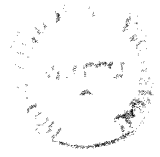
**University of Southampton**

# The Molecular Biology of Chlamydiaphages

Submitted by Sarah Ann Garner
for the degree of Doctor of Philosophy



**Faculty of Medicine**
**Division of Molecular Microbiology and Infection**

Submitted July 2003

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF MEDICINE
DIVISION OF MOLECULAR MICROBIOLOGY AND INFECTION

Doctor of Philosophy

# The Molecular Biology of Chp2

By Sarah Garner

Chlamydiae are obligate intracellular bacteria that have a unique developmental cycle. They are responsible for causing a wide range of human and animal diseases. As there is no natural system for gene transfer in chlamydiae and due to the inaccessibility of the bacteria within the host cell, studies into the molecular biology of chlamydiae have been slow. Bacteriophages infecting chlamydiae may provide the basis for a gene transfer system. Therefore study into the molecular biology of these bacteriophages is extremely interesting.

Chlamydiaphages belong to the *Microviridae* family of bacteriophages. These consist of lytic isometric phages with single-stranded circular DNA genomes. Members of the *Microviridae* family fall into two distinct sub-families. One sub-family that includes bacteriophages infecting enterobacteriaceae, and a second sub-family containing the chlamydiaphages, ϕMH2k and SpV4. Four chlamydiaphages have been discovered to date infecting various hosts; Chp1 (*C. psittaci*), Chp2 (*C. abortus*), ϕCPAR39 (*C. pneumoniae*) and ϕCPG1 (*C. caviae*). Whilst the genomes of Chp2, ϕCPAR39, ϕCPG1 remain highly conserved Chp1 shows high divergence throughout its genome.

A novel bacteriophage was discovery during this study (Chp3). Chp3 infects a chlamydial species not previously known to carry bacteriophages (*C. pecorum*). The genome of Chp3 is 4,554 bp and encodes eight open reading frames that are organised in the same relative context as other chlamydiaphages. Chp3 is closely related to the Chp2-like bacteriophages. Interestingly while the VP1 of the Chp2-like bacteriophages remains highly conserved two regions show high divergence between Chp2/Chp3 and ϕCPG1/ϕCPAR39. It has been hypothesised that these regions are the chlamydiaphage receptor-recognition domain. The host range of Chp3 is identical to the host range of Chp2 (*C. abortus*, *C. pecorum*, *C. felis* and *C. caviae*). ϕCPAR39 is also able to infect *C. abortus*, *C. caviae* and *C. pecorum*, however, it is unable to infect *C. felis* and can additionally infect *C. pneumoniae*. This suggests that these two sub groups of chlamydiaphages use different surface receptor molecules.

It has been predicted that the chlamydiaphage genome is arranged into eight major open reading frames (ORF1-8). Immuno-specific reagents were produced to the proteins encoded by ORF2-8 and a monoclonal antibody to VP1 was identified. These reagents were then used to screen Chp2 infected *C. abortus* inclusions and semi-purified Chp2 particles. The expression of VP1, VP3, ORF5 protein ORF4 protein and ORF7 protein was demonstrated.

It is likely that chlamydiaphages regulate their gene expression in a similar way to other bacteriophages in the *Microviridae*. By comparing the genome organisation of Chp2 to two members of this family of bacteriophages SpV4 and ϕX174 two promoter regions were identified in the Chp2 genome situated upstream of the ORF4 and ORF5 translational initiation codons. Using promoter selection vectors it was shown that these two promoter regions are functional in *E. coli*. The promoter situated upstream of the ORF5 translational initiation codon is a stronger *E. coli* promoter than the ORF4 promoter. This may be used as a mechanism for the regulation of gene expression and used to control when and how much of each protein is produced.

# Publications and Presentations

Sections of work presented in this thesis has been published:

Sarah A. Garner, J. Sylvia Everson, Paul R. Lambden, Bentley A. Fane, Ian N. Clarke. Isolation, molecular characterisation and genome sequence of a bacteriophage (Chp3) from *Chlamydophila pecorum*. Virus genes. In Press.

Everson,J.S., Garner,S.A., Lambden,P.R., Fane,B.A, Clarke,I.N. (2003). Host Range of Chlamydiaphages ϕCPAR39 and Chp3. Journal of Bacteriology *185*, 21, 6490-6492.

Everson,J.S., Garner,S.A., Fane,B., Liu,B.-L., Lambden,P.R., Clarke,I.N. (2002). Biological Properties and Cell Tropism of Chp2, A Bacteriophage of the Obligate Intracellular Bacterium *Chlamydophila abortus*. Journal of Bacteriology *184*, 2748-2754.

Garner, S., Everson, S., Fane, B., Lambden, P. R., and Clarke, I. N. Discovery and molecular characterisation of a bacteriophage infecting *C. pecorum*. 25-28. 2002a. Antalya, Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B.

Garner, S. A., Everson, J. S., Pickett, M. A., Lambden, P. R., and Clarke, I. N. Effects of Chlamydaphage 2 Infection on Host Transcription: Active promoters within the Chp2 genome. 619-622. 2002b. Antalya, Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., and Timms, P.

# Figures

# Tables

# Acknowledgements

I am very grateful to my supervisors Professor Ian Clarke and Dr . Paul Lambden for their continual advice and encouragement throughout my PhD studentship. My thanks also go out to Shirley Sawyers for providing a constant supply of sterile tips tubes and other materials and to my friends and colleagues for their assistance during the last three years. I would also like to thank the MRC for funding this project.

# Abbreviations

| | |
|---|---|
| Å | Ångstrom |
| $A_{260}$ | Absorbance 260 nanometres |
| $A_{550}$ | Absorbance 550 nanometres |
| $A_{600}$ | Absorbance 600 nanometres |
| aa | amino acid |
| Amp | Ampicillin |
| APS | Ammonium persulphate |
| ATP | Adenosine triphosphate |
| BCIP | 5-Bromo-4-chloro-3-indolyl phosphate |
| bp(s) | Base pair(s) |
| BSA | Bovine serum albumin |
| °C | Degrees centigrade |
| cDNA | complementary Deoxyribonucleic acid |
| $dH_2O$ | Deionised water |
| DMSO | Dimethyl sulphoxide |
| DNA | Deoxyribonucleic acid |
| ds | double stranded |
| DTT | Dithiothreitol |
| dATP | 2´deoxy adenosine 5´triphosphate |
| dCTP | 2´deoxy cytidine 5´triphosphate |
| dGTP | 2´deoxy guanosine 5´triphosphate |
| dTTP | 2´deoxy thymidine 5´triphosphate |
| dNTP(s) | 2' deoxynucleoside 5'-triphosphate(s) |
| ddNTP(s) | 2', 3' –dideoxynucleoside 5' -triphosphate(s) |
| EDTA | Ethylenediaminetetra-acetic acid |
| EIA | Enzyme immuno assay |
| ELISA | Enzyme linked immunosorbant assay |
| EM | Electron Microscopy |
| FCS | Foetal calf serum |
| FLC | Full length clone |
| $^{x}g$ | x gravity |
| g | gram |
| IEM | Immune electron microscopy |
| IgG | Immunoglobulin G |
| IMS | Industrial methylated spirit |
| IPTG | Isopropyl-β-D-thio-galactoside |
| Kb(s) | Kilobase(s) |
| KDa | Kilodalton |
| l | Litre |
| LB | Luria-Bertani |
| M | Molar |
| mA | milliamp |
| mg | milligram |
| ml | millilitre |
| mm | millimetre |
| moi | multiplicity of infection |
| $M_r$ | Relative molecular mass |

| | |
|---|---|
| mRNA | messenger ribonucleic acid |
| MW | molecular weight |
| NBT | Nitro-blue tetrazolium |
| nm | nanometre |
| nt | nucleotide |
| ORF | Open reading frame |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| RNA | Ribonucleic acid |
| rpm | Revolutions per minute |
| RT-PCR | Reverse transcription PCR |
| SDS | Sodium dodecyl sulphate |
| ss | single stranded |
| TB | Terrific broth |
| TBS | Tris buffered saline |
| TEMED | N,N,N′,N′ tetramethylethylenediamine |
| Tm | melting temperature |
| TNT | Transcription and translation |
| Tris | Tris (hydroxymethyl)aminomethane |
| TTBS | Tween Tris buffered saline |
| u | unit |
| UHQ | Ultra high quality (water) |
| UTR | Untranslated region |
| UV | Ultraviolet |
| V | Volts |
| v/v | volume per unit volume |
| w/v | weight per unit volume |
| µg | microgram |
| µl | microlitre |
| µM | micromolar |
| µm | micrometre |

**Single and Three Letter Amino Acid codes**

| | | |
|---|---|---|
| Ala | (A) | Alanine |
| Arg | (R) | Arginine |
| Asp | (D) | Aspartate |
| Asn | (N) | Asparagine |
| Cys | (C) | Cysteine |
| Gln | (Q) | Glutamine |
| Glu | (E) | Glutamate |
| Gly | (G) | Glycine |
| His | (H) | Histidine |
| Ile | (I) | Isoleucine |
| Leu | (L) | Leucine |
| Lys | (K) | Lysine |
| Met | (M) | Methionine |
| Phe | (F) | Phenylalanine |
| Pro | (P) | Proline |
| Ser | (S) | Serine |
| Thr | (T) | Threonine |
| Trp | (W) | Trytophan |
| Tyr | (Y) | Tyrosine |
| Val | (V) | Valine |

# Chapter 1
# Introduction

## 1.1 Chlamydiae

### 1.1.1 Historical perspectives

Scientific interest in chlamydiae came in flurries during the last century as associations were made between the organism and disease. Interest greatly increased in the 1970s when an association with genital disease was found (Grayston and Wang, 1975). The chlamydial inclusion was first described in 1907 and was named Chlamydozoa (mantle bodies) (Halberstaedter and von Prowazek, 1907). They were initially described as protozoans, but were later reassigned as large viruses when they were found to be the causal agent in the 1929 psittacosis outbreak (Potter *et al.*, 1979). Stanier and Lwoff defined chlamydiae as bacteria, based on their work that distinguished viruses from bacteria (Moulder, 1964; Moulder, 1966). The basic properties of bacteria, viruses and chlamydiae taken from Schachter and Caldwell (1980) can be seen in the table below (Schachter and Caldwell, 1980).

**Table 1.1 Basic properties of bacteria, chlamydiae and viruses**

| | Bacteria | Chlamydiae | Viruses |
|---|---|---|---|
| Size (nm) | 300-3,000 | 200-1,000 (EB-RB) | 15-350 |
| Intracellular parasites | - | + | + |
| Replication | Binary fission | Binary fission | Eclipse-synthesis-Assembly |
| Ribosome's present | + | + | - |
| Appearance | Varied | Coccoid | Symmetrical |
| Metabolic enzymes | + | + | - |
| Energy production | + | - | - |
| Complex cell wall | + | + | - |

Chlamydiae share some properties with viruses, however, as seen from the table above, the basic molecular properties of chlamydiae are bacterial.

### 1.1.2 Chlamydial developmental cycle

Chlamydiae are bacteria that have a unique obligate intracellular developmental cycle that takes place inside the cytoplasm of eukaryotic cells. Chlamydial infection of cells involves three basic stages, attachment and entry into the cell, interactions with the intracellular environment and replication, followed by exit from the cell. The infectious form of the organism is the elementary body (EB), which attaches to receptive cells via projections on the host cell surface. It is thought that different species of chlamydiae use different means for the initial interaction with the host cell (Kuo and Grayston, 1976). A heparan sulfate-like compound has been shown to function as an attachment receptor for chlamydiae, these are also common attachment means for other pathogens (Spears and Storz, 1979; Zhang and Stephens, 1992; Duensing *et al.*, 1999; Wuppermann *et al.*, 2001). Candidates for the chlamydial adhesion molecule include the major outer membrane protein (MOMP) and the 60kDa cysteine rich protein (Omc2), both of which are abundant on the chlamydiae EB and have been shown to bind heparin (Su *et al.*, 1990; Stephens *et al.*, 2001). Other studies have suggested that chlamydiae do not produce their own heparan sulfate-like ligand and instead parasitize these compounds from the host cell (Taraktchoglou *et al.*, 2001).

After attachment, interactions between the host cell and chlamydiae occur including endocytosis, tyrosine phosphorylation events and cellular cytoskeletal rearrangements as the EBs penetrate the host cell via either receptor mediated endocytosis or parasite-specified phagocytosis (Birkelund *et al.*, 1994; Fawaz FS *et al.*, 1997; Majeed *et al.*, 1999; Krull *et al.*, 1999). Chlamydiae grow and divide within a membrane structure known as an inclusion. To avoid host cell defences, viable chlamydiae can block lysosomal fusion and degranulation inside the inclusion (Eissenberg and Wyrick, 1981). In the first few hours after entering the cell the EB is re-organized into a larger structure called the reticulate body (RB). The bacteria then replicate by binary fission within the inclusion every 2-3 hours. To contain the increasing numbers of bacteria, the inclusion membrane enlarges extensively yielding a phagosome which takes up over three-quarters of the volume of the cell (Ward, 1983).

Chlamydiae are in constant interaction with the host cell whilst inside the inclusion and acquire nucleotides, amino acids and other precursors from the host cell (Hatch, 1975; McClarty, 1994). It has been shown that they can actively redirect their endocytic vacuole to an exocytic vesicular trafficking pathway to obtain certain lipids from the Golgi apparatus directly into the inclusion (Hackstadt *et al.*, 1995; Hackstadt *et al.*, 1996; Wylie *et al.*, 1997). Chlamydial proteins may interact with the host cell to alter the properties of the inclusion membrane to allow cellular trafficking. To transport the essential nutrients into the inclusion it is thought that chlamydiae parasitize the host cell for adenine triphosphate (ATP) (Hatch *et al.*, 1982). This is verified by the expression of two ATP transports that are expressed early in development (Shaw *et al.*, 2002). However, it has also been shown both by genomic and experimental evidence that they are able to synthesise ATP (Stephens *et al.*, 1998; Iliffe-Lee and McClarty, 1999; Read *et al.*, 2000a).

Approximately 12-18 hours after the initial infection the chlamydial developmental cycle becomes asynchronous and some RBs undergo a reorganization and condense back into the infectious form of the organism (Ward, 1983). The length of the developmental cycle varies depending on the chlamydial strain. At about 20 to 30 hours after the initial infection host cell degradation begins to occur and the bacteria exit the host cell (Ward, 1983). It has been demonstrated that different species of chlamydiae exit the host cell in different ways, some appear to leave the host cell intact and exit via exocytosis, other species exit by lysis and show an active necrosis effect on cells (Todd and Caldwell, 1985; Belland *et al.*, 2001). The EBs are then able to infect new host cells and begin the cycle again.

### 1.1.3 Morphology

The two forms of chlamydiae have been specifically adapted to their unique territory in the developmental cycle. The EB, which is the infectious form of the organism, is adapted for extracellular endurance and is rigid and uniform with an electro-dense core of DNA (Matsumoto, 1982). The EB rigidity is due mainly to two cysteine-rich proteins, which may perform the same function and act as a substitute for peptidoglycan (Hatch, 1996). EBs are spherical bodies with a diameter of 0.2-0.3 micrometer. The cell wall of the EB is composed of a granular outer layer and an inner layer that is composed of hexagonal array structures (Manire, 1966).

The RBs, which are the replicative form of the organism, are adapted for intracellular replication, and in contrast are more delicate and larger in size than the EB with a diameter of 0.5-1.6 micrometer. The cytoplasm inside the RB looks granular due to ribosome synthesis (Ward, 1983). The cell wall of RB resembles those of Gram-negative bacteria as it has an inner and outer membrane and the cell envelope is pleomorphic due to the absence of any forms of structural stabilisation (Matsumoto, 1982).

Conversion from the RB to the EB requires the synthesis of late phase proteins. These include histone-like proteins (Hackstadt, 1991) that bind to the chlamydial genome mediating condensation of the genome and inhibit transcription and translation. Plus the cysteine rich outer membrane proteins that are cross-linked by disulfide bonds, and form a complex structure together with MOMP that gives the EB envelope its rigid form (Hatch, 1996). Analysis of the chlamydial outer membranes have shown that all chlamydiae possess LPS, MOMP, Omp2 and Omp3 in this region. It has been shown that LPS, MOMP, Omp2 and the more recently discovered POMP proteins (Longbottom *et al.*, 1998b) are surface exposed in the outer membranes of both EBs and RBs (Birkelund *et al.*, 1988; Birkelund *et al.*, 1989; Longbottom *et al.*, 1998a). The *C. pneumoniae* Pmps have also been shown to be surface exposed (Vandahl *et al.*, 2002). Omp3 is not surface exposed and may be linked to Omp2 (Collett *et al.*, 1989; Hatch, 1996).

Cylindrical projections have been observed on the surfaces of chlamydiae and it is likely that these projects allow direct contact between the bacteria and the inclusion membrane. The projections were observed on both EBs and RBs but were more numerous on the RBs (Wyrick, 2000). It is hypothesised that these spikes may serve as a type III secretory system, serving as a channel allowing communication to the outside of the inclusion (Hsia *et al.*, 1997). Genes have been identified in *C. pneumoniae* and *C. trachomatis* that have homologies to type III secretory proteins (Stephens *et al.*, 1998; Kalman *et al.*, 1999).

Other secretory proteins have been discovered in chlamydiae including Inc proteins (Rockey *et al.*, 1995) and CPAF (Zhong *et al.*, 2001). Inc proteins are localized in the chlamydial inclusion membrane (Rockey *et al.*, 1995) and may have roles in inclusion development, vesicle trafficking, nutrient acquisition, avoidance of lysosomal fusion and structural reorganisations. The CPAF proteins (protease-like activity factor), are secreted into the host cytoplasm and are thought to provide protection against host immune recognition by activating transcription factors to degrade MHC class I and II molecules (Zhong *et al.*, 2001). Proteins distantly related to the Pomp/Pmp genes have also been identified. The pmp proteins resemble members of the autotransporter family of proteins (Mygind *et al.*, 2000; Henderson and Lam, 2001). These are proteins that are transported across the plasma membrane.

### 1.1.4 Classification and Taxonomy

Classification of chlamydiae has changed over the past years as more elaborate phylogenetic classification mechanisms have been developed. In 1980 when the approved lists of bacterial names was published (Euzéby, 1980) the family *Chlamydiaceae* contained one genus *Chlamydia* and two species *C. trachomatis* and *C. psittaci.* Classification was based on glycogen accumulate and sulfonamide susceptibility *C. trachomatis* strains accumulated glycogen in their inclusions and were sensitive to sulfadiazine. Each of these species contained a wide variety of organisms, and even before the list was published, strains of chlamydiae had been isolated which could not be placed in either of the two species (Dwyer *et al.*, 1972; Darougar *et al.*, 1980). During the 1980s new DNA based classification techniques were developed allowing the further differentiation of the chlamydial groups. These techniques, especially DNA-DNA reassociation, supported the presence of eight species of chlamydiae and led to the development of two new species in

the genus *Chlamydia*; *C. pneumoniae* and *C. pecorum* (Grayston, 1989; Fukushi and Hirai, 1992). In 1993 a ninth group of *Chamydiaceae* was identified from swine, described as a *Chlamydia trachomatis*-like *Chlamydia* (Kaltenboeck *et al.*, 1993). These nine groups of *Chlamydiaceae* are still supported by phylogenetic analysis, using the ribosomal operon and analyses of antigenicity, associated disease, phenotype, host range and biological data (Kaltenboeck *et al.*, 1993; Everett and Andersen, 1997; Pudjiatmoko *et al.*, 1997; Takahashi *et al.*, 1997).

In 1999 the order chlamydiales was revised based on 16S and 23S rRNA sequence similarity (Everett *et al.*, 1999). It was proposed that the family *Chlamydiaceae* be divided into two genera: *Chlamydia and Chlamydophila.* The genera *Chlamydia* contains three species: *C. muridarum; C. suis*, and *C. trachomatis*. The genera *Chlamydophila* contains six species: *C. caviae; C. felis; C. pecorum; C. pneumoniae; C. abortus* and *C. psittaci.* All members of the family *Chlamydiaceae* have high sequence identity (>90%) between their 16S rRNA gene. Other groups of *Chlamydia*-like organisms have also been identified which have high 16S rRNA sequence identity with chlamydiae (>80%) and have been assigned into additional families within the order *Chlamydiales*: *Parachlamydiaceae, Waddliaceae* and *Simkaniaceae* (Everett *et al.*, 1999).

The proposal to change the taxonomic nomenclature of the family chlamydiaceae (Everett *et al.*, 1999) has not been widely accepted by scientists in the field (Schachter *et al.*, 2001). A review (Stephens, 2002) highlights the main problems with the new nomenclature. It is accepted that *C. trachomatis, C. pneumoniae* and *C. psittaci* are separate species, but there is debate as to whether there is sufficient divergence to place them in separate genera. It was proposed that to be included in the genus *chlamydia* the 16S rRNA genes should have a 95% identity to *C. trachomatis*, based on observations seen in other free living bacteria. However, chlamydiae differ from free living bacteria in two main ways, their lack of horizontal gene transfer and faster mutation rates. This is due to their obligate intracellular developmental cycle and may make the 95% ruling theory inapplicable. Due to these reasons the rRNA clocks for free living bacteria and chlamydiae may not be synchronized, which has lead to an over estimation in the evolutionary difference. Stephens proposes that the rRNA clocks of chlamydiae are more like those of endosymbionts than free-living bacteria. Endosymbionts like chlamydiae only grow within cells, lack horizontal gene

transfer and have small genomes. They also accumulate fixed mutations quicker than free-living bacteria. Endosymbionts have been shown to accumulate mutations 1.7-2.7-times faster than free-living bacteria. Therefore phylogenetic differences of the chlamydial populations on a divergence scale related to free living bacteria is actually much lower (2-3%) which implies that *C. pneumoniae C. psittaci* and *C. trachomatis* are distinct species but not distinct genera (Stephens, 2002). It has also been predicted that the new families *Parachlamydiaceae, Waddliaceae* and *Simkaniaceae* are not sufficient to contain over a hundred new *Chlamydia*-like 16S sequences (Meijer and Ossewaarde, 2002).

I agree with the theory by Stephens that chlamydiae have a population structure and dynamic analogous to endosymbionts, and that the new families proposed to contain the *Chlamydia*-like species are probably not ample. I believe that the future will lead to the identification of many more chlamydia-like organisms infecting diverse hosts. The identification and analysis of these species plus the sequence analysis of more strains of known chlamydiae will allow the evolutionary history of chlamydiae to be defined more accurately, and may even lead to a reassignment of these two new genomic groups back into one common genera. During this thesis I am going to use the new taxonomic nomenclature for the family *Chlamydiaceae* (Everett *et al.*, 1999), that divides this family into two genera *Chlamydia* and *Chlamydophila*. This is because I believe the description of the separate species is probably accurate.

