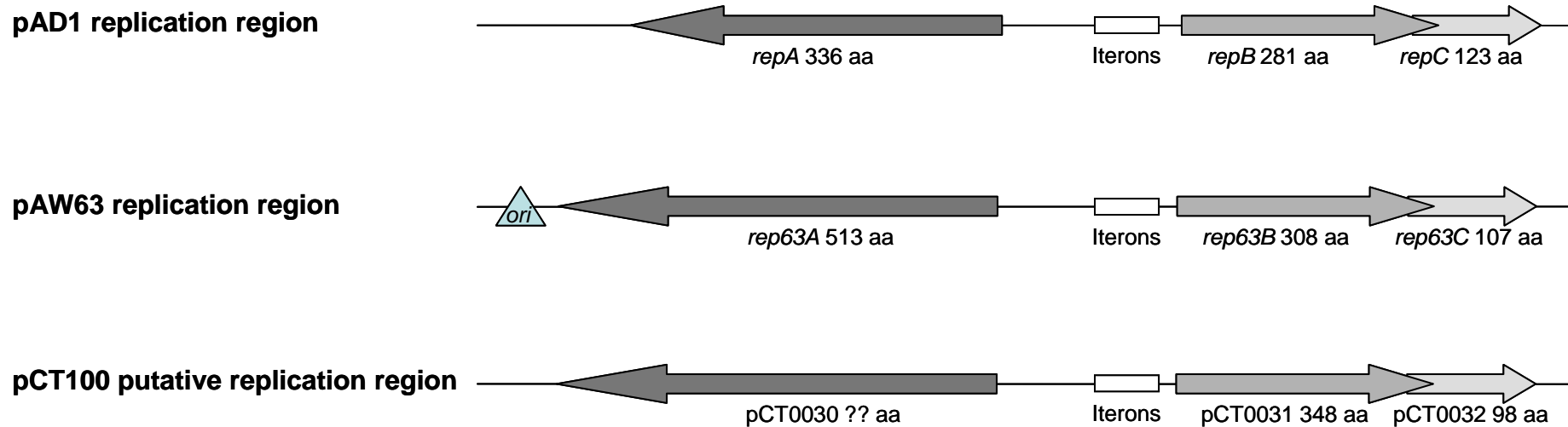
Figure 3.27: Genetic map of the replication regions of plasmids pAD1 and pAW63 and the putative replication region of plasmid pCT100.

The genetic organisation and orientation of genes that comprise the replication regions for plasmids pAD1 and pAW63 is shown in comparison to the putative replication region of plasmid pCT100, comprised of ORFs pCT0030, pCT0031 and pCT0032. The amino acid length of each ORF or gene is indicated, however, further sequence data is required to determine the length of pCT0030. The location of each iteron region for each replication region (repeated DNA motifs) is shown. The coding regions of the 'repB-type' and 'repC-type' genes overlap in each replicon. For the pAW63 replicon, replication is initiated an 85 bp origin (ori) region, a cis-functioning locus downstream of rep63A. See Figure 3.28 for the relevant nucleotide features of the putative replicon of pCT100.

Figure 3.28: Nucleotide sequence of the pCT100 putative plasmid replication region.

The nucleotide sequence begins at the start of ORF pCT0030 and extends to the stop of ORF pCT0032. TCTTTTTATC repeats located between the start of pCT0030 and the start of pCT0031 are underlined with arrows, reflective of the orientation. CTATATA repeats are underlined. Mismatches from the consensus sequence are indicated with small triangles. Regions highlighted in bold represent regions of high A+T content. Inverted repeats located immediately downstream of pCT0032 are indicated by converging thick lined arrows. A TCCTCTCTCCTCTCTTC repeat located immediately upstream of pCT0030 is indicated by a single think lined arrow.

```
 361 ACTTCCCTTTTTCCGAAAAGCAGTGATTCCTTCATGCACGACCGCATCGATCACGCTGCT
 421 ATATTGATACATCGCTTCGTTGTCAAACATCTGATCCTCTCTCCTCTCTTCGTTGATTCT
```