### 1.1.5 Pathogenesis

Members of the family *Chlamydiaceae* are responsible for causing disease in a wide range of species including humans, most domestic mammals and practically all avian species (Schachter and Caldwell, 1980). The main species associated with human infections is *C. trachomatis*. It is estimated that amongst the 38 million people who are blind in the world, 6 million are blind due to *C. trachomatis* and approximately one hundred and fifty million people, mostly children, have active infections (Thylefors *et al.*, 1995). Significant progress has been made to control trachoma in the last 50 years through trachoma control programmes that target endemic areas by donating zithromax and offering technical and financial support. These programmes have already helped to reduce disease in some endemic areas including Morocco and Tanzania, and it is the aim of the World Health Organisation (WHO) to eliminate trachoma by 2020 (Dawson and Schachter, 1999). *C. trachomatis* is also one of the most common sexually transmitted pathogens, and can be transmitted horizontally (sex) and vertically to neonates when passing through the cervix (Schachter and Caldwell, 1980). *C. trachomatis* is a major cause of infertility in both men and women. As the infection can be asymptomatic many cases are never detected (Peterson, *et al.*, 2000). *C. trachomatis* has also been associated with cervical cancer and possibly ovarian cancer (Hare *et al.*, 1982).

The development of nucleic acid amplification tests (NAATs) has revolutionised diagnosis of *Chlamydia trachomatis* infections, these tests are far more sensitive than other techniques, which include the isolation of the organism in tissue culture and using monoclonal antibodies to identify the causal organism (Schachter and Moncada, 2002). Four NAATs are currently available: a polymerase chain reaction-based test, a ligase chain reaction-based test, a transcription mediated amplification test and a strand displacement assay. However, these test are expensive and rarely performed in trachoma-endemic areas (Bird *et al.*, 2002). There have been problems in the use of these tests primarily in the general understanding of how to use them and their acceptance. But also due to the sensitivity of NAATs as problems with false positives due to contamination and false negatives due to inhibitors present in clinical samples have occurred (Schachter and Moncada, 2002).

*C. pneumoniae* is also an important human pathogen and is commonly associated as the causal agent in human upper and lower respiratory tract infections (Storz, 1988). The exact role it plays in upper respiratory tract infections has not been characterized but it is known to cause acute pneumonias and bronchitis that may lead to persistent reoccurring infections (Berdal and Scheel, 1993; Saikku, 2002). It has also been associated with chronic inflammatory diseases such as asthma and atherosclerosis (Berdal and Scheel, 1993; Cook *et al.*, 1998). A two fold greater risk factor has been identified in coronary heart disease in patients with antibodies to *C. pneumoniae* (Taylor-Robinson and Thomas, 2000). *C. pneumoniae* has been isolated from *C. trachomatis* negative cases of conjunctivitis, in a case of ulcerative skin lesions and as a causal agent for bone and joint diseases (Saikku, 2002). It may also be a risk factor for lung cancer and lymphoma (Paavonen, 2000). The diagnosis of *C. pneumoniae* has not changed in the past four years (Schachter and Moncada, 2002). There are commercially available serologic tests to measure antibodies to *C. pneumoniae* but a commercially available NAAT would significantly benefit the field (Schachter and Moncada, 2002).

*C. psittaci* isolates also infect humans and are responsible for causing a rare disease called psittacosis, which is a well known hazard to those working in duck processing plants or pet shops (Meyer, 1965; Schachter and Caldwell, 1980). *C. pecorum* has only been isolated from mammals (sheep, cattle, goats and swine) and has been identified as a causal agent of reproductive diseases, urinary tract diseases and infertility (Storz, 1988). *C. felis* infects house cats and has been associated with inflammation of the conjunctiva and respiratory problems (Studdert *et al.*, 1981). *C. caviae* infects guinea-pigs and causes ocular inflammation and eye discharge (Storz, 1988) and *C. abortus* is associated with cases of abortion and weak neonates and mainly infects ruminants (Stamp *et al.*, 1950; Griffiths *et al.*, 1995; Buxton *et al.*, 2002).

Members of the family *Parachlamydiaceae*, consists of endoparasites of free-living amoeba. Free-living amoeba, have been identified as playing a role in human disease and cause amoebic keratitis and meningoencephalitis (Horn *et al.*, 2000). There is increasing evidence that endoparasites increase the virulence of free-living amoeba and that some are causative agents for disease. This is indicated by both the presence of parachlamydia-like 16S rRNA sequences from bronchitis patients and the presence of antibodies in the blood of respiratory patients against chlamydiae-related endoparasites (Horn *et al.*, 2000).

### 1.1.6 Gene regulation

During the developmental transition of chlamydiae from the EB to the RB specific processes must be regulated. At the genetic level this involves primarily the transcriptional regulation of early and late genes (Stephens, 1993). Throughout the developmental cycle chlamydiae also have to respond to the environmental stimuli to survive, including heat shock, antibiotic attack and varying host cell environments.

Elements controlling gene regulation in chlamydiae are thought to include structural changes in DNA, for example, gene expression from ompl promoters is dependent on changes in DNA structure based on the presence of an AT-rich region upstream of the promoter (Tan and Engel, 1996). Codon usage is also hypothesised to play a role because highly expressed chlamydial genes have been shown to be located on the leading replication strand (Romero *et al.*, 2000). Quorum sensing and iron regulation, mechanisms used by other bacteria to regulate their behaviour, could also function in chlamydiae. No Quorum sensing homologs have yet been identified but proteins with Fur-like properties have been detected (Wyllie and Raulston, 2003). Interesting Fur proteins act as repressors of gene expression.

Defining chlamydial promoter regions has been troublesome as chlamydiae lack promoter homologies with other eubacteria and many genes are thought to contain multiple promoter regions (Mathews and Sriprakash, 1994; Tan and Engel, 1996; Douglas and Hatch, 1996; Shen *et al.*, 2000). The chlamydial genome encodes three eubacterial sigma factors and the core RNA polymerase subunits. It has been shown by RT PCR that each sigma factor is expressed at different stages throughout the developmental cycle (Mathews *et al.*, 1999). This could function as a control mechanism for the regulation of gene expression.

### 1.1.7 Molecular biology

Studies into the genetics of chlamydiae, has been hampered due to both the inaccessibility of the bacteria within the inclusion and the dependency of the bacteria on their host. Few chlamydial mutants exist and as yet no natural system for gene transfer has been discovered (Stephens, 1993). No host free system exists for the propagation of chlamydiae so it must be grown in mammalian tissue culture or yolk sacs. This limits the number of

phenotypic markers available to identify and study mutants. Genetic investigations have required the purification of chlamydiae from the host cell which could alter the parameters being measured. The development of a stable system for gene transfer will offer huge advances in studies into the molecular biology of chlamydiae.

In spite of these difficulties advances have been made in the field using recently developed techniques such as real-time RT-RCR and proteomics (Mathews *et al.*, 1999; Vandahl *et al.*, 2001). However the most significant advance in the past 5 years has been the genome sequencing of *C. trachomatis* and *C. pneumoniae* (Stephens *et al.*, 1998; Kalman *et al.*, 1999; Read *et al.*, 2000a). Homologies between the chlamydial ORFs and gene sequence similarity with other organisms have revealed potential mechanism used by chlamydiae for metabolism, structure, virulence and growth mechanisms. Examples of these mechanisms include ATP generating mechanisms and a type III secretory system (Hsia *et al.*, 1997; Stephens *et al.*, 1998).

## 1.2 Chlamydiaphages

### 1.2.1 Origin of chlamydiaphages

The first description of a bacteriophage infecting chlamydiae, was in 1976 when icosahedral particles of a bacteriophage were described in a crystalline lattice array in chlamydial-like organisms in molluscs (Harshbarger, 1976). The next description of a bacteriophage infecting chlamydiae was not until 1982 when 'Chlamydiaphage 1' (Chp1) was found in *C. psittaci* isolated from ducks. Chp1 was later lost, however in 1996 fourteen years after the discovery of Chp1, a bacteriophage infecting *C. caviae* was discovered 'Chlamydiaphage Guinea pig 1' (ϕCPG1) (Hsia *et al.*, 1996). A third bacteriophage infecting *C. abortus* was described in 2000 'Chlamydiaphage 2' (Chp2) (Liu *et al.*, 2000). Also in the same year a dsDNA extrachromosomal element was identified during the genome sequencing project of *C.pneumoniae* strain AR39 (Read *et al.*, 2000b). This was later shown to be the replicative form of 'Chlamydiaphage AR39' (ϕCPAR39) (Everson *et al.*, 2002).

It is not known how the chlamydiae originally became infected with bacteriophages, although it has been predicted by Moulder (Moulder, 1988) that bacteria closely related to chlamydiae, that lived in freshwater or marine habitats, first became infected with the bacteriophages. This led to a co-evolution between the bacteriophage and host as chlamydiae evolved into an intracellular parasite (Moulder, 1988). This hypothesis, however, assumes that there are physical barriers that exist between bacteriophage infection of chlamydiae (Hsia *et al.*, 2000b) and that once each bacteriophage has infected its host, prior to its commitment to the intracellular environment, no further infections could occur. However, it has been suggested, that ϕCPG1 may gain access to RBs by attaching to EBs before internalisation (Hsia *et al.*, 2000b). This hypothesis suggests that bacteriophages have had the opportunity to infect chlamydiae throughout their evolution. This is supported by the existence of a closely related bacteriophage Spiroplasma virus 4 (SpV4) (Renaudin *et al.*, 1984). It has been proposed, due to the similarities between the chlamydiaphages and ϕX174 that they have evolved from an ancestor that belonged to the *Microviridae* group of bacteriophages. The *Microviridae* group of bacteriophages are isometric bacteriophages and have a single stranded DNA genome. This family consists of

more than fifteen members that have been isolated from diverse hosts (Hayashi *et al.*, 1988). Some of the bacteriophages in this family have been extensively studied, due to their small genome and the simple structure of the mature bacteriophage. An example of this is ϕX174, which will be considered below in some detail as the prototype of this group of bacteriophages.

### 1.2.2 Introduction to chlamydiaphages

#### 1.2.2.1 Chp1

The first chlamydiaphage to be described in detail was Chp1; this bacteriophage infects an avian *C. psittaci* strain (Richmond *et al.*, 1982). The *C. psittaci* was isolated from ducks via swabs taken during an outbreak of ornithosis in workers at duck processing plants (Richmond *et al.*, 1982). The bacteriophage is icosahedral, 22nm in diameter and has a single stranded DNA genome, which is 4877 bases long with a buoyant density of 1.37g/ml in CsCl. SDS-PAGE analysis revealed three polypeptides approximately 75kDa, 30kDa and 16.5 kDa in size (Storey *et al.*, 1989a).

Preliminary analysis of the replicative cycle of Chp1 was described (Richmond *et al.*, 1982). RBs containing crystalline arrays of the bacteriophage 20 hours after the initial infection. After 40 hours some of the RBs had lysed releasing bacteriophage particles into the inclusion (Richmond *et al.*, 1982). After lysis the bacteriophage seemed to adhere to RBs, rather than EBs, and it was therefore suggested that bacteriophage infection occurs late in the developmental cycle (Richmond *et al.*, 1982).

1.2.2.2 ϕCPG1

The second chlamydiaphage to be identified was ϕCPG1; this bacteriophage was isolated from a *C. caviae* strain, which causes Guinea pig inclusion conjunctivitis (GPIC) (Hsia *et al.*, 1996). The bacteriophage is icosahedral approximately 25nm in diameter and has a single stranded DNA genome which is 4529 bases in length. SDS-PAGE analysis yielded three bands of polypeptide 62kDa, 26kDa and 13kDa in size (Hsia *et al.*, 2000a; Hsia *et al.*, 2000b).

The replication cycle of ϕCPG1 was analysed by investigating the developmental cycle at different time points by EM. The bacteriophage was observed throughout the investigation and the infection rate was very high (100%), when compared to the infection rate of Chp1 (20%) (Hsia *et al.*, 2000a). The replication cycle of the bacteriophage was investigated by TEM (Hsia *et al.*, 2000a) and a mechanism for infection and transmission was proposed. It will be considered below as an example of the replication cycle of a chlamydiaphage.

During the early stages of infection the bacteriophage did not seem to alter the normal stages of GPIC infection. In the intracellular vacuoles, bacteriophage could be observed bound to EBs. Infection occurred later after the conversion from EBs to RBs, which could be due to the bacteriophage being unable to penetrate the robust EB membranes. Division in bacterial cells infected with bacteriophage seemed to be inhibited, leading to the generation of abnormally large RBs. It was proposed that bacteriophage replication occurred within these large RBs. Due to the association seen between the bacteriophage and lysed chlamydiae membranes, it was concluded that ϕCPG1 is a lytic bacteriophage, which causes the lysis of both the inclusion and cellular membranes, resulting in the release of copious amounts of bacteriophage. The bacteriophage is then able to attach, and consequently infect, other chlamydiae. Interesting features observed during the analysis were the 'pearl necklace' structures, which could be seen when the bacteriophage particles became trapped between the outer membranes.

#### 1.2.2.3 Chp2

Chp2 which was discovered in 1998 infects *C. abortus* (Liu *et al.*, 2000). It is a featureless bacteriophage which is 25 nm in diameter and has a single stranded circular DNA genome which is 4567 bases long (Liu *et al.*, 2000).

Chp2 particles were observed attached to EBs in crude cell extracts. This observation conforms to the hypothesis for the infection process of ϕCPG1, (i.e. that the bacteriophage first binds to EBs, then is transported as a passenger into the inclusion, and infects the more fragile RBs). Interestingly, in contrast to the other two bacteriophages, Chp2 does not form paracrystalline arrays, and is evenly distributed within the chlamydiae RBs (Liu *et al.*, 2000). It also seemed to have a distinct cytopathic effect on chlamydiae RBs, which is different from that observed in the other chlamydiaphages. In contrast to the other chlamydiaphages, Chp2 infected RBs show no significant difference in shape or size from uninfected RBs. Chp2 seems to be a lytic bacteriophage as it can be identified in association with membranous material (Liu *et al.*, 2000).

#### 1.2.2.4 ϕCPAR39

Evidence for a fourth chlamydiaphage was discovered during the genome sequencing of the *C. pneumoniae* AR39, during which a 4524 nucleotide circular molecule was unexpectedly discovered (Read *et al.*, 2000a). This molecule showed homology to the *Microviridae* class of bacteriophages (Read *et al.*, 2000a). An interesting feature observed during the sequencing of *C.pneumoniae* strain AR39 was a truncated version of the bacteriophages ORF4 integrated into the *C. pneumoniae* chromosome (Read *et al.*, 2000a). It is unlikely that this gene is functional as it does not contain some of the core motifs shared by the bacteriophage Rep proteins (Read *et al.*, 2000a). The presence of this gene in the *C. pneumoniae* genome suggests that either the bacteriophage has taken a bacterial gene into its genome, the bacteria have adopted a bacteriophage gene into its genome, or that at some time in the past the bacteriophage was integrated in the bacterial genome. A related copy of this gene was also observed integrated into the CWL029 *C. pneumoniae* genome (Read *et al.*, 2000a).

The discovery of ϕCPAR39, a chlamydiaphage that can infect *C. pneumoniae* AR39 raises numerous questions about how these phages interact with chlamydiae during infection. *C. pneumoniae* is an important human pathogen and is commonly associated as the causal agent in human upper and lower respiratory tract infections. ϕCPAR39 may contribute to the virulence of *C. pneumoniae* and has being observed in clinical isolates of *C. pneumoniae* (Karunakaran *et al.*, 2002).

### 1.2.3 Genomic analysis of the chlamydiaphages

It has been predicted via computer analysis that the chlamydiaphage genome is organised into eight ORFs each of which have both an 'ATG' start codon and a region which could function as a ribosome-binding site, suggesting that they encode proteins (Storey *et al.*, 1989b) (figure 1.1). VP1, VP2, and VP3 are the three main structural proteins assigned to all chlamydiaphages by analogy to the three structural proteins of Chp1 (Storey *et al.*, 1989b). It is also very likely that ORF4 and ORF5 encode proteins as they remain largely conserved between the chlamydiaphages. Two of the smaller open reading frames present in the bacteriophage genomes (ORF7 and ORF8) are situated in similar genomic locations in all four chlamydiaphage genomes and could potentially encode small polypeptides.

All of the chlamydiaphage open reading frames are in the same orientation suggesting that the genes are transcribed as a single operon (Read *et al.*, 2000b). No promoters have been identified in the chlamydiaphage genome that show homology to the consensus promoter in *E. coli* (Storey *et al.*, 1989a). Regions in the bacteriophage genome, which could form hairpin structures, have been identified by computer analysis. In Chp1 a hairpin region was identified in a region between ORF1 and ORF2, which shows homology to the hairpin between F, and G in ϕX174 (Storey *et al.*, 1989a).

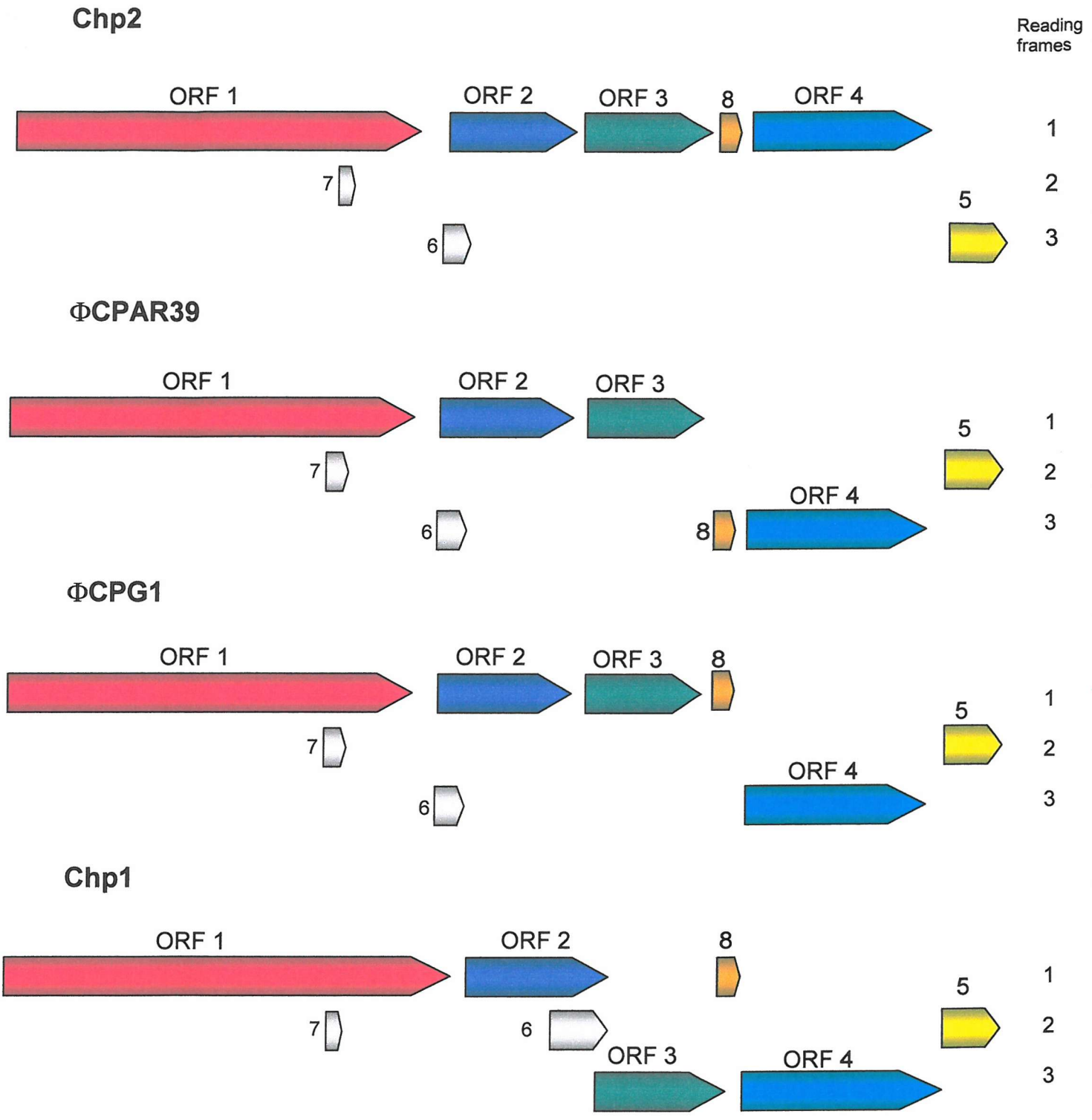



**Figure 1.1** A linear representation of the genome organisations of Chp2, ϕCPAR39, ϕCPG1 and Chp1.

However, this region is not present in the other chlamydiaphages (Read *et al.*, 2000b). Regions that could form hairpins have also been identified in the intergenic region between ORF3 and ORF4. It has been suggested, by analogy with ϕX174, that the hairpins may be important for recognition by bacteriophage replication apparatus (Read *et al.*, 2000b). The genome of Chp1 has a similar genome organisation to Chp2, ϕCPAR39 and ϕCPG1, but has diverged significantly at the nucleotide level throughout its genome.

### 1.2.4 Cell tropism of chlamydiaphages

All of the chlamydiaphages discovered to date have been identified in species belonging to the genus *Chlamydophila*. No bacteriophages have been found in the genus *Chlamydia*. This suggests that it was a member of the genus *Chlamydophila* that first became infected with these bacteriophages and that members of the *Chlamydia* genus do not have the correct receptor or rep protein these bacteriophages require for attachment or replication.

The structural proteins of the chlamydiaphages show the highest homology with the SpV4 coat proteins. In SpV4 'mushroom-like' structures can be seen on the virus surface, which correspond to an insertion loop. Similar insertion loops (IN5) can be identified by alignment of protein sequence from the chlamydiaphage VP1 (Storey *et al.*, 1989a; Liu *et al.*, 2000; Hsia *et al.*, 2000b). This region is not present in the genome of ϕX174. It has been suggested (Chipman *et. al.*, 1998) that mushroom like structures could be involved in DNA injection and might encode the chlamydiaphage receptor-binding site (Read *et al.*, 2000b).

Although VP1 of ϕCPG1 ϕCPAR39 and Chp2 remain highly conserved, the IN5 loop region of Chp2 is highly divergent to the IN5 loops of ϕCPAR39 and ϕCPG1. Another region in the VP1 of the chlamydiaphages located at residue 462-467 in Chp2 (Ins) is also highly divergent in Chp2 compared to ϕCPG1 and ϕCPAR39. This region is not present in SpV4 or ϕX174 and it has been predicted by computer analysis to be located in close contact to IN5 on the capsid protein (Read *et al.*, 2000b).

## 1.3 Other members of the Microviridae family of bacteriophages

### 1.3.1 ϕX174

ϕX174 has a single-stranded DNA genome, 5386 nucleotides long (Sanger *et al.*, 1978), which contains nine key genes A, B, C, D, E, F, G, H and J (Tessman and Tessman, 1978). The gene orientations observed in this bacteriophage are similar to those seen in chlamydiaphages and based on amino acid similarities and/or regions of homology it may be possible to attribute roles for chlamydiaphage proteins based on comparisons with the ϕX174 proteins.

The mature bacteriophage capsid of ϕX174 contains four proteins: a major capsid or coat protein (protein F); a major spike protein (protein G); a minor spike protein (protein H), and an internal protein (protein J). The functions of the proteins encoded by the other genes include; single stranded DNA synthesis (protein C); lysis (protein E); inhibition of host cell replication (protein A*); an endonuclease (protein A); capsid precursors (protein B and protein D) and a protein that stimulates the production of bacteriophage particles (protein K) (Freymeyer *et al.*, 1977; Denhardt, 1977; Pollock *et al.*, 1978a; Tessman *et al.*, 1980; Aoyama *et al.*, 1983a; Aoyama *et al.*, 1983b; Hamatake *et al.*, 1985; Aoyama and Hayashi, 1986).

Three promoters can be found in the ϕX174 genome: $P_A$ which is situated before the gene A; $P_B$ which is situated before gene B and $P_D$ which is located before the gene D. Four terminator sequences are also present in the genome: $T_J$ that is positioned after gene J; $T_F$ after gene F; $T_G$ after gene G and $T_H$, which is situated after gene H (Hayashi and Hayashi, 1970; Clements and Sinsheimer, 1975; Hayashi *et al.*, 1988).

#### 1.3.1.1 DNA replication

An outline of the infection process and a diagrammatic representation of the replicative cycle (figure 1.2) taken from 'Biology of Bacteriophage ϕX174' (Hayashi *et al.*, 1988) is given below.

Infection of the bacteriophage into the host cell occurs in three steps:

i) Attachment- the bacteriophage binds to a lipopolysaccharide receptor on the bacteria surface.
ii) Eclipse- the DNA is ejected from the bacteriophage capsid
iii) Penetration- the DNA enters the bacterial cell.

Attachment of ϕX174 to host cells is via LPS containing a terminal sugar molecule and is dependent on $Ca^{++}$ (Incardona and Selvidge, 1973). Two regions in the bacteriophage spike proteins G and H are predicted to be the bacteriophage receptor regions as host range mutants have amino acid changes in these regions (Newbold and Sinsheimer, 1970a; Newbold and Sinsheimer, 1970b). It has also been predicted that a second receptor region is required for penetration, as the mutations in bacteriophage resistant host mutants do not map strongly to the glucose-binding site (Tessman, 1965; Sinsheimer, 1968; Dowell *et al.*, 1981). The DNA is then injected into the host cell from the five-fold vertices of the bacteriophage. It has been predicted that the DNA is injected directly into the cytoplasm, as bacteriophage particles have been observed by EM embedded into the host outer membrane (Bayer and Starkey, 1972). A pilot protein (protein H), enters the host cell with the bacteriophage DNA (Jazwinski *et al.*, 1975). After the bacteriophage DNA has entered the host cell, replication takes place. Replication occurs in three stages. After injection into the host cell the ssDNA is converted in to RF DNA. This process requires 13 host proteins and is called stage I replication. The protein recognition site is a stem loop structure present between F and G, this region allows the initiation of primosome assembly (Shlomai and Kornberg, 1980). Other host cell proteins join the complex and the primosome then migrates in the 5'>3' direction along the ssDNA to produce RNA primers (Kornberg, 1980).

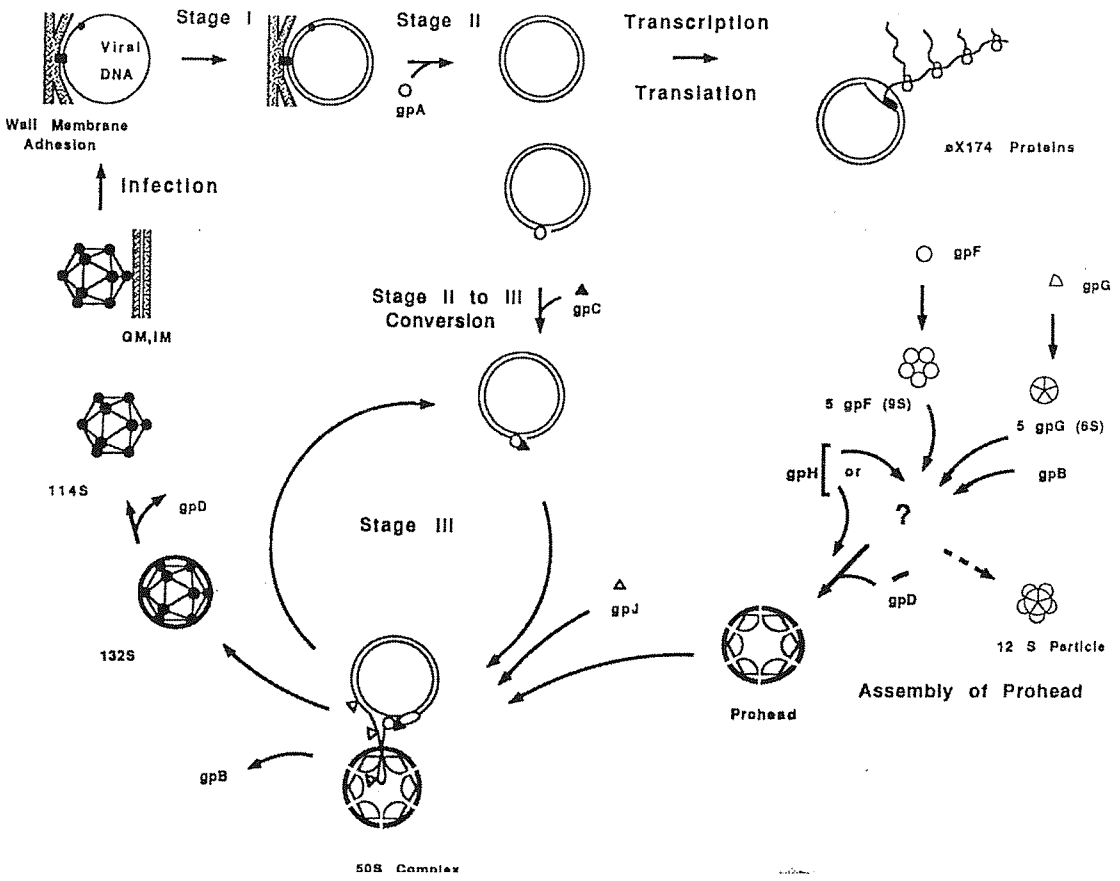



Figure 1.2 A diagrammatic representation of the replication cycle of ϕX174 taken from Biology of the Bacteriophages.

(Hayashi, Aoyama, *et al* 1988)

The RF I DNA is amplified during stage II by semi-conservative replication (Shlomai *et al.*, 1981). This process requires the 13 host proteins plus the *E.coli* rep protein and protein A. Protein A cleaves the RF DNA at the origin of replication, then anneals to the 5'end. (RF-protein A complex). The RF complex joins to the prohead and forms a structure called the 50S complex (replication assembly). (The RF-protein A complex consists of the proteins; A, B, C, D, F, G, H, J and host proteins, rep protein and DNA polymerase III holoenzyme). A bacteriophage protein encoded by gpC (protein C), directs the change from stage II replication to DNA packaging (stage III replication). Stage III replication involves concurrent synthesis and packaging of circular bacteriophage DNA from the RF-protein A complex (Eisenberg *et al.*, 1976; Hayashi, 1978). Once the RF complex has associated with the prohead, rolling circle replication occurs and the replicated viral strand is packaged into the virus prohead (Aoyama and Hayashi, 1986). The bacteriophage protein J is also incorporated into the bacteriophage prohead with the replicated DNA and protein B is eliminated. When a complete circuit of replication is completed the protein A cleaves the viral DNA at the origin and the two ends are ligated together (Fujisawa and Hayashi, 1976). The bacteriophage protein D is then removed from the complex. This involves a conformational change in the capsid proteins (Fujisawa and Hayashi, 1977) resulting in the production of a mature 114S bacteriophage particle. The steps in this process are tightly regulated to the production of mature bacteriophage.

#### 1.3.1.2 Gene expression of ϕX174

Transcription and translation occur continuously throughout the replication cycle of the bacteriophage and no switch of gene expression is observed. Host transcription and translation are inhibited but not shut off during bacteriophage infection (Linney *et al.*, 1972). Thus ϕX174 does not seem to have any elaborate mechanisms to regulate gene expression. However, a simple attenuation control mechanism exists in the form of termination events, based on the arrangement of the promoters and terminators across the bacteriophage genome (Pollock *et al.*, 1978b). For example, the promoters $P_B$ and $P_D$ are situated before the two most frequently transcribed and consequently translated genes, which are required in large quantities, while gene H (which is positioned at the distal end of a operon) is transcribed and consequently translated in the smallest amounts. This is

because the polymerase has had to bypass three terminators before it reaches gene H. The translation of overlapping regions does not follow this theory, for example, the translation of gene E is controlled by a weak ribosome-binding site (Buckley and Hayaski, 1987). Stability of RNA transcripts also functions as a control mechanism for protein expression as each mRNA decays at a characteristic rate (Hayashi *et al.*, 1976). ϕX174 is a lytic bacteriophage, gene E encodes the lysis protein, which is inserted into the host's inner membrane and inhibits the host cell enzyme involved in peptidoglycan biosynthesis (Hayashi *et al.*, 1988; Bernhardt *et al.*, 2000; Bernhardt *et al.*, 2001).

### 1.3.2 SpV4

Spiroplasma virus (SpV4) is a small isometric virus (Renaudin *et al.*, 1984), which infects a honeybee spiroplasma called *S. melliferum* strain B63 (Stamburski *et al.*, 1990). It has a circular, single-stranded DNA genome 4421 nucleotides in length (Renaudin and Bove, 1994).

The genome of SpV4 has nine open reading frames, which are situated in three reading frames and have a start codon, a termination codon and a Shine-Dalgarno ribosome binding site sequence (Stamburski *et al.*, 1990). The viron is approximately the same size as ϕX174 although it lacks the projections observed on the ϕX174 surface. The capsid protein of SpV4 (VG1) shares sequence identity with the Chp2 VP1 (33.4%) and VG2 shows sequence identity with Chp2 ORF4 protein (23.1%). VG4 shares conserved features to the Chp2 VP2 (Stamburski *et al.*, 1990).

Transcript mapping of SpV4 revealed that the virus has three promoters and one terminator (Stamburski *et al.*, 1990). Transcription of SpV4 is thought to be controlled by simple attenuation mechanism, based on the arrangement of promoters and terminators (Renaudin and Bove, 1994).

The starting point for this work was to develop a basic tool kit of reagents to allow study of chlamydiaphage gene expression. By contrast to ϕX174 and even SpV4 very little is known about the chlamydiaphages. The study of bacteriophages that infect chlamydiae are of great interest as they may provide the foundation for a genetic system for gene transfer. The purpose of this project was to investigate the basic biology of Chp2. The specific objectives of this project are laid out below.

- To express the five major open reading frames in *E. coli* to enable the production of polyclonal antisera, to allow expression of the ORFs to be studied *in vivo*.
- To investigate the transcriptional mechanisms of Chp2.
- To detect other single-stranded DNA bacteriophages, with homology to Chp2.

# Chapter 2

# Materials and Methods

## 2.1 Chemicals and Solutions

### 2.1.1 Water

General-purpose solutions were prepared using deionised water ($dH_2O$) produced by reverse osmosis using the Elga water purification system. Solutions for nucleic acid manipulations were made using Ultra High Quality water (UHQ $H_2O$), where $dH_2O$ is further purified by reverse osmosis to a resistance of 18 mega-Ohms and autoclaved at 15 psi for 15 minutes.

### 2.1.2 Chemicals and Enzymes

The chemicals were purchases as Merck grade AnalaR®, or equivalent unless stated otherwise. DNA modifying enzymes including restriction endonucleases were purchased from Promega Southampton; Life technologies Middlesex (Gibco BRL); Boehringer Corporation, London, and New England Biolabs Inc, Beverly, USA. Bacterial culture medium was purchased from Difco laboratories, Detroit, USA. All chemicals and enzymes were handled and stored as recommended by the manufacturer.

### 2.1.3 Sterilisation

Antibiotics and other heat sensitive substances were filter sterilised by filtration using a 0.22μm Millipore filter disc. Other solutions including bacterial growth media were sterilised by autoclaving at 15 psi for 15 minutes.

### 2.1.4 Plastics and Glassware

Sterile plastics were purchased from Bibby Sterilin (Staffordshire, England). Disposable polypropylene 1.5ml tubes and pipette tips were purchased from Sarstedt (Leicester, UK) then autoclaved at 18 psi for 30 minutes to reduce the possibility of contamination by nucleases. Glassware was sterilised by heating to 160°C for at least one hour.

## 2.2 Bacterial Growth Media and Solutions

### 2.2.1 Luria-Bertani (LB) Medium

Bacto-tryptone (10g/l)
Bacto-yeast extract (5g/l)
NaCl (10g/l)

The tryptone, yeast extract and NaCl were dissolved in $dH_2O$ and the pH was adjusted to 7.5 with 5M NaOH. The medium was then autoclaved and stored at 4°C.

### 2.2.2 LB Agar

1.5% (w/v) Bacto-agar (Difco) was added to LB media (2.2.1), autoclaved, cooled to 50°C and poured into plastic sterile Petri-dishes (20ml per 90mm diameter plate). The agar plates were then allowed to set at room temperature then the surface was air-dried in an incubator (45-50°C). Unused plates were stored inverted at 4°C until required for a maximum of two weeks and air-dried prior to use. LB agar plates were required for the growth and short-term maintenance of *E. coli* strains (table 2.1).

### 2.2.3 LB Soft Agar Overlays

For the growth and assay of bacteriophage M13mp18/T7, 3ml 0.65% soft LB agar was used as an overlay medium.

### 2.2.4 LB-Glycerol

For long term storage of bacterial suspensions at -70°C LB-glycerol was used. LB media (2.2.1) was prepared containing 10% (w/v) glycerol and was sterilised by autoclaving in 10ml aliquots. The medium was stored indefinitely at room temperature.

### 2.2.5 Antibiotics

Ampicillin (Sodium salt, Sigma) was dissolved in UHQ $H_2O$ to 50mg/ml, kanamycin (Sulfate salt, Sigma) was dissolved in UHQ $H_2O$ to 10mg/ml, and chloramphenicol (Crystalline, sigma) was dissolved in 100% EtOH at 34mgs/ml, then filter sterilised through a 0.2-micron filter (Gelman sciences, Northampton), aliquoted then stored at -20°C.

### 2.2.6 LB agar and Medium with Antibiotics

Medium containing ampicillin was used to maintain ampicillin-resistant *E.coli*. Ampicillin (2.2.5) was added either to molten LB agar (2.2.2) or to LB medium (2.2.1), at final concentrations of 50μg/ml immediately prior to use. All agar plates were poured, dried and stored as described previously (2.2.2).

### 2.2.7 2YT Medium

For the growth of *E.coli* JM109 infected with M13mp18/T7 helper phage 2YT medium was used. The medium contained:
Bacto-tryptone (16g/l)
Bacto-yeast extract (10g/l)
NaCl (5g/l)

Prior to autoclaving and storage in 50ml aliquots, the pH was adjusted to 7 with NaOH.

### 2.2.8 Minimal E Agar

The defined minimal E salt solution described by Vogel and Bonner (1956) (Vogel and Bonner, 1956) was prepared as a 50× stock containing citric acid (100g/l), $K_2HPO_4$ (500g/l), $MgSO_4.7H_2O$ (10g/l), and $NH_4NaHPO_4$ (175g/l) dissolved in $dH_2O$. Aliquots (20ml) were sterilised by autoclaving and stored at room temperature. Bacto-agar [1.5% w/v)] and 0.2% (w/v) glucose were added to $dH_2O$, autoclaved and allowed to cool to 50°C. Sterile minimal E salts were added to the glucose-agar to make a 1× final concentration.

### 2.2.9 Xgal Agar Plates

Ampicillin agar plates (2.2.6) containing 2% BCIG {5-brom-4-chloro-3-indolyl-β-D-galactoside} were utilised in the detection of promoter regions functional in *E. coli*.

## 2.3 Bacteria, Bacteriophage, and plasmids

The plasmid vectors and strains of *E.coli* used in this study are listed in tables 2.1 and 2.2.

### 2.3.1 Storage and Growth of *E.coli*

Most strains, except *E.coli* JM101, were kept on LB agar as working stocks at 4°C and passaged at intervals of 3-4 weeks. Working stocks of *E.coli* JM101 were maintained on minimal E agar plates containing 1μg/ml thiamine hydrochloride. Any *E.coli* strains containing plasmid encoded resistance were maintained on LB agar plates containing the required antibiotic, at 4°C.

Liquid cultures of *E.coli* were grown at 37°C at 180rpm on a shaker for 16 hours in LB broth (2.2.1). For long-term maintenance, a suspension of each strain of *E.coli* was stored in LB medium containing 10% glycerol at -70°C.

Some methods needed a static overnight starter culture of the required strain of *E.coli*. To prepare the starter culture, 10ml of LB broth was inoculated with the required strain of *E.coli* and incubation at 37°C overnight.

### 2.3.2 Preparation of Competent *E.coli*

The transformation efficiency of *E.coli,* is dependent on the preservation of all materials and reagents at 0°C. A static 10ml starter culture in LB broth of the recipient *E.coli* strain was first grown overnight at 37°C. Then 1ml of the initial culture was used to inoculate 25ml of LB broth and incubated at 37°C on a shaker at 180rpm, until the $A_{550}$ value was between 0.4-0.5 (approximately 2 hours). The culture was then left on ice for 20 minutes and all following procedures performed at 4°C.

To pellet the cells the recipient culture was then centrifuged at 2,000×g for 10 minutes and the supernatant discarded. The bacterial pellet was immediately re-suspended in 10mls of ice-cold sterile 0.1M $MgCl_2$ and centrifuged as above. The supernatant was

discarded and the bacterial pellet re-suspended in 1ml of ice-cold sterile 0.1M $CaCl_2$. The *E.coli* cells were then kept on ice for a minimum of 1 hour, until use as competent cells.

### 2.3.3 Transformation of Competent *E.coli* cells

For each transformation, competent *E.coli* were dispensed into 100µl aliquots in ice-cold eppendorf tubes. To each aliquot of competent cells ice-cold DNA was added and the tubes were left on ice with occasional shaking for 30 minutes. To allow the uptake of DNA, the competent *E.coli* cells were heat shocked for 2 minutes at 42°C. The cells were left for a further 30 minutes to allow the cells to recover on ice, prior to spreading on LB agar plates containing the relevant selective antibiotic (2.2.7), then incubation at 37°C overnight.

**Table 2.1** ***Escherichia coli*** **Strains**

| Strain | Relevant Genotype | Reference |
|---|---|---|
| JM101 | *sup*E, *thi*, Δ(*lac-pro*AB), [ F', *tra*D36, *pro*AB,$^{+}$ *lac*I$^{q}$ lacZΔM15] | (Yanisch-Perron, Vieira, and Messing, 1985) |
| DH5α | *F*$^{-}$, ϕ80 *lac*ZΔM15, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd* R17 ($r_K^-$, $m_K^+$), *Sup*E44, *rel*A1, *deo*R, Δ(lacZYA-*arg*F)U169 | (Hanahan, 1983) |
| TOP 10 | *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC), ϕ80 *lac*ΔM15, Δ *lac*X74, *deo*R, *rec*A1 *ara*D139 Δ(ara leu)7697, *gal*U, *gal*K, λ-*rps*L, *end*A1, *nup*G | (Grant *et al.*, 1990) |
| BL21(DE3)p lysS | *F*$^{-}$, *ompT*, *hsdS*$_B$($r^-_B$,$m^-_B$), *gal*, *dcm*, *met*, *(DE3)*, *pLysS* | (Studier, 1991) |
| JM109 | *end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17($r_k$-,$m_k^+$), *rel*A1, *sup*E44, Δ(*lac-pro*AB) | (Yanisch-Perron, Vieira, and Messing, 1985) |

**Table 2.2 Plasmid and Bacteriophage Vectors**

| Plasmid | Relevant features | Reference |
|---|---|---|
| M13mp18/ T7 | Helper phage that carries the T7 RNA Polymerase gene driven by *E.coli lac* promoter induced by IPTG | Invitrogen |
| pRSET ABC | Specialised cloning vectors T7 polymerase promoter, Ampicillin resistance Six histadine metal binding domain | Invitrogen |
| PGEX-4T1 | Fusion protein expression vector ptac promoter Ampicillin resistance Fused protein to GST | Pharmacia Biotech |
| PMAL-c3 | Fusion protein expression vector ptac promoter Ampicillin resistance Fused protein to a MBP | New England Biolabs |
| pSP73 | Specialised cloning vector T7 and SP6 RNA polymerase promoters Ampicillin resistance | Promega |

## 2.4 Isolation and Purification of Nucleic acids

### 2.4.1 Ethanol Precipitation of DNA and RNA

Sodium acetate was added at a final concentration of 0.3M to aqueous RNA or DNA solutions by the addition of 0.1 volumes of a 3M stock (adjusted to pH 5.0 with glacial acetic acid). Two volumes of absolute ethanol (at -20°C) were added and the RNA or DNA precipitated for between 20 minutes and 16 hours (overnight) at -20°C. The precipitated nucleic acid was recovered by centrifugation (20 minutes, microfuge) and washed with 80% ethanol to remove any excess salts. The pellet was vacuum dried for 2-3 minutes and re-suspended in the necessary volume of sterile UHQ $H_2O$.

### 2.4.2 Isolation of Plasmid DNA

#### 2.4.2.1 Qiagen Midiprep

A Midiprep kit (Qiagen) was use for large-scale purification of plasmid DNA. An overnight 25ml culture of bacteria in LB (2.2.1) containing the appropriate antibiotic was pelleted by centrifugation for 20 minutes (centrifuge, 2,000xg). The bacterial pellet was resuspended in 4mls of supplied buffer P1 (50mM Tris.Cl, pH 8.0; 10mM EDTA; 100μg/ml RNaseA), containing RNase A. Following lysis and neutralisation (4mls buffer P2 (200mM NaOH, 1% SDS), 5 minutes, and 4mls pre-chilled buffer P3 (3.0M potassium acetate, pH 5.5) on ice for 15 minutes) cellular debris and chromosomal DNA were removed by centrifugation (2,000×g) at 4°C for 30 minutes. To make sure all particulate matter had been removed from the supernatant, it was centrifuged again for a further 15 minutes.