pCT0030 start ←

```
 481 CACTTTTAGCAAACAAAAAAGGACAAGATCCCTCCCTGCAGAGAAATCTTGTCCTTTTAA

 541 ACAAAAATCTCGCGCAGGAAAAGGAAACTTTTTCACATGCACGAAACGTTCAAAAACCTT

 601 ATAAAAATGCTTGAAAACTTGCTTTTAAGCATTGTCAAATAAGCCTACTTATGCTATAAT

 661 TTTTTTATCAATTTTAAATTAATAGCTTGTAGGGTTTCGTCTTTTTAATGTGTTTTTGGT

 721 CGAACACATTAAAGTGCATTTTGGTCAATGCATGGGCGATTCCTTTTTTATTTACTTGTT

 781 GTTTACACAATACAACAAGTTTTTTTTTTTGTAAACTATCATTTTTTCTCACAATAAGCT

 841 TACTCCTATATAATATTTGATAAAAAGATATAAATATCTTTTTATCTTTTTATCTATATA

 901 TACTTTTATCTTTTTATCTCTTTATCTTTTTATCTATATATCTCATTATCTTTTTATTTT

 961 TTTATCTATATATCTTTTTATCTCCTGTAAAGTATTGCTGTTATGCAATTTACAATTAAT

1021 TTTATAAGTTACTATGTTACTTTATACTACAAATAATTTATATAACCTCCGAATATTTAT

1081 CTTTTTTCGCTTTACATAAAATTCAAGGAATGTTATCATGAGTTTATCAGATAAATTTAT
```

→ pCT0031 start

```
1141 CTTTTTATCCTTTTATCTTTATTACAAATATATTACAGGAGGCGAGAGAATGTTTAATAG
1201 TCCGGCTTATATAGAAGAATATAAAAAAGTGCAGAAAGTACTAAGTGATAATCGATGCAG
1261 AGTTATTATTAATGCAAATCAAAAAGGTGGAGTAGGGAAAACTACCAACACCACTATGGA
1321 GGCAATTATTCTTGCGCATGTCTTTAATAAAAAAGTTTTATTGATTGATGAAGATATGCA
1381 AGGTAATGAAACTAGTTTCATGAGTAAAACTTATGATGTTATAGAGTTCCCCATGTCACT
1441 AATGAAAGCAATTGAAGAAGGAGACTTAACAAAAGCGATTGTTAATTTAGACGAAAATAT
1501 TGATATTATCCCTGGTAGTTACGATATGAGAAAAATGGTTAATTTTTTAATTGGTAAATT
1561 CAAAACTGAAGAAGCACAAACTTTTTATCTCTCTTCACTTATAGACAAAATTCGTGATGA
1621 CTATGATTTTATTTTTATTGATGTTCCACCTTCTACAGATTTAAAAGTAGATAATGCTGT
1681 AGTTGCTGCTGATTACTTAGTTGTGGTTCAAGAAACTCAACAATTTGCATTAGAAGGCTC
1741 AAACACCTTTATTTCAACTTATCTTAATACCTTAGTAGATGATTTTGGCAGTCGCTTCAA
1801 AACAGAAATAATTGGTATTTTACCAGTTTTACTACAGAAAAAAAGAAAATTGCATGAAAA
1861 AATACTAGAAACTACAAAAGAAAAATTTGGTAAAGAAAATATATTTGGAACTACAATAAA
```

pCT0031 stop ←

```
1921 CAATCACGCAAGACTAGAGTTATATGGTAAATTTGGCGTTCAATTTGAAGATCATTGGGA

1981 TCGAACAATGTTTGCACTCTATACTGATATAGTACAAGAGATGCTTGAACGTATGCAGCT

2041 TATTGAAAACGGCAGTGATTTAGATAACTATCAATATAAACCAATTTTTTTACAATCCAAA

2101 AATTGGTAAAGTAACTGCTTTGGGAAAGGAGATTGTTGTTAATGGCTCGATTAATAGATC
```

→ pCT0032 start

```
2161 GTAAAGACAAGGAACAAGCACAAGTAAAAACTCAAGCAGATACATCGAGAAAAGTTAATA
2221 CAAAGTCCAACTCCGAAAAAAAAGTTAATGCAAAAGACAGGAAAAATATTAAGGTTAACC
2281 CTGACGATATGTTATTAATTGATACAATAGCTAATAATTTCTTTGAAACAAAATTAAAAC
2341 ATTATGAATTGGTAAATCTCTTATTAGAAACTGGGTGGATACTAAGCTAGAACCAAGAC
2401 AACAAAAGATACTTAAGGATTTATTGAAAACAGATAGTTTTATCTAAAGATAATTAGAT
```

pCT0032 stop ←   →

```
2461 AAAAAGATAGTTTTATCTAATTATCTTTTTATCCTTTATGCCCACCGACCAAATCAGATC
```

→   ←

```
2521 GTTTCAATCCTGTAGCCCCGGCTACTAAGCTATTGGCATGCATCAATGCTTTAAATCCAT
```