A QIAGEN-tip 100 (supplied) was used to isolate the plasmid DNA. The tip was first equilibrated by the addition of 4mls of QBT buffer (750 mM NaCl; 50mM MOPS, pH7.0; 15% isopropanol; 0.15% Triton X-100) and emptied by gravity flow. Next the bacterial supernatant was applied to the QIAGEN-tip and allowed to enter the resin by gravity flow, the flow through was kept for later examination. The resin in the QIAGEN-tip was next washed by the addition of 2 × 10mls buffer QC (1.0 M NaCl;

50mM MOPS, pH 7.0; 15% isopropanol), and the DNA was eluted from the resin with 5mls of the supplied buffer QF (1.25 M NaCl; 50mM Tris.Cl, pH8.5; 15% isopropanol). The eluted DNA was then precipitated with 0.7 volumes of room temperature isopropanol then centrifuged (2,000×g) at 4°C for 30 minutes to pellet the DNA. The supernatant was then discarded and 2mls of 70% ethanol was added to wash the pellet. To pellet the DNA, the solution was centrifuged (2,000×g) for 10 minutes. The ethanol was then removed and the DNA pellet vacuum-dried for 5 minutes to remove any remaining ethanol then re-suspended in 100µl UHQ $H_2O$. The DNA was analysed by agarose gel electrophoresis (2.5.2) to determine its concentration.

### 2.4.2.2 Qiagen Spin Minipreps

A miniprep kit (Qiagen) was used for small-scale plasmid DNA purification. 4.5mls of a 10 ml overnight culture of bacteria in LB containing the appropriate antibiotic was pelleted by centrifugation (microfuge, 5 minutes, 17, 900xg) and the cell pellet re-suspended in 250µl of the resuspension solution P1 (50mM Tris.Cl, pH 8.0; 10mM EDTA; 100µg/ml RNaseA). Following lysis and neutralisation, (250µl cell lysis solution P2 (200mM NaOH, 1% SDS) then 350µl neutralisation solution N3 (3.0M potassium acetate, pH 5.5)) chromosomal DNA and cellular debris were eliminated by centrifugation (10 minutes, 17,900xg). To isolate the plasmid DNA the supernatant was applied to the minicolumn then centrifuged (1 minute, microfuge, 17,900xg). The DNA was washed by addition of 750µl buffer PE (1.0 M NaCl; 50mM MOPS, pH 7.0; 15% isopropanol) to the column and centrifugation (1 minute, 17,900xg). Any remaining wash buffer was removed by a further 1 minute microfugation (17,900xg). The bound plasmid DNA was then eluted in 50µl buffer UHQ (stand room temperature 1 minute, microfuge 1 minute, 17,900xg). Yields of 1-10µg could be achieved.

### 2.4.2.3 **Purification of Plasmid DNA by STET** (Holmes and Quigley, 1981)

The STET (Sucrose, Triton, EDTA, Tris) method (Holmes and Quigley, 1981) is a rapid technique used for screening large numbers of *E.coli* transformants. STET buffer (100ml) contains 5% (w/v) Triton X-100, 50mM EDTA, 50mM Tris-HCl pH 8.0, 8% (w/v) sucrose.

An inoculating loop of *E.coli* from a plate culture of a single transformant, was vigorously re-suspended in 100µl STET buffer to obtain a suspension. To lyse the cells 8µl lysozyme solution (10mg/ml in UHQ $H_2O$) was added and incubated for 5 minutes at room temperature. Cell lysis was achieved by boiling the sample for 40 seconds. Cell debris was removed (microfuge, 10 minutes) and the supernatant, containing the plasmid DNA, recovered and precipitated with an equal volume of isopropanol (-20°C for 1 hour).

The DNA was recovered (microfuge, 15 minutes), then washed with 80% (v/v) ethanol, vacuum dried for 2-3 minutes and re-suspended in 15µl UHQ $H_2O$. 5µl of each sample was sufficient for examination by agarose gel electrophoresis (2.5.2) or for restriction endonuclease digestion (2.6.1.1).

### 2.4.3 Isolation of Total RNA

A SV total RNA isolation system (Promega) was used to extract total RNA from cells. 525µl of SV lysis buffer (Promega) was added to a 25cm$^2$ flask of cells to cause cell lysis. 1100µl of SV RNA dilution buffer was then added, mixed by inversion and incubated at 70°C for 3 minutes. Cell debris was remove by centrifugation (14 300 x g for 10 minutes), and the cleared lysate was transferred to a clean microfuge tube then 200µl of 95% ethanol/500µl supernatant was added. To isolate the RNA the supernatant was applied to a spin column then centrifuged (14 300 xg for 1 minute). The RNA was then washed, by the addition of 600µl SV RNA wash solution to the column and centrifugation (1 minute, 14 300 xg).

A DNase incubation step was then carried out by applying 40µl of yellow core buffer, 5µl of 0.09M $MnCl_2$ and 5µl DNase I then incubating for 15 minutes at room temperature. The incubation was stopped by the addition of 200µl SV DNA stopping solution to the column and centrifugation (1minute, 14 300 xg). The RNA was then washed twice by addition of 600µl SV RNA wash solution to the column and centrifugation (1 minute, 14 300 xg), and then the addition of 250µl SV wash solution to the column and centrifugation (1 minute, 14 300 xg).

The bound plasmid RNA was eluted in 100µl of nuclease free water (stand room temperature 1 minute, microfuge 1 minute, 14 300 xg). The eluted RNA was stored at -70°C.

## 2.5 Analysis of Nucleic acids

### 2.5.1 Materials

#### 2.5.1.1 TAE Buffer (tris-acetate EDTA buffer)

A 1× working dilution was used prepared from a 50x pH 8.3 stock of TAE (2M Tris-acetate, 50mM EDTA; pH at 25°C: 8.2–8.4, Promega) for horizontal submerged agarose gels as the electrophoresis buffer.

#### 2.5.1.2 Agarose Gel Loading Buffer

Gel loading buffer was used for DNA/RNA samples analysed by horizontal submerged agarose gel electrophoresis. The loading buffer contained two dyes, xylene cyanol and bromophenol blue, these acted as markers to monitor the migration rate of the DNA through the gel. A 10× stock of the loading buffer was prepared in 20ml 1× TAE, which comprised of ficoll molecular weight 400,000 Sigma (5g), bromophenol blue (50mg) and xylene cyanol (50mg). Before use, the gel-loading buffer was heated to 70°C to ease sterilisation by filtration, and was diluted by addition to a DNA sample to give a final 1× concentration.

### 2.5.2 Agarose Gel Electrophoresis of DNA

Agarose gel electrophoresis was used for the characterisation, quantitation and separation of DNA fragments, and to enable the purification of specific fragments of DNA. Different concentrations of agarose were used to analyse different size fragments of DNA. High strength ultrapure analytical grade agarose (Biorad) was suitable for most purposes, Nusieve agarose (FMC Bioproducts, Flowgen) was used to analyse fragments of DNA smaller than 200 bases (table 2.3).

**Table 2.3 Agarose Gel Concentrations**

| Agarose | Ranges of DNA Analysed | |
|---|---|---|
| | lower range | Upper range |
| 0.7% | 1 kb | 10 kb |
| 0.9% | 500bp | 7 kb |
| 1.5% | 400bp | 6 kb |
| 2.0% | 200bp | 4 kb |
| 3-4% nusieve | 100bp | 3 kb |

Agarose gels were cast in perspex trays (Bio-Rad) of various sizes. The two ends of the trays were sealed with tape and the tray was placed on a level surface. The sample wells were formed in the gel using a plastic comb, this was placed in the tray approximately 2mm above the base of the tray and 1cm from the top of the gel. The agarose was melted in 1× TAE by heating to 100°C in a microwave oven. The sizes of gels used, and the volumes of melted agarose were:

6.5cm×10cm - 40ml,

15cm×10cm - 100ml,

15cm×20cm - 150ml.

The melted agarose/buffer was cooled to approximately ~50°C, then poured into a gel mold and allowed to set for at least 1 hour before use. Once the gel had set, the tape

and comb were removed carefully and the gel was submerged in 1× TAE buffer. DNA samples containing 1× gel-loading buffer were then loaded into the wells under the buffer. An aliquot (1μl) of a 1kb plus DNA ladder (Gibco BRL) was used as a marker for both the quantity and size of the analysed DNA. Electrophoresis was carried out for approximately 1 hour until the bromophenol blue dye had migrated 1cm from the end of the gel at 90-100v.

The gel was then removed and stained in 1μg/ml ethidium bromide for 30 minutes. DNA was visualised on a 2UV™ transilluminator, BioDoc-It™ System (UVP).

## 2.6 Manipulation of Nucleic Acids

### 2.6.1 DNA Cutting and Modification

#### 2.6.1.1 Restriction Enzyme Digests

Enzymes and buffers (10× concentrates) were stored at -20°C. Digests typically contained 1-2μg of DNA and 5-10 units of enzyme per μg DNA in a total reaction volume of 10-50μl. Reactions were incubated at 37°C (unless otherwise indicated by the manufacturer) for 2-4 hours. Attention was taken to ensure that the final concentration of glycerol in the reaction did not exceed 5% (v/v), to prevent the reduction of specificity of the enzyme (star activity).

#### 2.6.1.2 Dephosphorylation of DNA

Before cloning linearised vector DNA was dephosphorylated. Calf intestinal alkaline phosphatase (CIAP; Promega) was added to the digested DNA (1U CIAP/μg DNA) and incubated at 37°C for 15 minutes, followed by 15 minutes at 55°C and 5 minutes at 72°C. The 55°C incubation step caused the ends of the DNA duplex to separate slightly allowing the enzyme access to the 5′ terminal phosphate groups and the 72°C incubation step inactivated the enzyme. The dephosphorylated DNA was purification with Qiagen PCR preps (2.7.3.1) before cloning.

### 2.6.1.3 Ligation of DNA

Covalent joining of molecules of DNA was performed using T4 DNA ligase. A ligation reaction contained 10ng vector and a three to five molar excess of insert DNA. The other components of the reaction were 1µl acetylated BSA (1mg/ml), 1µl 10× ligase buffer (700mM Tris-HCl pH7.5; 70mM $MgCl_2$; 10mM DTT; 1mM ATP) and 1µl T4 DNA ligase. The ligations were incubated at either 14°C (cohesive ends) or 4°C (blunt ends) overnight.

## 2.7 Amplification of DNA by Polymerase Chain Reaction (PCR)

Bio-X-Act™ DNA Polymerase is a mixture of polymerases that possess 5′-3′ DNA polymerase activity and 3′- 5′ proof-reading activity which prevents mis-incorporations during primer extension, and therefore increases the fidelity of the product. Bio-X-Act was used to amplify all DNA used for sequencing and cloning.

### 2.7.1 Reaction Conditions

DNA samples were amplified using a Perkin Elmer Cetus 9600 thermal cycler or a PTC-255 Peltier Thermal Cycler (MJ Technologies, USA). The template DNA was amplified in cycles which consisted of three steps: 1) template DNA denaturation which was usually 20 seconds at 94°C, 2) primer annealing, which was based on $T_m$°C -10°C for 20 seconds [where $T_m$°C=4(G+C)+2(A+T)], and 3) primer extension/polymerisation which was 72°C for all reactions, for 15 seconds per 150 bases amplified. Each step within a cycle was carried out at a specific temperature for a specified period of time with a rapid thermal ramp between temperatures. Each cycle was repeated for 25-35 times. After the completion of the thermal cycling, all samples were held at 10°C until removed from the thermal cycler.

In general a 50µl PCR sample contained: 50mM KCl, 10mM Tris-HCl pH 8.8, 1.5mM $MgCl_2$, 0.1% (v/v) Triton X-100 (all supplied as a 10× buffer), 0.2mM of each deoxynucleotide, 250ng each primer, 0.5-5µl template DNA and 1.25U Bio-x-Act polymerase.

### 2.7.2 Primer Design

Oligonucleotides were bought re-suspended and desalted (Cruachem). The sequence designed based on the following criteria: 1) that they should not contain any palindrome sequences, 2) the forward and reverse primers should not form any 3′-5′ concatenates, 3) the primers should be 17- 40 bases in length, 4) the termini of the primers should be not be complementary and 5) the forwards and reverse primers should have similar $T_m$ values. PCR primers containing restriction sites tails were used to introduce enzyme sites into certain amplification products to facilitate cloning.

### 2.7.3 Purification of Amplified Products

Before sequencing or cloning, the DNA fragments were purified to remove contaminating nucleotides and primers. Purification was accomplished using the Qiaquick gel extraction kit (Qiagen) or using the Qiaquick PCR purification kit (Qiagen).

#### 2.7.3.1 Qiaquick PCR Purification Kit

The DNA to be purified was mixed with 5 volumes of buffer PB (supplied) then added to a Qiaquick spin column. The product was bound to the column by centrifugation (1 minute), washed by addition of 750µl buffer PE (supplied) and centrifugation for 1 minute. The column was dried by a further 1-minute centrifugation and vacuum dessication for 1 minute. The purified product was eluted in 30µl UHQ by centrifugation (stand 1 minute, microfuge 1 minute).

#### 2.7.3.2 Qiaquick Gel Extraction Kit

The DNA was cut from the agarose gel using a clean scalpel blade. Three volumes of buffer QG (supplied) was added and incubated at 50°C for 10 minutes until the gel slice had completely dissolved. One gel volume of isopropanol was then added to the sample and mixed, the sample was then loaded into a Qiaquick spin column and the DNA bound by centrifugation for 1 minute. Excess agarose was removed by addition of 0.5mls buffer QG and centrifugation (1 minute, microfuge). The column was

washed by addition of 750µl buffer PE (supplied) and left for 5 minutes before centrifugation (1 minute, microfuge). The column was dried by a further 1-minute centrifugation and vacuum dessication for 1 minute. The purified product was eluted in 30µl UHQ by centrifugation (stand 1 minute, microfuge 1 minute).

### 2.7.4 Two-Step Real time PCR

#### 2.7.4.1 cDNA Production

cDNA samples were prepared using a Perkin Elmer Cetus 9600 thermal cycler or a PTC-255 Peltier Thermal Cycler (MJ Technologies, USA). Immediately after the total RNA had been extracted from the cells, cDNA was prepared using a TaqMan$^{R}$ Gold RT-PCR Kit (Applied Biosystems). A two-step RT PCR was carried out, 20µl reactions were set up for each tube, containing 2µl 10X Taqman RT Buffer, 4.4µl 25mM Magnesium chloride, 4µl deoxyNTPs, 0.4µl RNase Inhibitor, 0.5µl MultiScribe Reverse Transcriptase, and 1µl Random hexamers. PCR cycling parameters involved three steps, first a 5 minutes incubation step at 25°C to allow all the hexamer primers to be extended by the reverse transcriptase, a 30 minutes incubation step at 48°C for reverse transcription and then a 5 minutes incubation step at 95°C to denature the reverse transcriptase. The products were stored at -20°C until required.

#### 2.7.4.2 Real Time PCR

Real-time PCR utilises the 5' nuclease activity of AmpliTaq Gold DNA Polymerase to cleave a TaqMan probe during PCR. The probe contains a quencher dye at the 3' end and a reporter dye at the 5' end. During PCR reactions cleavage of the probe separates the quencher and reporter dye which results in an increase in reporter fluorescence. The accumulation of PCR products is monitored directly by detecting the increase in fluorescence of the reporter dye. A 50µl reaction was carried out for the second step of the two step process containing, 5µl 10X Taqman Buffer A, 11µl 25mM MgCl, 1.5µl of 10mM deoxyATP/CTP/GTP/UTP, 1µl 5µM GAPDH probe and 1µl of 10µM forward and reverse primers, 0.5µl AmpErase UNG 0.25µl Ampli

Taq Gold DNA polymerase (5.0U/μl) and 21.25μl of RNase-free water were added to 5μl of cDNA. PCR parameters were first a 2-minute UNG incubation at 50°C (to degrade dUTP, thus prevent contamination from previous real-time PCR reactions), a 10-minute AmpliTaq Gold Activation at 95°C, then a Denaturation step consisting of 15 cycles of 95°C for 15 seconds followed by an Anneal/Extend step 60°C for 1 minute.

## 2.8 Automated DNA Sequencing

Automated DNA sequencing was performed using an ABI PRISM Model 377 DNA Sequencing System with an ABI PRISM BigDye Terminator Cycle sequencing Ready Reaction Kit. This technique relies on four fluorescent dye-labelled di-deoxynucleotide terminators and employs a thermostable DNA polymerase (AmpliTaq DNA Polymerase, ABI PRISM) to determine DNA sequences by the chain termination method (Sanger, Nicklen, and Coulson, 1977). The Sanger method of DNA sequencing can be summarised in three steps: i) annealing of the primer DNA to the template DNA, ii) labelling of the DNA during polymerisation, and iii) termination of chain extension. Each sequence determination is carried out as a single reaction, containing all four deoxynucleoside triphosphates (dNTP). The ddNTP lacks the necessary 3′-OH group required for chain elongation, thus the oligonucleotide chain is terminated selectively at A, C, G, or T depending on the particular ddNTP incorporated. The resulting fragments have a common origin but are terminated at different nucleotides along the extension product and therefore can be separated according to size by high-resolution denaturing gel electrophoresis. In automated sequencing, a fluorescent dye label specific to A, T, C, or G is incorporated into the DNA along with the terminating base. A laser scans across a region of the gel during the electrophoresis process, detecting the fluorescently labelled DNA passing through this region.

### 2.8.1 Materials

To avoid any extraneous fluorescence, reagents of the highest quality and UHQ $H_2O$ were used. All glassware was cleaned carefully by treatment with concentrated hydrochloric acid then rinsed with UHQ $H_2O$.

A ready to use stock solution of 4.8% Gene Page Plus was purchased from Amersco, containing 6M urea and 1x TBE buffer which was stored at 4°C. A 5X liquid concentrate of TBE buffer was purchased (Amersco) and stored at room temperature

### 2.8.2 Sequencing Reactions

The sequencing reactions were performed using a Model 9600 thermocycler (Perkin Elmer Cetus) or a PTC-255 Peltier Thermal Cycler (MJ Technologies). The DNA sequencing kit (ABI PRISM) was stored at -70°C until required, then at -20°C between uses.

High quality template DNA is required for sequencing. The table below shows the amount of template DNA used in a cycle sequencing reaction.

**Table 2.4 The Amount of Template used in a Cycle Sequencing Reaction**

| Template | Quantity |
|---|---|
| **PCR product:** | |
| 100-200 bp | 1-3 ng |
| 200-500 bp | 3-10 ng |
| 500-1000 bp | 5-20 ng |
| 1000-2000 bp | 10-40 ng |
| >2000 bp | 40-1000 ng |
| **Double-stranded** | 200-500 ng |

In general the higher quantities of DNA template used gave higher signal intensities. The reaction was mixed in a 0.6ml double-snap-cap microcentrifuge tube with 8µl of the supplied premix (Big Dye Terminator RR Mix, ABI PRISM) and 3.2pmol of primer in a total volume of 20µl. The primers were generally 18-25bp oligonucleotides. The tubes were placed in the thermal cycler preheated to 96°C to reduce false priming. Denaturation was at 96°C for 10 seconds, annealing at $T_m$-10°C for 5 seconds and extension at 60°C for 4 minutes, for 25 cycles. The temperature of the thermocycler was altered by a rapid thermal ramp. The reactions were held at 10°C after thermo-cycling was complete. The extension products were purified by, ethanol precipitation, rinsing with 70% ethanol and vacuum dessication for 1 minute. The DNA pellet could be stored at -20°C for up to four weeks.

### 2.8.3 Gel Preparation

The Gene Page Plus gel was cast between two glass plates separated by spacers of 0.2mm constant thickness. To avoid fluorescent contamination the glass plates were washed thoroughly with the detergent Alconox (Merck Ltd), rinsed with hot water then methanol (Hypersolv) and air-dried. The plates and spacers were assembled and secured in the supplied gel cassette, then placed horizontally on a level surface to form a mold for the gel.

150µl of 10% APS solution and 15µl of TEMED solution were added to 25ml of Gene Page Plus solution to initiate polymerisation. A single flat-edged comb (PE Biosystems) was inserted into the top of the gel. The gel was then allowed to polymerise for 2 hours.

### 2.8.4 Gel Electrophoresis

Once the gel had set, the comb was carefully withdrawn and the laser-scanning region cleaned by wiping with methanol (Hypersolv) and 3mm chromatography paper (Whatman International). The plates were scanned to ensure the absence of any background fluorescence. Lower buffer reservoir on which the bottom edge of the plates rests was placed in the electrophoresis chamber. The plates were secured and the upper buffer chamber placed at the top of the plates. A 36 well disposable sharks

tooth comb (PE Biosystems) was then inserted carefully between the plates so that the teeth were inserted to a depth of 0.2 mm and the buffer reservoirs filled with 1× TBE.

The sample pellets were re-suspended in 4µl of loading buffer (5 deionised formamide: 1 blue dextran). The samples were denatured by heating at 95°C for two minutes and 2µl loaded into alternate wells that were pre-washed with 1×TBE. Electrophoresis was carried out for 12 hours at 2500 volts, 40 mAmps, 30 Watts and at 40°C. The sequence data were automatically collected and analysed by computer (Applied PRISM and DNA Star software).

### 2.8.5 Computer Analyses

DNA sequence editing, restriction site analyses, translations, open reading frame locations, protein molecular weight predictions, DNA and protein sequence alignments and amino acid compositions were performed using DNAstar LaserGene software.

## 2.9 Protein expression systems

### 2.9.1 XPRESS System™ for Protein Expression in *E.coli*

The pRSET vectors (Invitrogen) are designed for the high-level protein expression and purification of cloned genes in *E.coli* cells. DNA fragments were cloned into the pRSET vectors under the control of the T7 promoter and positioned in frame with a sequence that encoded an N-terminal fusion peptide. This sequence included an ATG translation initiation codon, a sequence encoding a series of six histidine residues that functioned as a metal binding domain in the translated protein, a transcript stabilising sequence from gene 10 of phage T7 and the enterokinase cleavage recognition sequence.

The metal binding domain of the fusion peptide allowed simple one step purification of recombinant proteins by Immobilised Metal Affinity Chromatography. The enterokinase cleavage recognition site on the fusion peptide between the metal

binding domain and the recombinant protein allowed for removal of this N-terminal fusion peptide from the purified recombinant protein.

To propagate the pRSET expression plasmids, ligation mixes of the pRSET constructions were transformed into JM101 cells (table 2.1). Transformants could be selected with ampicillin plates, and the correct constructions were confirmed by restriction enzyme analysis and sequencing.

#### 2.9.1.1 Cloning into the Expression Vector

All inserted fragments were cloned into the pRSET multiple cloning site (MCS), this has ten unique restriction enzyme sites positioned down stream of the enterokinase cleavage recognition sequence. pRSET A, B and C differed only in the spacing between the sequence that coded for the N-terminal peptide and the MCS. This enables DNA fragments to be cloned in frame with the first ATG of the pRSET vector.

#### 2.9.1.2 M13mp18/T7 Helper phage

T7 RNA polymerase was produced using M13mp18/T7 helper phage (Invitrogen). This carried out high-level expression of recombinant protein from the pRSET vector. The vector containing the inserted DNA was propagated into an *E.coli* strain, which didn't produce T7 polymerase. Then when expression was required, the *E.coli* cells were infected with an M13 phage, which contained the T7 RNA polymerase gene. The *E.coli* lac promoter drives this phage gene. IPTG was also required in expressions, as it prevents the repression of the *lac* promoter and hence induced T7 RNA polymerase expression from the M13mp18/T7 phage.

#### 2.9.1.3 Determination of the Titre of M13mp18/T7 Phage

LB buffer was used to prepare ten-fold serial dilutions ($10^{-7}$ to $10^{-14}$) of the M13mp18/T7 phage stock in eight sterile glass bijoux. Eight minimal E plates (2.2.8) were also labelled with the dilutions $10^{-7}$ to $10^{-14}$. 100µl of each dilution and 100µl of JM109 (table 2.1) "plating cells" (overnight LB culture of JM109 at 37°C) were

added to the relevant glass bijoux containing 3ml of molten LB soft-agar (2.2.3) at 55°C. Each bijoux was then gently mixed then the contents poured swiftly onto the corresponding minimal E plate. The plate was swirled gently to ensure that the soft-agar was equally distributed over the whole plate. To allow soft agar to set the plates were incubated at room temperature for five minutes, then inverted and incubated at 37°C for 12-16 hours.

A plate containing a sufficient number of plaques (20-200), was then selected and the number of plaque forming units (pfu) per millilitre of stock calculated according to following formula:

**pfu/ml = number of plaques × 1/dilution factor × 1/vol plated (ml)**

#### 2.9.1.4 Expression of Recombinant Protein

25ml of LB containing ampicillin was inoculated with an overnight culture of a single colony known to contain the recombinant plasmid. The culture was then incubated at 37°C with shaking for approximately 2 hours until the $O.D._{600}$ was 0.3. Then IPTG (final concentration 1 mM) was added and the culture was incubated for a further hour at 37°C. IPTG enters the cells then binds to the *lac* repressor, which induces the *lac* promoter. The cells were then infected with the M13/T7 phage at a ratio of 5pfu/cell the amount of phage to be added was determined by taking the $OD_{600}$, and assuming that $OD_{600}$ = 1.0 is equivalent to $10^9$ cells.ml. The M13/T7 phage is essential for expression as it contains the T7 RNA polymerase gene. The culture was then incubated for a further 4 hours at 37°C. At the end of the expression period 1ml of the culture was harvested centrifuged and the pellet re-suspended in 200µl SDS-sample buffer. A 20µl sample was analysed by a SDS-PAGE.

For a large scale expression a single recombinant *E.coli* colony was inoculated into 2ml of LB Amp and the culture grown overnight at 37°C with shaking. 0.3 ml of this overnight culture was then used to inoculate 50ml LB Amp, which was then incubated at 37°C until the $OD_{600}$ reached 0.3. 1 mM IPTG was then added and the culture was

incubated for a further hour at 37°C. The $OD_{600}$ of the culture was then determined, and the cells were infected with M13/T7 phage at a ratio of 5pfu/cell. The culture was then grown for approximately 4 hours at 37°C with shaking.

#### 2.9.1.5 Expression of Recombinant Proteins in BL21 (DE3) pLysS *E. coli.*

Expression of genes in pLysS hosts provides extra stability for target genes due to the expression of a small amount of T7 lysozyme, which is a natural inhibitor of T7 DNA polymerase. Expression in BL21 (DE3) pLysS *E.coli* cells (Stratagene) was carried out as described in section 2.9.1.4 except that the M13/T7 phage that contains the T7 RNA polymerase gene was not required as DE3-derivatives of BL21 contains a plasmid that encodes the T7 RNA polymerase gene under the control of the lacUV5 promoter. As such, addition of IPTG to medium induces expression of T7 polymerase with consequent expression of the gene of interest

#### 2.9.1.6 ProBond$^{TM}$ Resin

ProBond$^{TM}$ Resin (Invitrogen) was used to purify the recombinant fusion proteins that were expressed from pRSET vectors.

#### 2.9.1.7 Purification of Recombinant Protein

##### 2.9.1.7.1 Preparation of Denatured *E.coli* Lysate

Cells from the 50ml culture were harvested by centrifugation (2,000×g for 15 minutes) and re-suspended in 10ml of Guanidinium HCl Buffer (pH 7.8). The cell suspension was mixed for 5-10 minutes at room temperature to ensure thorough cell lysis. The cell lysate was then sonicated on ice with three 20-second pulses at a high intensity setting to shear the DNA and RNA. The insoluble debris was removed by centrifugation (microfuge) at 2,000×g for 15 minutes. The sheared lysate was subsequently transferred into a fresh tube and stored on ice or at -20 until use.

#### 2.9.1.7.2 Preparation of ProBond$^{TM}$ Column

3ml of ProBond$^{TM}$ resin (Invitrogen) was loaded into a glass column (0.6cm in diameter and 15cm in length). The column was then washed with 10ml of UHQ $H_2O$ and equilibrated with 10ml of Denaturing Binding Buffer by gravity flow prior to sample application.

#### 2.9.1.7.3 Sample Application, Column Washing and Elution under Denaturing Conditions

The resin was then re-suspended with 5-10 ml of denatured *E.coli* cell lysate by gently rocking the column for 10 minutes to allow the expressed protein containing the poly-Histidine tag to bind. The resin was then allowed to settle by gravity flow. The column was washed with 8ml of Denaturing Binding Buffer, 8ml of Denaturing Wash Buffer pH6.0 followed by 8 ml of Denaturing Wash Buffer pH5.3. Applying 6ml of Denaturing Elution Buffer eluted the protein, which was subsequently collected in 500µl fractions. The elution was monitored by SDS-PAGE gel (2.13.1). The fractions containing the protein were then combined and dialysed against 10mM Tris, pH 8.0, 0.1% Triton X-100 overnight at 4°C to remove the urea, then concentrated by evaporation in dialysis tube at 4°C for 48 hours and stored at -20°C.

**Table 2.5 Solutions used in the XPRESS System for protein purification**

| Buffers | Contents | Total Volume + pH |
|---|---|---|
| Solution A | 200 mM $NaH_2PO_4$ | 1000ml |
| | 5M NaCl | |
| | UHQ | |
| Solution B | 200mM $Na_2HPO_4$ | 1000ml |
| | 5M NaCl | |
| | UHQ | |
| Guanidinium HCl Buffer | 57.3g GuHCl | 100ml |
| | 0.58ml solution A | pH 7.8 |
| | 9.42ml solution B | |
| | 70ml UHQ | |
| Denaturing Binding | 48.1g Urea | 100ml |
| Buffer | 0.58ml solution A | pH 7.8 |
| | 9.42ml solution B | |
| | 70ml UHQ | |
| Denaturing Wash Buffer | 48.1g urea | 100ml |
| | 7.38ml solution A | pH 6.0 |
| | 2.62ml solution B | |
| | 70ml UHQ | |
| Denaturing Wash Buffer | 48.1g Urea | 100ml |
| | 9.17ml solution A | pH 5.3 |
| | 0.83ml solution B | |
| | 70ml UHQ | |
| Denaturing Elution | 48.1g Urea | 100ml |
| Buffer | 10ml solution A | pH 4.0 |
| | 70ml UHQ | |

## 2.9.2 GST Gene Fusion System Expression in *E. coli*

The Glutathione S-transferase (GST) Gene Fusion System is designed for the high level of expression and purification of fusion proteins produced in *E.coli*. DNA fragments that are cloned into the pGEX-4T1 vectors were under the control of a *tac* promoter and positioned in-frame with a fragment that encoded GST. Expressions in *E.coli* subsequently yielded fusion proteins with the GST protein bound at the amino terminus. The fusion protein allows the expressed protein to be purified from the bacterial lysates by affinity chromatography using immobilised glutathione (Frangioni and Neel, 1993;Smith, 1993).

### 2.9.2.1 Cloning into the Expression Vector

All inserted fragments were cloned into a multiple cloning site (MCS) that had six unique restriction enzyme sites positioned downstream of the thrombin cleavage recognition sequence.

To propagate the pGEX-4T-1 expression plasmids, the ligation mixes of the pGEX-4T-1 were transformed into *E.coli* BL21. Transformants were selected with ampicillin plates and the correct constructs were confirmed by restriction enzyme analysis and sequencing.

### 2.9.2.2 Screening pGEX Recombinants for Expression of Recombinant Protein

Expression from the pGEX-4T-1 vector is under the control of a *tac* promoter, which is inducible by the addition of IPTG. The pGEX vectors contain a *lac*$I^q$ gene. The product of this gene is a repressor protein that binds to the operator region of the *tac* promoter. This prevents the expression until IPTG induction.

Several colonies of *E.coli* transformed with the pGEX recombinants were picked and transferred into separate tubes containing 2ml of 2x YTA medium. (16g tryptone, 10g yeast extract, and 5g NaCl in 900ml $dH_2O$). The bacteria were grown at 37°C

with vigorous agitation until the $OD_{600}$ was between 0.6-0.8. The expression of protein was then induced by the addition of 2µl 100mM IPTG (0.1mM final concentration). The bacteria were then incubated for 1-2hours at 37°C. After this time 1.5ml of the culture was removed and the bacteria were pelleted (microfuge, 5 minutes). The supernatant was then discarded and the bacterial pellet re-suspended in 300µl of ice cold PBS.

The bacterial cells were then lysed by one 10 second pulse using a sonicator at high intensity, followed by centrifugation in a microcentrifuge for 5 minutes to remove insoluble material. 20µl of Glutathione Sepharose 4B was then added to each supernatant and mixed gently for 5 minutes at room temperature. 100µl of PBS was then added, votexed and centrifuged for 5 seconds to sediment the Glutathione Sepharose beads. The supernatant was discarded and the PBS wash repeated in total three times. The fusion proteins were eluted by the addition of 10µl elution buffer and incubation for 5 minutes at room temperature. The beads were then sedimented (microcentrifuge, 5 minutes) the 10µl of the elution buffer analysed by SDS PAGE.

#### 2.9.2.3 Expression of the Recombinant Protein

2ml of 2xYTA Amp was inoculated with a single recombinant *E. coli* colony and the culture grown overnight at 37°C with shaking. The culture was then diluted 1:100 into pre-warmed 2x YTA medium and grown at 37°C until the $A_{600}$ reaches 0.5-2. IPTG was then added to a final concentration of 0.1-1.0 mM and incubated for an additional 2-6 hours. The culture was then spun at 8000rpm for 10min at 4°C to sediment the bacteria cells, and then the pellet was placed on ice. The bacterial pellet was then re-suspended in 50µl of ice cold PBS per ml of culture used, and then sonicated on ice for three 20-second pulses at a high intensity.

### 2.9.2.4 Purification of Recombinant Protein

The GST fusion proteins were purified from the bacterial lysate by affinity chromatography using Gluthione Sepharose 4B resin.

The Glutathione Sepharose 4B slurry was first washed in ice cold PBS, then 2ml of 50% slurry was added to each 100ml of bacterial sonicate, followed by a 30 minute incubation at room temperature with gently agitation. The solution was then loaded into a glass column (0.6cm in diameter and 15cm in length) allowed to settle then drained by gravity flow. The Glutathione Sepharose 4B slurry was then washed by adding 10 bed volumes of 1 x PBS for a total of three washes. The bound protein was then eluted in 1ml of elution buffer per 1ml bed volume, incubated for 10 minutes at room temperature then collected. The elution was repeated twice more, and 10µl of each elution was analysed by SDS PAGE.

**Table 2.6 Solutions used for the purification of GST fusion proteins.**

| Buffers | Contents | Total Volume |
|---|---|---|
| Elution Buffer | 50mM Tris-HCL<br>10mM reduced glutathione pH8.0 | store -20°C<br>1-10ml<br>aliquots |
| 100mM IPTG | 500mg IPTG<br>20ml $dH^2O$<br>filter sterilised. | 1ml aliquots<br>store -20°C |

## 2.9.3 Maltose Binding Protein System for Expression in *E. coli*

The Maltose Binding Protein (MBP) Gene Fusion System is designed for the expression and purification of a protein by fusing it to a MBP, which is encoded by the *malE* gene. DNA sequences that are cloned into the pMAL-c2 vectors were placed under the control of a *tac* promoter and positioned in frame with the 3' end of

the *mal*E gene. This results in the expression of an MBP-fusion protein. The fusion protein can subsequently be purified by a one step affinity purification system specific for MBP (Kellerman and Ferenci, 1982).

### 2.9.3.1 Cloning into the Expression Vector

All inserted fragments were cloned into a multiple cloning site (MCS) that had seven unique restriction enzyme sites positioned downstream of the factor Xa cleavage recognition sequence.

To propagate the pMAl-c2 expression plasmids, the ligation mixes were transformed into *E.coli* JM101. Transformants were selected with ampicillin plates and the correct constructs were confirmed by restriction enzyme analysis and sequencing.

### 2.9.3.2 Expression of the Recombinant Protein

25ml of LB containing ampicillin and 0.2% glucose was inoculated with an overnight culture of a single colony known to contain the recombinant plasmid. The culture was then incubated at 37°C with shaking for approximately 2 hours until the $O.D._{600}$ was 0.4-0.6. Then IPTG (final concentration 0.3 mM) was added to the culture IPTG enters the cells then binds to the *lac* repressor, which induces the *lac* promoter. The culture was then incubated for a further 4 hours at 37°C. At the end of the expression period 1ml of the culture was harvested centrifuged and the pellet re-suspended in 200µl SDS-sample buffer. A 20µl sample was analysed by SDS-PAGE.

For a large scale expression a single recombinant *E.coli* colony was inoculated into 2ml of LB Amp (50µg/ml) containing 0.2% glucose and the culture grown overnight at 37°C with shaking. 0.3 ml of this overnight culture was then used to inoculate 50ml LB, Amp, and 0.2% glucose, which was then incubated at 37°C until the $OD_{600}$ reached 0.4-0.6). IPTG was then added (final concentration 0.3 mM) and the culture was incubated for a further hour at 37°C. The culture was then grown for approximately 4 hours at 37°C with shaking. Finally the bacterial cells were harvested (microcentrifuge, 20minutes, 4000 x g) at 4°C and the supernatant

discarded. The bacterial lysate was then sonicated using short bursts to release the protein (1-3 minutes), then centrifuged 20 minutes at 14 000 xg at 4°C and the pellet was discarded. 10µl of the supernatant was then analysed by SDS PAGE.

### 2.9.3.3 Purification of the Fusion Protein

The MBP fusion proteins were purified from the bacterial lysate by affinity chromatography using amylose resin. Amylose resin was poured into a glass column (0.6cm in diameter and 15cm in length) allowed to settle then washed with 8 column volumes of column buffer. The crude protein extract was then loaded on to the column and allowed to empty by gravity flow. The column was then washed with 12 column volumes of column buffer, then the fusion protein was eluted with column buffer/10 mM maltose, collecting ten-twenty 3ml fractions. Each fraction was then analysed by SDS PAGE. The fractions containing protein were combined and dialysed against 10mM Tris, pH 8.0, 0.1% Triton X-100 overnight at 4°C to remove the maltose, then concentrated by evaporation in dialysis tube at 4°C for 48 hours and stored at -20°C.

**Table 2.7 Solutions used for the purification of MBP fusion proteins.**

| Buffers | Contents | Total Volume |
|---|---|---|
| Column buffer | 20ml 1.0 M tris.Cl pH 7.4<br>11.7g NaCl (0.02M final)<br>2.0ml 0.5 M EDTA (1mM final)<br>store 1 month RT. | 1000ml |
| IPTG 0.1M | 1.41 g IPTG<br>50ml $H_2O$<br>store 1 year 4°C | 50ml |

## 2.10 Protein Preparation for Immunisation

Adjuvants were used to enhance the immune response when the amount of antigen is limited. They act by prolonging the half-life of the antigen by protecting it from degradation and minimise any direct toxic effects. They also increase the efficiency of uptake of the antigen by macrophages by allowing a slow sustained release from the site of immunisation. The first injection consisted of recombinant protein combined with Freund's complete adjuvant and heat killed and dried *Mycobacterium tuberculosis*. Further injections consisted of recombinant protein combined with Freund's incomplete adjuvant (paraffin oil and mannide monooleate). Five subcutaneous injections were carried out at ten-day intervals. A pre-immune bleed was obtained prior to the first injection and tested by immunoblotting (2.12.5). The first tail bleed was taken ten days after the last injection, to check the level of antibody produced. A second tail bleed was taken prior to cardiac puncture. The blood was stored at room temperature for one hour and then placed at 4°C overnight. The serum was removed and stored in 50% glycerol at –20°C.

## 2.11 Electron Microscopy

### 2.11.1 EM of Tissue Culture Supernatant

Formvar coated copper EM grids were obtained from the Electron Microscopy department, Southampton General Hospital.

The electron microscope grid was inverted directly onto a drop of tissue culture supernatant on a sheet of parafilm and left for two minutes. The grid was gently blotted using damp 3MM chromatography paper and inverted onto a drop of phosphotungstic acid stain (pH6.0) for 30 seconds. The grid was gently blotted and air dried before examination with a Hitachi 7000 Electron Microscope,

### 2.11.2 EM of Bacterial Cells

A 1ml sample taken from an overnight culture was pelleted (microcentifuge, 1 minute) then fixed in 3% glutaldehyde followed by postfixation by the Electron Microscope department, Southampton General Hospital with 1% osmium tetroxide (OsO4) in 0.1 M cacodylate buffer, dehydration in acetone and embedding in epon. Thin sections were obtained in a Reichert ultramicrotome OmU3, and were observed stained with aqueous uranyl acetate and lead citrate and examined in a Hitachi 7000 Electron Microscope.

## 2.12 Analysis of Proteins

### 2.12.1 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE separates linear polypeptide-SDS complexes by electrophoretic molecular sieving in polyacrylamide gels. The proteins are first denatured using SDS and β-mercaptoethanol, which destroys any secondary and tertiary structures and inter- and intra-chain disulphide bonds respectively. The treated proteins have approximately equal charge to mass ratios and are separated solely on the basis of molecular size.

All glass plates and spacers (1.5mm) used to form the gel mold were cleansed and degreased with IMS (industrial methylated spirits). SDS-PAGE was made with 10% separating and 4% stacking acrylamide gels, unless otherwise stated, and prepared as shown in table 2.8. All buffers and solutions were prepared using UHQ $H_2O$ (table 2.9). Polymerisation of acrylamide and the cross-linker, N,N,-methylenebisacrylamide (Electran grade; Merck) to form gels was accelerated by free radicals from N,N,N'-N'-tetra-methylenediamine (TEMED) produced by the action of the initiator ammonium persulphate (APS).

The gel solutions were degassed to reduce the oxygen content, as this is an inhibitor of the acrylamide polymerisation. The separating gel was cast first allowing room for the 4% stacking gel to be cast on top. The separating gel was overlaid with isobutanol

to ensure the gel surface was flat, and air was excluded. After the separating gel had set, the isobutanol was removed and the surface of the gel was washed several times with UHQ $H_2O$. The stacking gel was then poured and a comb (which is used to make the sample wells) carefully inserted. Polymerisation was complete after ~30 minutes.

The well-comb and bottom spacer, were removed from the gel and the gel was assembled in an electrophoresis tank. The required volume of electrode buffer (Table 2.5) was freshly prepared. Protein samples were diluted in SDS sample buffer then denatured by boiling for 5 minutes and then loaded onto the gel. Electrophoresis was carried out at 40mA constant current per gel. On completion, the proteins within the gels were stained using PAGE blue 83 (Coomassie) stain (2.12.3).

**Table 2.8 Materials used for SDS-PAGE gels**

| Solution | Contents |
|---|---|
| Acrylamide monomer solution | 40% stock solution |
| SDS stock solution | 10% (w/v) |
| 5 X Separating Buffer | 23g Tris dissolved in 100ml UHQ $H_2O$ → pH 8.9 |
| 5 X Stacking Buffer | 6.78g Tris in 100ml UHQ $H_2O$ → pH 6.7 |
| N,N,N',N'-tetramethylethylenedamine | TEMED |
| Ammonium Persulphate solution | 10% (w/v) - freshly prepared |
| Electrode Buffer | 20mM Tris, 192mM glycine, 0.1% (w/v) SDS - freshly prepared |
| Sample Bufffer | 125mM Tris-HCl pH6.8, 4% (w/v) SDS, 20% (w/v) glycerol, 10% (v/v) 2-mercaptoethanol, 10μg/ml bromophenol blue |

**Table 2.9 The Components of SDS-PAGE Gels**

| | 10 % Gel | | 12.5 % Gel | |
|---|---|---|---|---|
| | Separating Gel (10%) | Stacking Gel (4%) | Separating Gel (12.5%) | Stacking Gel (5.3%) |
| Acrylamide/ Bis solution (37:1) | 5ml | 1ml | 6.65ml | 1.33ml |
| 5 X Separating Buffer | 4ml | | 4ml | |
| 5 X Stacking Buffer | | 1ml | | 1ml |
| 10% SDS | 0.2ml | 0.1ml | 0.2ml | 0.1ml |
| UHQ $H_2O$ | 8.75ml | 6.92ml | 7.1ml | 6.95ml |
| TEMED * | 37.5µl | 25µl | 37.5µl | 25µl |
| Ammonium Persulphate (5mg/ml) * | 2ml | 1ml | 2ml | 1ml |

* add after degassing and immediately before casting gel.

### 2.12.2 SDS-PAGE Protein Markers

The molecular weights and concentrations of proteins were estimated by comparing electrophoretic mobility with proteins of known molecular weight (tables 2.10 and 2.11). A lyophilised mixture of protein markers (Sigma), with molecular weights ranging from 14,200 - 205,000 kDa (tables 2.10 and 2.11) were re-suspended in 1ml sample buffer, boiled for 5 minutes and 15µl loaded into one well. Each band of protein contained approximately 1.7µg. Alternatively 10µl of a Broad Range Protein Marker, (Promega) was added (table 2.12)

**Table 2.10 Low Molecular Weight Protein Markers**

| Low molecular weight protein markers (Da) | |
|---|---|
| Albumin, bovine | 66 000 |
| Albumin, egg | 45 000 |
| Glyceraldehyde-3-P-dehydrogenase | 36 000 |
| Carbonic anhydrase, bovine | 29 000 |
| Trypsinogen, bovine pancreas | 24 000 |
| Trypsin inhibitor, soybean | 20 000 |
| α-Lactalbumin, bovine milk | 14 200 |

**Table 2.11 High Molecular Weight Protein Markers**

| High molecular weight protein markers | (Da) |
|---|---|
| Myosin, rabbit muscle | 205 000 |
| β-Galactosidase, *E.coli* | 116 000 |
| Phosphorylase b, rabbit muscle | 97 400 |
| Albumin, bovine | 66 000 |
| Albumin, egg | 45 000 |
| Carbonic anhydrase, bovine | 29 000 |

**Table 2.12 Protein Markers, Broad Range (Promega)**

| Protein markers | MW (Da) |
|---|---|
| Myosin | 212 000 |
| MBP-β-galactosidase | 158 194 |
| β-galactosidase | 116 351 |
| Phosphorylase b | 97 184 |
| Serum albumin | 66 409 |
| Glutamic dehydrogenase | 55 561 |
| MBP2[*1] | 42 710 |
| Lactate dehydrogenase | 36 487 |
| Triosephosphate isomerase | 26 625 |
| Trypsin inhibitor | 20 040-20 167 |
| Lysozyme | 14 313 |
| Aprotinin | 6 517 |
| Insulin A, B chain | 2 340-3 400 |

### 2.12.3 Protein Staining

The glass dish used for the staining procedure was washed and degreased with IMS and the gel was placed in Coomassie staining solution [0.1% (w/v) PAGE Blue 83, 20% (v/v) isopropanol, 10% (v/v) acetic acid] for 1 hour. Excess stain was removed by repeated washes in a destain solution [20% (v/v) isopropanol, 10% (v/v) acetic acid] until the gel background was clear. The gel was then placed in gel drying solution [25% (v/v) methanol, 5% (w/v) glycerol] for 1 hour then dried between two sheets of gel drying film (Promega gel drying kit).

### 2.12.4 Electrophoretic Transfer of Proteins

Protein samples were electrophoresed through a SDS polyacrylamide gel as described earlier (2.12.1). Then transferred onto Hybond-C nitrocellulose membrane (Amersham International) in the presence of the SDS-PAGE running buffer diluted 4:1 with methanol. A Trans-Blot semi-dry blotting apparatus (Biorad) was used for the transfer of proteins as described by the manufacturer. Before transfer the high molecular weight marker was sliced off the gel to check that the gel had run correctly. A sheet of nitrocellulose and eight pieces of 3mm chromatography paper (Whatman International) were cut to give a 0.5cm margin around the gel. The gel and membrane were then sandwiched between eight sheets of buffer-soaked 3mm paper and any trapped air bubbles removed by rolling a glass Pasteur pipette over the surface. The gel/membrane assembly was then placed on the horizontal anode electrode plate, with the membrane on the anode side relative to the gel. The cathode was then placed over the top of the assembly and electrophoresis performed for 60 minutes at 0.8 mA per $cm^2$ of gel.

The efficiency of transfer was checked by staining the post-transfer gel (2.12.3) and a strip of the membrane containing one of the molecular weight markers in methanol (50%) and glacial acetic acid (10%) containing 0.5% (w/v) amido black. Amido black staining was performed for 10 minutes at room temperature after which time the membrane strip was de-stained with methanol (50%) and glacial acetic acid (10%) until the background staining was reduced to adequate level.

### 2.12.5 Immunological Detection of Proteins

The immunological detection of proteins was performed by first incubating the Hybond-C membrane in a "blocking" solution (5% (w/v) Marvel, 20mM Tris-HCl pH 7.5, 0.5M NaCl and 0.05% (w/v) sodium azide) overnight (16 hours). The milk protein in the Marvel bound to, or "blocked", the free protein binding sites of the membrane, preventing non-specific binding of antibodies.

Any excess Marvel was removed with 2×5 minute washes in TTBS (Tris-HCl pH 7.5, 0.5M NaCl, 0.05% (w/v) Tween-20). The transferred proteins were then detected by incubating the membrane overnight at 4°C in a primary antibody solution containing 10% normal goat serum in TTBS plus primary antibody (mouse or rabbit serum diluted 1 in 100 or 1 in 1000). The residual unbound antibody was removed with 3×5 minute washes in TTBS and 1×5 minutes in TBS (20mM Tris pH7.5, 0.5M NaCl).

The primary antibody was then detected by incubating the membrane in a secondary antibody-alkaline phosphatase conjugate at a dilution recommended by the manufacturer (1 in 1,000 for goat anti-rat and goat anti-mouse, Bio-Rad) in 1% gelatine and TTBS at room temperature for 1 hour. Residual unbound antibody was removed by 3× 5-minute washes in TTBS, followed by 2× 5 minute washes in TBS. The final washes with TBS removed all traces of Tween-20 that may interfere with the enzyme activity of the alkaline phosphatase.

The transferred proteins were indirectly detected by the action of the alkaline phosphatase-conjugate on the substrates NBT and BCIP. After the final TBS washes the membrane was incubated in 19.6ml carbonate buffer (0.1M $NaHCO_3$, 1mM $MgCl_2$, pH 8.9) containing 200µl NBT [30mg/ml in 70% (v/v) DMF] and 200µl BCIP (15mg/ml in 100% DMF) at room temperature in the dark until the colour was sufficiently developed. The enzyme reaction was stopped, by washing the membrane in $dH_2O$. The membrane was finally air-dried and stored at room temperature protected from light.

## 2.13 Assay for β-Galactosidase activity

Assays for β-galactosidase were performed on intact bacteria, using the method described by Nicolaidis and Drabble (1979) (Nicolaidis and Drabble, 1979).

Cells were first permeabilised by the addition of 0.5ml cetyl trimethylammonium (CTAB, 80ug/ml) to 0.5ml cell culture and mixed on a vortex for 5 seconds. 4ml of 3mM ortho-nitrophenyl-β-D-galactoside (ONPG) in 0.05M sodium phosphate buffer (pH 7.0), prewarmed to 37°C, was then added and mixed by vortexing. Incubation was at 37°C for 5-60minutes, the reaction being terminated by the addition of 1.0 ml sodium carbonate (1.0M). The absorbance at 420nm (due to ortho-nitrophenol), and 550nm (plus a 600nm reading was taken at the time the sample was taken for assay) was measured. β-galactosidase activity was measure in 'Miller units' (Miller, 1972) using the equation below;

$$\text{Miller Units} = 1000 \times A_{420} - 1.75(OD_{550})/t \times v \times OD_{600}$$

$A_{420}$ and $OD_{550}$ are read from the reaction mixture.
$OD_{600}$ is the culture optical density at the time the sample was taken for assay.
T is the reaction time, in minutes.
V is the volume of culture assayed, in ml.

## 2.14 *In vitro* Coupled Transcription and Translation

### 2.14.1 TNT T7 Quick System

*In vitro* coupled transcription and translation was carried out using the Promega TNT® T7 Quick system. The reaction was set up as follows: 100-500ng Plasmid DNA (Qiagen miniprep or maxiprep), 20µl TNT® T7 Quick master mix, 1µl $^{35}S$ methionine and UHQ $H_2O$ to 25µl. The reaction was incubated at 30°C for 2 hours. Then 1-2µl of the reaction products analysed by SDS-PAGE (2.12.1). The gel was Coomassie stained (2.12.3) and destained as usual, then placed for 30 minutes in a solution containing 32g sodium salicylate, 100ml methanol and 100ml $dH_2O$, which enhances the signal obtained on autoradiography. The gel was then vacuum dried

onto a sheet of 3mm chromatography paper (Whatman International) and the reaction products were detected by autoradiography at -70°C (2.16).

### 2.14.2 *E. coli* T7 Extract System for Circular DNA

Coupled *in vitro* transcription/translation of circular DNA was carried out using a Promega *E. coli* T7 Extract System for Circular DNA. The reaction was set up as follows: 1μg plasmid DNA (Qiagen miniprep or maxiprep), 5μl amino acid mixture minus methionine, 20μl S30 premix without amino acids, 1μl $^{35}S$ methionine,15μ T7 S30 extract, circular, and nuclease free water to a final volume of 50μl. The reaction was vortexed then centrifuged for 5 seconds and incubated for 1-2 hours at 37°C. The reaction was then stopped by incubating the samples in an ice bath for 5 minutes. 5μl of the reaction was then acetone precipitated then 10μl analysed by SDS PAGE. The gel was coomassie stained and destained as usual, then soaked in a solution of sodium salicylate, 100ml methanol and 100 ml $dH_20$. This enhances the signal obtained by autoradiography. The gel was then vacuum dried onto a sheet of 3mm chromatography paper (whatman International) and the reaction products were detected by autoradiography at -70°C.

To avoid background staining an acetone precipitation was carried out on the reaction products to remove PEG from the S30 extract. 5μl of reaction product was added to 20μl of acetone then incubated for 15 minutes on ice. The sample was then centrifuged at 12,000 x g for 5 minutes and the pellet vacuum dried for 15 minutes. Once the pellet was fully dry it was re-suspended in 20μl of SDS PAGE sample buffer, heated for 2-5 minutes to 100°C and 1-5μl analysed by SDS PAGE gel.

## 2.15 Radio-Immune Precipitation Assay (RIPA)

Products from the *in vitro* transcription/translation systems (5μl) were incubated at 37°C with 5μl undiluted rabbit or mouse antisera in 800μl of RIPA buffer {10mM Tris-HCL, pH7.5, 1mM EDTA, 0.15mM NaCl, 0.1% SDS, 0.5% Empigen BB [N-dodecyl-N, N-dimethylglycine], 0.1mM phenylmethylsulphonylfluoride} for 1 hour. After incubation beaded agarose coated with goat anti-mouse IgG (Sigma) or goat

anti-rabbit IgG (Sigma) was added and incubated at room temperature for 1 hour to adsorb immune complexes. The beads were then washed three times in RIPA buffer and once in PBS prior to derivatisation in sample dissociating buffer and separation by SDS-PAGE gel. The gels were stained and incubated in 1M sodium salicylate/50% methanol to prepare for autoradiography. Then vacuum dried and exposed to Kodak XAR-5 film at -70°C for 16 hours.

## 2.16 Autoradiography

Dried SDS-PAGE gels were marked with fluorescent paint, which orientates the film during autoradiography. The dried gel was then placed with a sheet of film (Kodak X-Omat scientific imaging film, Sigma) in an autoradiography cassette (Kodak) containing an intensifying screen. After 2-24 hours at -70°C, the film was developed using a Kodak automated developer.

## 2.17 Northern Blotting

### 2.17.1 RNA Gels Containing Formaldehyde

The method was adapted from Lehrach et al (1977) and Seed et al (1982) (Lehrach *et al.*, 1977;Seed, 1982). Formaldehyde vapours are toxic so all solutions containing formaldehyde were prepared in a chemical hood. Gels were also run in the hood with the air pressure on and the glass screen closed. 150ml of gel mix was prepared. For this 1.5g of agarose was melted in 110mls $dH_2O$ after which 15mls of 10X MOPS was added.

10X MOPS 0.2M'(N-morpholino) propane sulfonic acid
0.05M sodium acetate
0.01M $Na_2EDTA$ pH 7.0

The mixture was cooled to 50°C at which point 25ml formaldehyde was added. The gel was immediately poured and allowed to set for at least one hour before running. The gel was run overnight at 25 volts in 1X MOPS.

### 2.17.2 Denaturation of RNA

Less than 20μg of RNA was used in each reaction. To this 2μl of 10X MOPS was added, 10μl of deionised formamide and 3.5μl of formaldehyde. If any yellow colour was present the formamide was deionised further by adding Dowex XG8 mixed-bed resin and stirred before adding. The reaction was heated to 65°C for five minutes and snapped cool on ice. All reactions were performed in eppendorf tubes. To each tube 4μl of filtered ficoll dyes was added prior to loading on the gel.

### 2.17.3 Vacuum Blotting

The LKB2016 VacuGene Blotting system was used to transfer the formaldehyde gel to a membrane. The gel was placed onto the transfer membrane, which was held within a 'window' in a waterproof mask. The mask was supported on the porous screen on the base of the VacuGene unit. After locking the upper frame to the base, the base was evacuated. The gel was then flooded sequentially with 10XSSC for 1 hour. By continuously subjecting the gel to vacuum, transfer to the membrane occurs in one direction only. Hybond-N nitro-cellulose membranes were used. Once blotting was complete the Hybond-N was baked for two hours at 80°C in a baking oven to fix the membrane before hybridisation was started.

## 2.18 $^{32}$P Radiolabelling of Double Stranded cDNA

The radiolabelling of cDNA probes were performed using the GibcoBRL random primers DNA labelling system. cDNA amplified by PCR was purified using the Qiagen gel extraction kit. The DNA was denatured at 90°C for 5 minutes followed by immediate transfer to ice, and the labelling reaction assembled as follows; 25ng denatured template, 10μl 5x labelling buffer, 2μl non-labelling dNTPs (1μl 1.5mM

dATP, dTTP, dGTP mixed), 2µl BSA 1mg/ml, 5µl [$\alpha^{32}$P] dCTP (Amersham), 1µl Klenow enzyme (5 units).

The 5X labelling buffer contains random hexadeoxyribonucleotides, which bind randomly along the single-stranded DNA template. Klenow enzyme (DNA polymerase I lacking 3' to 5' exonuclease activity) fills in the gaps between primers using dNTPs, one of which is labelled with $^{32}$P. This produces dsDNA in which one of the two strands is radiolabelled. The reaction was stopped by boiling the samples at 90°C for five minutes.

### 2.18.1 Prehybridisation

100x Denhart's solution was made containing 2g BSA (Sigma), 2g ficoll molecular weight 400,000 (Sigma) and 2g polyvinylpyrrolidone (BDH) in 100ml UHQ $H_2O$ and stored at-20°C.

An appropriate volume of pre-hybridisation solution was made containing 5x Denhart's solution, 6x SSC and 0.5% SDS. The pre-hybridisation solution was warmed to the hybridisation temperature; 65°C for homologous probes. Hybridisation bottles, 300x32mm (Hybaid) had 10ml of warmed prehybridisation solution added. 20µl 10mg/ml denatured sonicated herring sperm DNA (Promega) was added to give a final concentration of 20µg/ml. Prehybridisation was carried out for at least 4 hours at 65°C in a Hybaid hybridisation oven, to block non-specific binding of the probe DNA.

### 2.18.2 Hybridisation

The $^{32}$P probe DNA (from a 50µl labelling reaction) was denatured at 90°C for 5 minutes and transferred directly to ice, the 10µl was added to each hybridisation bottle. This concentration of labelled DNA (5ng) gave a clear background. A control of 25-50ng unlabelled probe DNA on a separate filter was included in each experiment to ensure correct labelling and hybridisation had taken place. The hybridisation was carried out overnight at 65°C for homologous probes.

### 2.18.3 Removal of Excess Probe

The hybridisation solution was discarded (using an isotope disposal sink) and the filter's were gently rinsed with approximately 10ml pre-warmed 65°C 2x SSC solution. This was immediately discarded and 10ml 2x SSC wash was added, for 15 minutes. The final wash contained 10ml 2xSSC and 0.1% SDS and was left on for 30 minutes. Filters were removed from the hybridisation bottles and covered in Saran wrap (Dow Chemical Company), then attached to 3mm chromatography paper (Whatman International), before autoradiography.

## 2.19 Immunofluorescent Antibody Stain

### 2.19.1 Propagation of BGMK cells

"Buffalo Green" monkey kidney cells were used for the growth of *C. abortus* MA (carrying Chp2). All procedures were performed in a Class II safety cabinet to protect the cells. BGMK cells were grown on coverslips in six-well trays in Dulbecco modified Eagle medium supplemented with 10% (vol/vol) fetal calf serum.

### 2.19.2 Infection of BGMK Cells with *C. abortus*

Cells were infected with *C. abortus* MA (carrying Chp2) by centrifugation at 1,000 x g for 1 hour in medium containing cycloheximide (1μg/ml) and gentamicin (25μg/ml). All procedures were carried out in a Class I safety cabinet.

### 2.19.3 Fixing Coverslips in 24 Well Trays

The medium was removed from each well and the cells were washed in PBS. Cold methanol was then added to each well (this fixes the cells and inactivates any infectious chlamydiae). The cells were then incubated at 0°C for 20 minutes. The methanol is then removed and 1ml of PBS containing 0.1% sodium azide was added to each well. The cells were then stored at 4°C for up to 4 weeks.

### 2.19.4 Staining the Coverslips with Fluorescent Antibody

The Chp2 particles were detected by incubating the coverslips on a 50µl drop of a primary antibody solution (monoclonal antibody 55 (1:5) or polyclonal antisera raised to the proteins encoded by the eight major ORFs (1:50)) then incubated in a humid box for 40 minutes at 37°C. The coverslips were then washed for 3 minutes in PBS for a total of 3 washes, and the primary antibody was detected by incubating the coverslips on a 50µl drop of a secondary antibody FITC conjugate at a dilution recommended by the manufacturer (1:50) (diluted in 0.0025% Evans blue). The coverslips were placed in a humid box and incubated at 37°C for 30 minutes. Then finally washed 3 times for 3 minutes in PBS then once in UHQ $H_2O$, and mounted onto a labelled slide (1:1, glycerol:PBS). The cells were then examined using a Leitz fluorescence microscope.

# Chapter 3

## Search for new chlamydiaphages: discovery of Chp3

### 3.1 Introduction

The family *Chlamydiaceae* consists of two separate genera, *Chlamydophila* and *Chlamydia* that contain six and three chlamydial species respectively. The obligate intracellular developmental cycle of chlamydiae reduces the opportunity for interactions with bacteriophages of free-living bacteria. Therefore it was surprising when bacteriophages were discovered infecting chlamydiae. The first chlamydiaphage to be fully characterised was chlamydiaphage 1 (Chp1). Chp1 was isolated from an avian *C. psittaci* strain that was the causal agent in an outbreak of ornithosis (Richmond *et al.*, 1982). The second chlamydiaphage to be described chlamydiaphage guinea pig 1 (ϕCPG1), was isolated fourteen years later from a *C. caviae* strain that causes guinea pig inclusion conjunctivitis (GPIC) (Hsia *et al.*, 1996). Then in 2000, two novel chlamydiaphages were described; chlamydiaphage 2 (Chp2) isolated from *C. abortus* (Liu *et al.*, 2000) and the dsDNA genome of chlamydiaphage AR39 (ϕCPAR39), which was discovered in *C. pneumoniae* AR39 as a DNA contaminant during genome sequencing (Read *et al.*, 2000a). ϕCPAR39 was later shown to be an independently replicating bacteriophage (Everson *et al.*, 2002). These four bacteriophages share similar features, they are small icosahedral particles that contain a circular, single-stranded DNA genome (Storey *et al.*, 1989a; Liu *et al.*, 2000; Hsia *et al.*, 2000; Read *et al.*, 2000a).

Molecular characterisation has shown that chlamydiaphages are distantly related to members of the *Microviridae* family of bacteriophages (Storey *et al.*, 1989a; Liu *et al.*, 2000; Hsia *et al.*, 2000; Read *et al.*, 2000a). These bacteriophages have been isolated from diverse hosts, which include proteobacteria (Sanger *et al.*, 1978), spiroplasma (Ricard *et al.*, 1982) and chlamydiae (Storey *et al.*, 1989a). The evolutionary history of this group of viruses, and the breadth of hosts they infect is unknown. It has been suggested (Fane *et al.*, 2002) that the evolution of these

bacteriophages is unlike that of double stranded bacteriophages, which, rely on horizontal gene transfer (Hendrix *et al.*, 1999) and may involve species jumping. This idea would rationalize the presence of closely related single-stranded DNA bacteriophages in unrelated hosts.

The aim of this study was to search for new chlamydiaphages. This would help to determine the extent and nature of bacteriophage variation.

## 3.2 Rationale of chlamydiaphage infection of chlamydiae

Chlamydiaphages have already been identified in four different species of *Chlamydophila* (*C. psittaci, C. caviae, C. pneumoniae and C. abortus).* I therefore wanted to test the hypothesis that infection of chlamydiae with closely related *Microviridae* is widespread. The aim of this investigation was to determine the extent of bacteriophage infection of chlamydiae and to identify novel single-stranded DNA bacteriophages in other species of *Chlamydiaceae*.

### 3.2.1 Probing chlamydial samples for bacteriophages

To determine the extent of chlamydiaphage infection of *Chlamydiaceae*, the entire laboratory chlamydial collection consisting of thirty-nine different live chlamydial strains representing all nine species, were screened for the presence of bacteriophage. Due to the obligate intracellular development cycle conventional methods that have previously been used to screen bacteria for the presence of bacteriophages such as plaque assays on agar plates, could not be used. There are two possible approaches for screening samples of chlamydiae for bacteriophages; hybridisation analysis or screening isolates by immunofluorescence using a chlamydiaphage-specific reagent.

Hybridisation analysis is a simple technique that could be used to screen a large number of chlamydial isolates for the presence of bacteriophage using the chlamydiaphage genome as a probe. However, it has been shown from genome sequencing projects that a truncated chlamydiaphage ORF4 insert is present in a

'plasticity zone' of the *C. pneumoniae* AR39 genome (Read *et al.*, 2000a). This was surprising, especially the identification of an identical truncated insert in the chromosome of a *C. pneumoniae* CWLO29 that does not carry a chlamydiaphage (Read *et al.*, 2000a). Therefore, hybridisation using the Chp2 genome as a probe could not be used as a simple method to screen chlamydial samples for the presence of bacteriophages, due to the possibility that other random fragments of the chlamydiaphage genome may be integrated into a variety chlamydial genomes, which, have subsequently been naturally 'cured' of bacteriophage.

Immunofluorescence screening requires a chlamydiaphage-specific reagent that would allow the detection of bacteriophage antigen infecting metabolically active chlamydial isolates. A monoclonal antibody (55) specific to the chlamydiaphage VP1 protein has recently been developed which reacts with the coat protein of several chlamydiaphages by immunofluorescence (Everson *et al.*, 2002). This monoclonal was subsequently used to screen the laboratory culture collection of chlamydiae for chlamydiaphages. A single isolate that reacted with the monoclonal antibody was identified in the collection (figure 3.1). Chlamydiae from this sample were grown in cell culture to expand the amount of biological material for further analysis. The presence of a 'Chp2-like' bacteriophage in the cell culture supernatant was verified by negative stain electron microscopy (figure 3.2). Typical clusters of small round featureless bacteriophage particles with a diameter of approximately 25nm similar to the previously described chlamydiaphages were observed in the sample confirming the presence of a new bacteriophage termed Chp3.

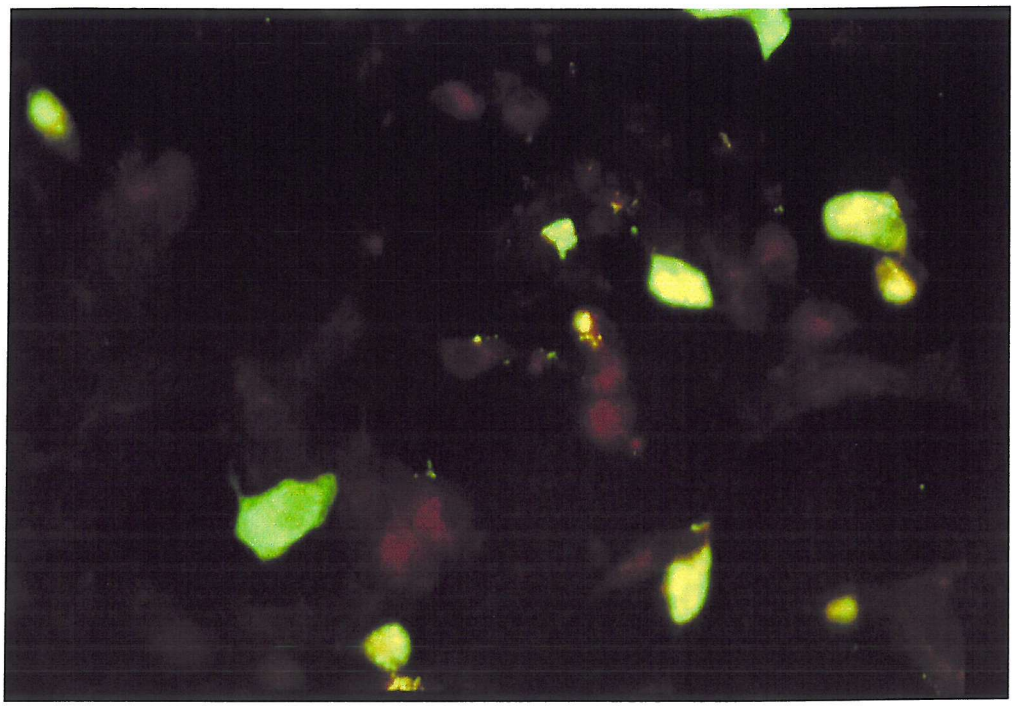


**Figure 3.1** Fluorescent micrograph of Chp3 infected chlamydial inclusions using monoclonal antibody 55.

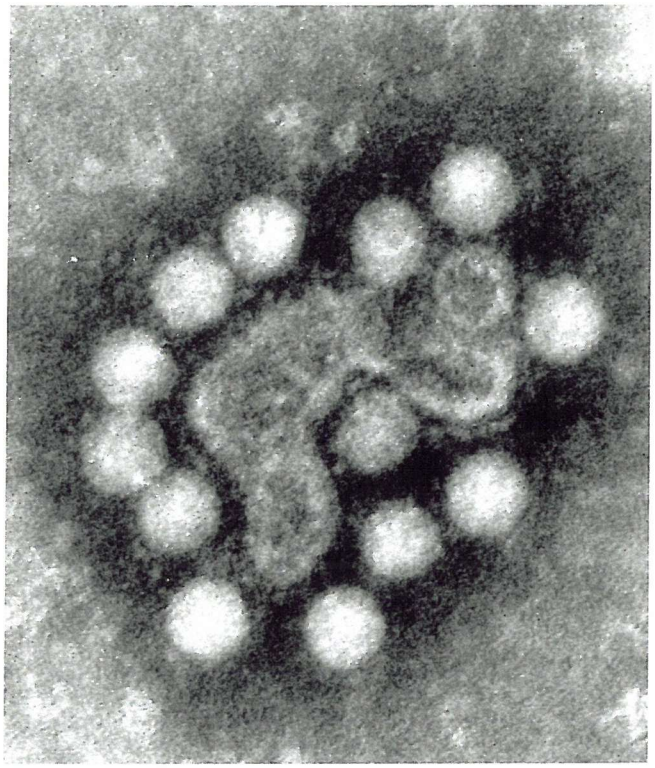



**Figure 3.2** Negative stain electron micrograph of Chp3 particles.

## 3.3 Identification of the Chp3 host

The identity of the host chlamydiae for Chp3 was determined using primer pair U23F (5' GATGCCTTGGCATTGATAGGCGATGAAGGA3') and 23SIGR (5' TGGCTCATCATGCAAAAGGCA 3') designed to amplify a fragment of the 23S rRNA gene from families of bacteria in the order *Chlamydiales* (Everett *et al.*, 1999). The template chlamydial DNA was isolated using a DNA extraction kit as described in section 2.7.3.2. PCR was performed using Bio-x-Act DNA polymerase for 30 cycles of 97°C for 30 seconds, 61°C for 15 seconds and 72°C for 30 seconds (2.7). The PCR product was analysed by gel electrophoresis (2.5.2) and purified to remove any remaining template, nucleotides or primers by absorption to a silica membrane (2.7.3.2). The PCR amplified a 600bp fragment of DNA that consisted of the region downstream of the chlamydiae 23S rRNA. The 600bp PCR product was then subjected to sequence analysis using the primers U23F and 23SIGR (2.8). A sequence of 540bp was generated internally from the 600bp PCR product. A BLAST search of this sequence was made into the GenBank database. This gave a 100% match with *C. pecorum* strains I, Z, L71 1 and 1710S (Everett and Andersen, 1997) showing that the host strain for Chp3 is an isolate of *C. pecorum*.

## 3.4 Extraction of Chp3 DNA

DNA was extracted from purified EBs of the *C. pecorum* strain containing Chp3 using a mini prep kit (2.4.2.2). A low yield of bacteriophage DNA was obtained using this method, as unlike Chp2 DNA extractions, RF bacteriophage DNA could not be visualised in the sample following gel electrophoresis. To verify the presence of chlamydiaphage DNA in this sample, a primer pair was designed based on conserved chlamydiaphage sequence. Initially, the nucleotide sequence of the four chlamydiaphage genomes were aligned and regions conserved between all four chlamydiaphages were used to design primers. Primer ChpDetectF (5'-ATGAAGTTCTYCCTGGAGATAC-3') and primer ChpDetectR (5'-GCAACTTTAGTAGGTAGACCAA-3') were designed to amplify a 300bp fragment from the VP1 region of the chlamydiaphage genome.

A PCR was performed with these primers using the Chp3 DNA as template for 30 cycles of 20 seconds at 94°C, 20 seconds at 44°C and 15 seconds at 72°C (2.7). Bio-x-Act DNA polymerase was used for the PCR amplification to minimise the risk of random mutations because it has 3'-5' proof-reading activity which prevents misincorporations during primer extension. The fragment generated in the PCR was then purified by absorption to a silica membrane (2.7.3.2) and sequenced.

Comparisons between the DNA sequence of the fragment amplified from Chp3 using the primers ChpdetectF and ChpdetectR with the other chlamydiaphages can be seen in figure 3.3.

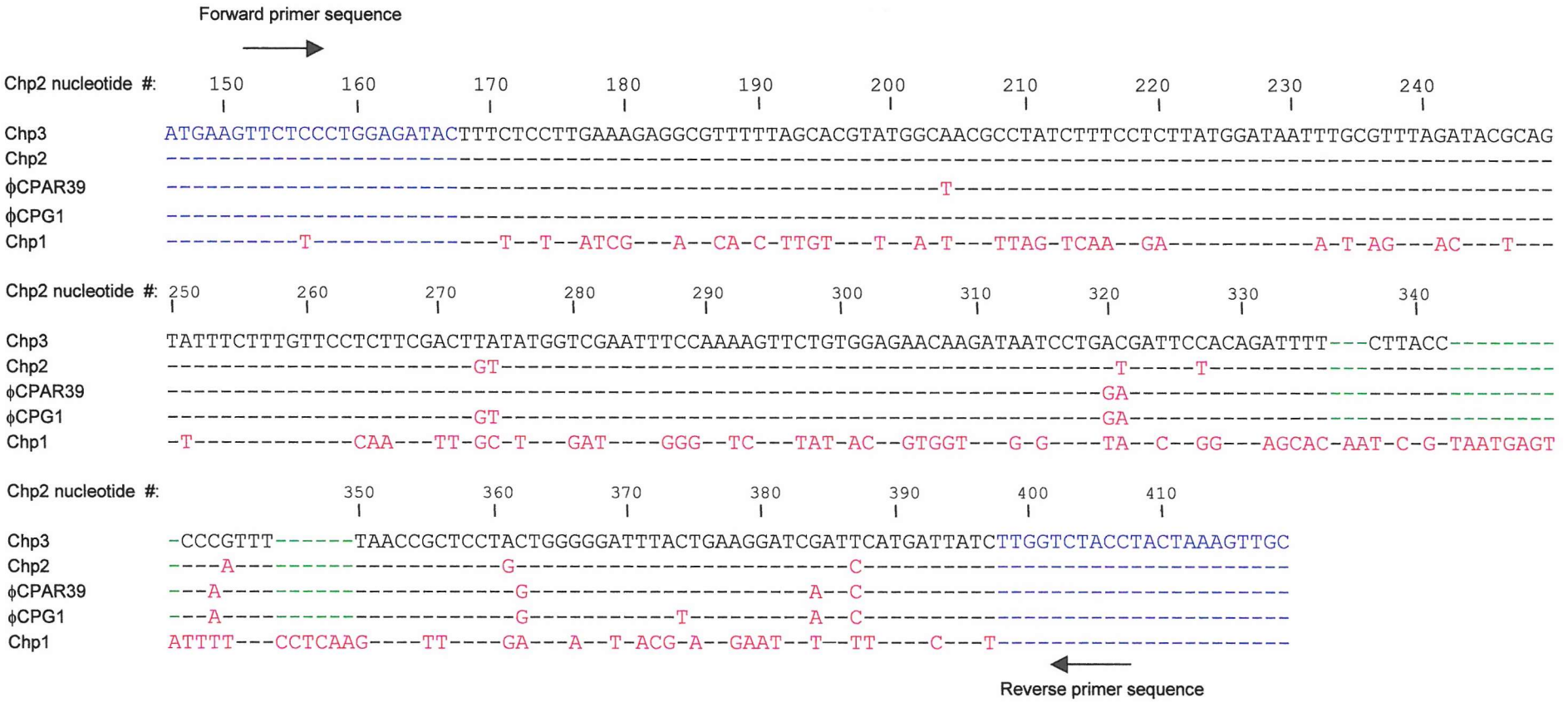



**Figure 3.3** Nucleotide sequence alignment between the DNA fragment amplified from the Chp3 genome (with the chlamydiaphage detection primers), to the corresponding regions in the other chlamydiaphage genomes. The nucleotides that are divergent to the Chp3 sequence in this region are indicated and conserved nucleotides are indicated by black dashes. Green dashes indicated gaps between the sequences. The sequence of the two primers are shown in blue.

Chp3 shows high nucleotide sequence identity (97.6%) to Chp2 and ϕCPAR39 in this region (table 3.1). This indicates that Chp3 is more closely related to the Chp2-like bacteriophages (Chp2, ϕCPG1 and ϕCPAR39) than Chp1.

**Table 3.1 Nucleotide sequence identity (%) between Chp3 and the other chlamydiaphages (Chp2, ϕCPG1, ϕCPAR39, Chp1) in the region amplified by the chlamydiaphage detection primers.**

| | **Chp3** |
|---|---|
| **Chp2** | 97.6 |
| **ϕCPG1** | 96.9 |
| **ϕCPAR39** | 97.6 |
| **Chp1** | 56.9 |

The similarities in genome sequence verified the original name of this bacteriophage as chlamydiaphage 3 (Chp3) as the next in sequence in concordance with the discovery of other similar bacteriophages in the UK e.g. chlamydiaphage 1 (Chp1) and chlamydiaphage 2 (Chp2).

### 3.4.1 Amplification of Chp3 DNA for Sequencing

Attempts were made to grow Chp3 so that genomic DNA could be isolated in large quantities. However, the yield of Chp3 was low, thus to amplify the bacteriophage genome in large enough quantities for sequence analysis, the original primers that were produced to sequence Chp2 (Liu *et al.*, 2000) were utilised. The whole Chp3 genome was amplified as nineteen 400-800bp fragments. The primer pairs used to amplify the fragments of Chp3 and the primers used to sequence each region are shown in table 3.2.

**Table 3.2 Primer pairs used to amplify the genome of Chp3**

| Primer Pairs | Size amplified in Chp3 (bp) | Primers used to sequence each region |
|---|---|---|
| ph23→ph12 | 400 | ph23, ph12, ph24 |
| ph24→ph26 | 440 | ph24, ph26, ph25 |
| ph25→ph28 | 800 | ph25, ph28, ph10, ph11 |
| ph27→ph29 | 800 | ph27, ph29, ph13, ph14 |
| ph17→ph20 | 500 | ph17, ph20, ph30 |
| ph30→ph31 | 500 | ph30, ph19, ph31 |
| ph19→ph18 | 400 | ph19, ph18, ph3 |
| ph7→ph2 | 300 | ph7, ph2 |
| ph32→ph21 | 800 | ph32, ph21, ph1, ph9 |
| ph9→ph15 | 500 | ph9, ph16, ORF4R, ORF5F, ORF5R, ph23, ph15 |
| ph7→ph22 | 500 | ph7, ph32, ph2, ph22, ORF4F |
| ph11→ph14 | 750 | ph14, ph28, ph27, ph11 |
| ph24→ph10 | 700 | Ph26, Chp3-1 |
| ph11→ph5 | 700 | ph11, Chp3-2, ph14 |
| ph13→ph29 | 400 | ph29, Chp3-4 |
| ph17→ph20 | 400 | ph17, Chp3-5 |
| ph7→ph8 | 700 | ph22, Chp3-6, ph2, Chp3-7, ph7, Chp3-8 |
| ph23→ph12 | 500 | Chp3-9 |
| ph27→ph29 | 700 | Chp3-3 |

The nucleotide sequence and genomic location of all primers in the Chp3 genome can be seen in appendix 1

The PCR conditions for each primer pair consisted of 30 cycles of 20 seconds at 94°C, 20 seconds at 46°C followed by 48 seconds at 72°C. The PCR products were analysed by gel electrophoresis (2.5.2) and purified using a PCR purification kit (2.7.3.1). The PCR products were then subjected to sequence analysis using the primers in the table above. All sequencing reactions were carried out as described in section 2.8.

The genome sequence of Chp3 was assembled using the Chp2 genome as a backbone. Sequence analysis was performed on both strands (on average three times in each direction). Appendix 2 shows the complete nucleotide sequence of the Chp3 genome.

## 3.5 Genomic Analysis of Chp3

The Chp3 genome is 4554 bp in length with a nucleotide composition of A(28.7%); C(18.4%); G(22.6%) and T(30.3%). The G+C content of Chp3 (41%) is similar to the other chlamydiaphages (37-40%), and identical to the host chlamydial genomes (41%) (Read *et al.*, 2000b).

### 3.5.1 Genomic Organisation of Chp3

The Chp3 ORFs were predicted using the Lasergene software (DNASTAR Inc. Madison, WI), which identified all of the Met-Ter ORFs greater than 99 nucleotides. The table below shows the results of the analysis; names have been assigned to the eight main ORFs, by analogy to Chp2 (figure 3.4). A small ORF was also identified within ORF4 (ORF9), which is also present in the other Chp2-like chlamydiaphages (table 3.3).

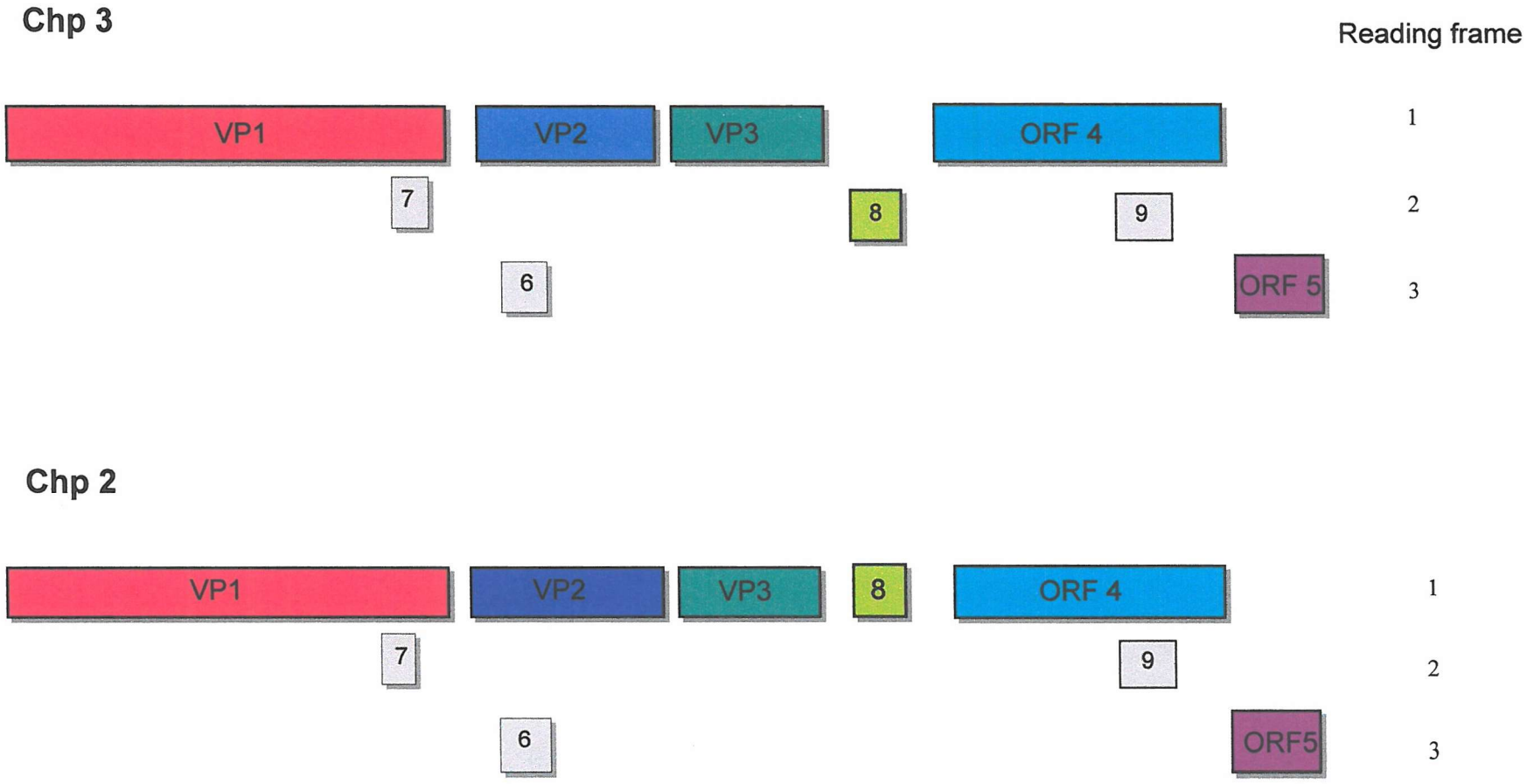



**Figure 3.4** A linear representation of the Chp2 and Chp3 Genomes.

**Table 3.3 Chp3 Open Reading Frames**

| Reading frame 1 | | |
|---|---|---|
| **ORF** | **Range** | **Nucl. Length** |
| ORF1 | 1-1698 | 1698 |
| ORF2 | 1804-2364 | 561 |
| ORF3 | 2404-2811 | 408 |
| ORF4 | 3316-4263 | 948 |

| Reading frame 2 | | |
|---|---|---|
| **ORF** | **Range** | **Nucl. Length** |
| ORF7 | 1406-1504 | 99 |
| ORF8 | 3011-3133 | 123 |
| ORF9 | 3998-4150 | 153 |

| Reading frame 3 | | |
|---|---|---|
| **ORF** | **Range** | **Nucl. Length** |
| ORF6 | 1791-1949 | 159 |
| ORF5 | 4293-4547 | 255 |

These ORFs all begin with an ATG translational initiation codon and are positioned in the same genomic context as the corresponding ORFs in Chp2, ϕCPAR39 and ϕCPG1. There are five untranslated regions in the Chp3 genome, the largest is between ORF3 and ORF8 (200nt) and the smallest is between ORF4 and ORF5 (30nt). It was possible to identify regions that could function as ribosome-binding sites upstream of all 9 ORFs so all could potentially code for proteins. The predicted ribosome-binding sites for each ORF are summarised in table 3.4. The ribosome-binding sites were identified in front of the Chp3 ORFs by identifying regions that were identical or similar to the well characterised ϕX174 ribosome-binding sites (gpA-GGAGG, gpA* GGAGG, gpB AGGAG, gpE, GAGG).

**Table 3.4 Chp3 Ribosome binding sites**

| Open reading frame codon | Predicted ribosome-binding site | Position of start |
|---|---|---|
| ORF1 | ACAAACAGTAGGGAAGAT**ATG**GTT | 1 |
| ORF2 | AAAAATGTGGAGAATTT**ATG**AAT | 1804 |
| ORF3 | AAGGAAGGTAG**ATG**TTT | 2365 |
| ORF4 | TTTAAATGAGGCATTA**ATG**ACG | 3316 |
| ORF5 | AATAGGAGCTTTTTTCA**ATG**AAA | 4293 |
| ORF6 | AGCTGAGGCGTAAAA**ATG**TGG | 1791 |
| ORF7 | TGGAGGAGAGATTGTGG**ATG**AGC | 1406 |
| ORF8 | AATGAGGTTAAAA**ATG**GCA | 3011 |
| ORF9 | GAATAGGAGCGGATTGGT**ATG**AGA | 3998 |

ATG start codons are in bold.
___, predicted ribosome-binding sites, sequences shown in green are similar to the ϕX174 RBS. Sequences shown in red are identical to the ϕX174 gpE RBS and sequences shown in blue are identical to the gpB RBS.

The predicted ribosome binding sites are conserved between Chp3, Chp2, ϕCPAR39 and ϕCPG1.

### 3.5.2 Genomic Comparisons

Comparisons were made between the genome sequences of Chp3, Chp2, Chp1, ϕCPG1 and ϕCPAR39, and between the amino acid sequences of the predicted proteins encoded by the nine major ORFs. All alignments were carried out using the Jotun Hein method, which is most commonly used for performing multiple alignments between related sequences (Hein, 1990).

Alignment between the chlamydiaphage genomes revealed a high nucleotide sequence identity between Chp2, Chp3, ϕCPAR39 and ϕCPG1 (92.7-97.1%) (table 3.5).

**Table 3.5 Nucleotide sequence identity (%) between the ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3 genomes**

| | Chp1 | Chp2 | ϕCPAR39 | ϕCPG1 |
|---|---|---|---|---|
| Chp3 | 52.4 | 97.1 | 93.4 | 93.3 |
| Chp1 | | 52.1 | 52.2 | 52.6 |
| Chp2 | | | 92.7 | 93.3 |
| ϕCPAR39 | | | | 97.1 |

The nucleotide sequences of the genomes of Chp3/Chp2 and ϕCPAR39/ϕCPG1 are highly conserved (97.1%). The genome of Chp3 is only 9nt shorter than the genome of Chp2 this is attributable to an 11nt deletion situated in a non-coding region of the genome 4nt downstream of the Chp3 ORF3 terminator codon (position nt2815) and a 2nt insertion situated downstream of ORF8 (position nt3235). The 11nt deletion causes a reading frame-shift in the Chp3 ORF8 compared to the Chp2 ORF8 from reading frame 1 to reading frame 2. The 2nt insertion after ORF8 restores the reading frame usage as for the Chp2 ORF4. The genomes of ϕCPAR39 and ϕCPG1 differ by 3nt due to there being 6nt insertions in the genome of ϕCPAR39 (positions nt3178-9 and nt3276-9) and only 3nt insertions in ϕCPG1 (positions nt2986 and nt3326-7). The reading frame of ORF8 is also different between ϕCPAR39 and ϕCPG1 moving from frame 1 to 3 due to the nucleotide insertion at position nt2986 in ϕCPG1. The 2nt insert at nt3326-7 restores the reading frame usage as for the ϕCPAR39 ORF4.

Chp1 although exhibiting a similar genomic organisation to the other chlamydiaphages, shows divergence throughout its genome (52.1-52.6% nucleotide sequence identity) and seems to form a distinct group within the chlamydiaphages.

### 3.5.2.1 VP1

Based on a region of homology between the VP1 of Chp1 and the major capsid protein of ϕX174 (figure 3.5) this protein has been identified as the virus major coat protein (Storey *et al.*, 1989b). Table 3.6 shows the results from an amino acid sequence alignment between VP1 of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3. A high amino acid sequence identity (84.8-99.5%) occurs in this region between the Chp2-like group of chlamydiaphages (Chp2, Chp3, ϕCPG1 and ϕCPAR39).

**Table 3.6 Amino acid sequence identity (%) between the VP1 of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 53.2 | 52.5 | 53.3 | 53.5 |
| **Chp2** | | 98.1 | 85.2 | 85.2 |
| **Chp3** | | | 85.2 | 84.8 |
| **ϕCPAR39** | | | | 99.5 |

The alignments of VP1 revealed two areas of divergence situated between amino acids 216 to 299 (IN5 loop) and 462-467 (Ins) (figure 3.6) (Read *et al.*, 2000b). It has been hypothesised that the larger of these two regions, the IN5 loop, is surface exposed and forms mushroom-like protrusions on the virion surface (Chipman *et al.*, 1998). The second smaller region, Ins, is also considered to be surface exposed and has been predicted to be situated in close proximity and may interact with the IN5 loop (Read *et al.*, 2000b). One explanation for the role of the IN5 loop is that it acts as the chlamydiaphage receptor-recognition domain. Interestingly, the IN5 loop and Ins of ϕCPG1/ϕCPAR39 and Chp3/Chp2 are very similar. The VP1 of Chp1 is highly divergent to other chlamydiaphages (52.5-53.5%).

**Figure 3.5** Amino acid alignment of a region in the Chp1, Chp2, Chp3, ϕCPG1 and ϕCPAR39 major capsid proteins that shows homology with a region in the ϕX174 major capsid protein (F protein). The amino acid residues that are conserved between the ϕX174 F protein and the chlamydiaphage VP1 are indicated in red type. Gaps between the sequences are indicated by green dashes.

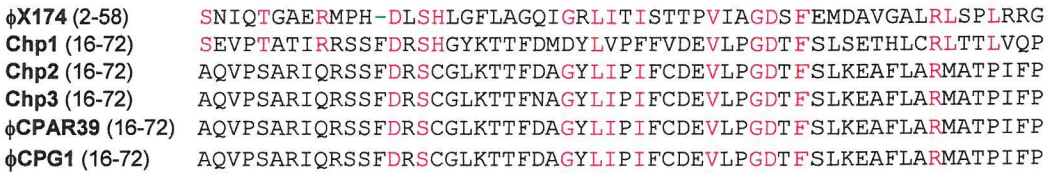
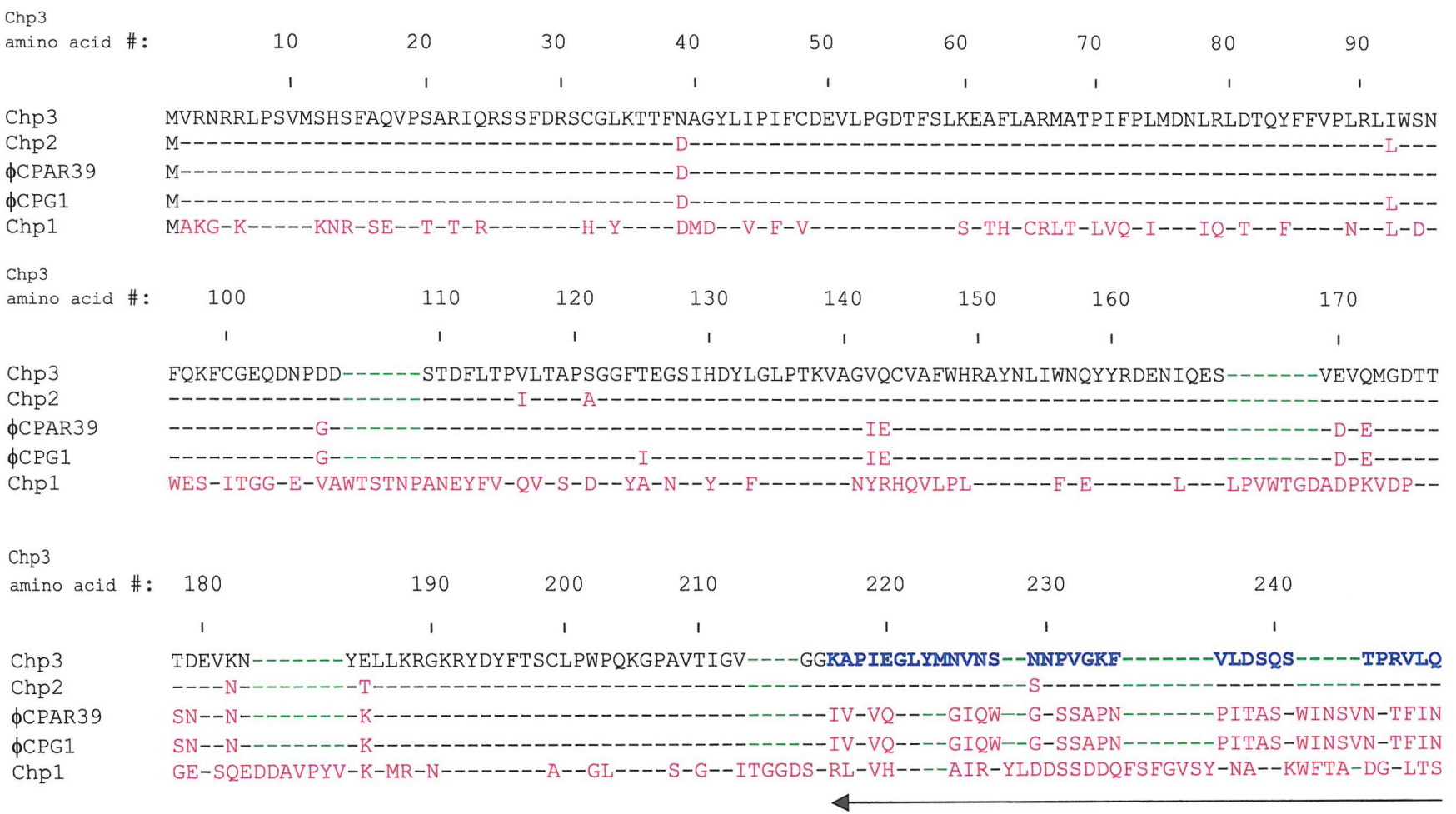
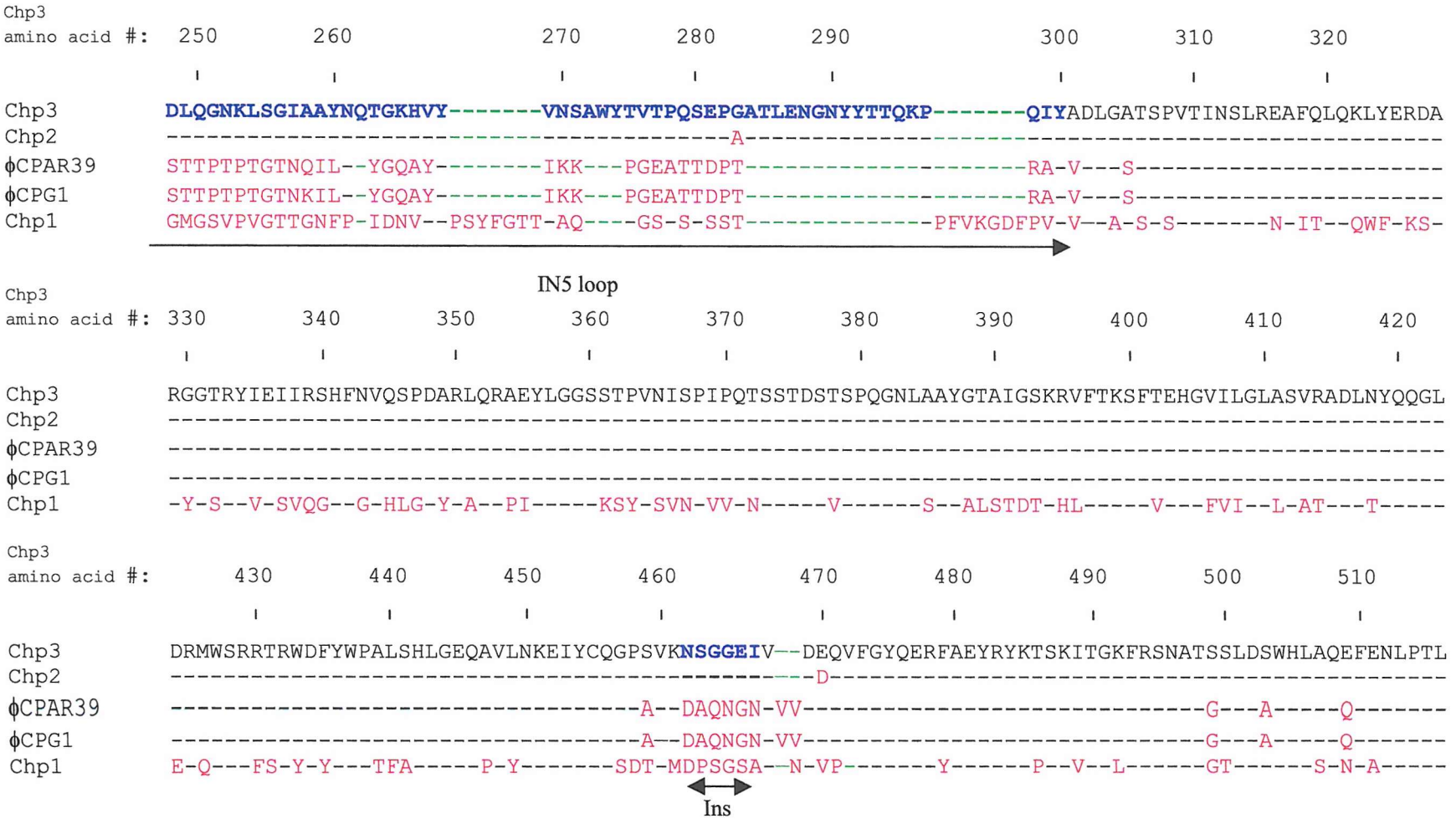
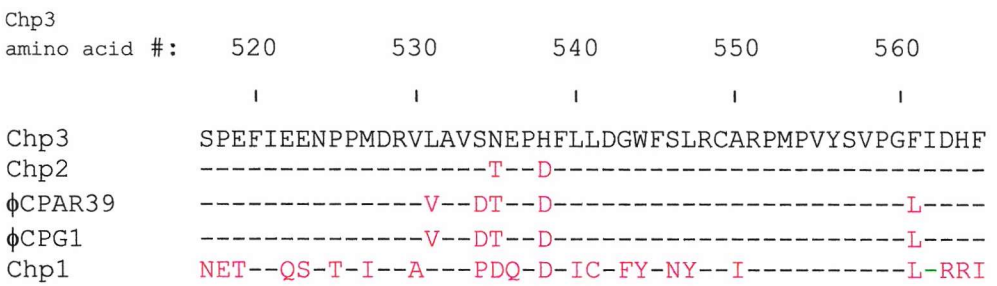


**Figure 3.6** Amino acid sequence alignment between the chlamydiaphage VP1. Amino acids residues divergent to the Chp3 VP1 are indicated. Gaps between the amino acid sequences are indicated by green dashes. Conserved residues are indicated by black dashes. The amino acid residues corresponding to the IN5 loop and Ins regions are shown in blue in the Chp3 amino acid sequence.




Chp3
amino acid #: 250 260 270 280 290 300 310 320
Chp3 DLQGNKLSGIAAYNQTGKHVY-------VNSAWYTVTPQSEPGATLENGNYYTTQKP-------QIYADLGATSPVTINSLREAFQLQKLYERDA
Chp2 ----------A----------
ϕCPAR39 STTPTPTGTNQIL--YGQAY--------IKK---PGEATTDPT----------RA-V---S----------
ϕCPG1 STTPTPTGTNKIL--YGQAY--------IKK---PGEATTDPT----------RA-V---S----------
Chp1 GMGSVPVGTTGNFP-IDNV--PSYFGTT-AQ----GS-S-SST----------PFVKGDFPV-V--A-S-S-------N-IT--QWF-KS-
IN5 loop
Chp3
amino acid #: 330 340 350 360 370 380 390 400 410 420
Chp3 RGGTRYIEIIRSHFNVQSPDARLQRAEYLGGSSTPVNISPIPQTSSTDSTSPQGNLAAYGTAIGSKRVFTKSFTEHGVILGLASVRADLNYQQGL
Chp2 ----------
ϕCPAR39 ----------
ϕCPG1 ----------
Chp1 -Y-S--V-SVQG--G-HLG-Y-A--PI-----KSY-SVN-VV-N-----V------S--ALSTDT-HL-----V---FVI--L-AT---T-----
Chp3
amino acid #: 430 440 450 460 470 480 490 500 510
Chp3 DRMWSRRTRWDFYWPALSHLGEQAVLNKEIYCQGPSVKNSGGEIV--DEQVFGYQERFAEYRYKTSKITGKFRSNATSSLDSWHLAQEFENLPTL
Chp2 ----------D----------
ϕCPAR39 ----------A--DAQNGN-VV----------G---A-----Q-------
ϕCPG1 ----------A--DAQNGN-VV----------G---A-----Q-------
Chp1 E-Q---FS-Y-Y---TFA-----P-Y-------SDT-MDPSGSA--N-VP-------Y------P--V--L------GT------S-N-A-----
Ins


Chp3
amino acid #: 520 530 540 550 560
Chp3 SPEFIEENPPMDRVLAVSNEPHFLLDGWFSLRCARPMPVYSVPGFIDHF
Chp2 ----------T--D----------
ϕCPAR39 ----------V--DT--D----------L----
ϕCPG1 ----------V--DT--D----------L----
Chp1 NET--QS-T-I--A---PDQ-D-IC-FY-NY--I----------L-RRI

### 3.5.2.2 VP2

VP2 is thought to be the equivalent of the ϕX174 protein H (Liu *et al.*, 2000). This was determined by examining the Chp1 structural proteins that had been separated by SDS PAGE and analysed using NIH Image technology. By assuming that there are 60 copies of VP1 per virion it was approximated that there are 15 copies of VP2. This is consistent with the ratio of H proteins present per virion in the coliphage ϕX174. In addition it was found that VP2 has a feature in common with DNA pilot proteins in that it is glycine rich (Liu *et al.*, 2000). The ϕX174 H protein is present in the mature virion, it is required for DNA penetration and guiding the DNA into the cell (Jazwinski *et al.*, 1975). VP2 is highly conserved (98.4-99.5% amino acid sequence identity) between the Chp3, Chp2, ϕCPAR39, and ϕCPG1 (table 3.7).

**Table 3.7 Amino acid sequence identity (%) between the VP2 of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 40.4 | 39.8 | 40.3 | 40.3 |
| **Chp2** | | 98.4 | 98.9 | 99.5 |
| **Chp3** | | | 98.4 | 98.9 |
| **ϕCPAR39** | | | | 99.5 |

There are only 4 amino acid residue changes in the Chp2, ϕCPAR39 and ϕCPG1 VP2 compared to the Chp3 VP2 (figure 3.7). One change in ϕCPAR39 from asparagine to serine located at residue 8. One change in Chp2 at residue 153 from valine to isoleucine and two changes that occurs in Chp2, ϕCPAR39 and ϕCPG1 from isoleucine to leucine and arginine to lysine at residues 170 and 175 respectively. The VP2 of Chp1 is highly divergent to other chlamydiaphages (39.8-40.3% amino acid sequence identity).


**Figure 3.7** Amino acid sequence alignment between the chlamydiaphage VP2. Amino acids residues divergent to the Chp3 VP2 are indicated. Conserved residues are indicated by black dashes. Gaps between the amino acid sequences are shown by green dashes.

### 3.5.2.2 VP3

It has been predicted that VP3 is the chlamydiaphage internal scaffolding protein based on conservation between the internal scaffolding-coat protein interactions in the atomic structure of the ϕX174 procapsid and proposed Chp2 VP3-coat interactions (Liu *et al.*, 2000). High levels of amino acid sequence identity (96.6-98.7%) occur between the VP3 of Chp2, Chp3, ϕCPAR39 and ϕCPG1 (table 3.8).

**Table 3.8 Amino acid sequence identity (%) between the VP3 of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 30.2 | 30.9 | 30.9 | 30.2 |
| **Chp2** | | 98 | 98.7 | 97.3 |
| **Chp3** | | | 98 | 96.6 |
| **ϕCPAR39** | | | | 98.7 |

There are only 6 amino acid residue changes in the Chp2, ϕCPG1 and ϕCPAR39 VP3 compared to the Chp3 VP3 (figure 3.8). The first is in Chp2 at residue 62 from an methionine to isoleucine, the second is in Chp2, ϕCPAR39 and ϕCPG1 at residue 85, from isoleucine to valine. The third change is in ϕCPG1 at residue 125, from glutamine to histidine, the fourth is in Chp2, ϕCPG1 and ϕCPAR39 at residue 127 from leucine to phenylalanine. The fifth is in ϕCPG1 at residue 128, from phenylalanine to cysteine, and finally the sixth one occurs in both ϕCPG1 and ϕCPAR39 at residue 131, from glycine to aspartic acid. The lowest amino acid sequence identity (30.2-30.9%) can be observed between the Chp1 VP3 and the other chlamydiaphages VP3 (table 3.8).

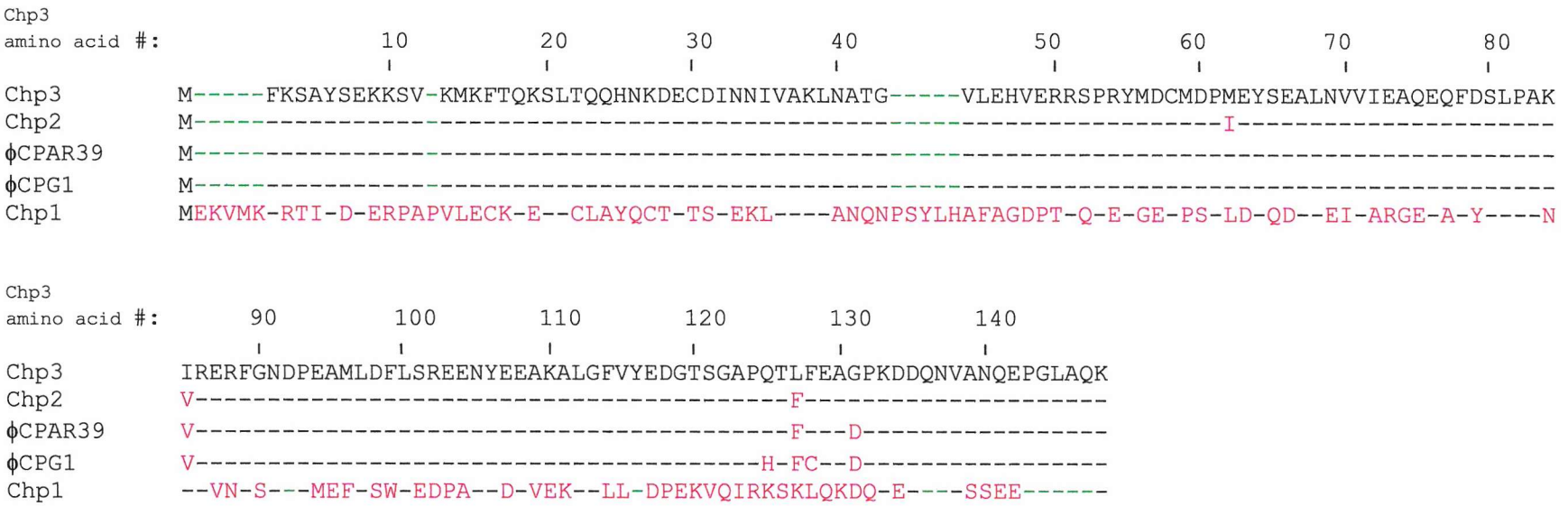



**Figure 3.8** Amino acid sequence alignment between the chlamydiaphage VP3. Amino acids residues divergent to the Chp3 VP3 are indicated. Conserved residues are indicated by black dashes. Gaps between the amino acid sequences are shown by green dashes.

### 3.5.2.4 ORF4 Protein

The protein encoded by ORF4 has two regions of homology with the C terminus of the ϕX174 gene A protein (Storey *et al.*, 1989b) (figure 3.9). The ϕX174 A protein is a site-specific endonuclease that initiates and terminates the ϕX174 rolling circle RF DNA replication. It first nicks the RF DNA at the origin of replication, which creates a free 3'-OH primer for DNA synthesis, following DNA replication the A protein cleaves the newly synthesised DNA at the regenerated origin and re-circularises it before releasing the newly synthesised viral DNA (Tessman, 1966; Langeveld *et al.*, 1978). It is likely that the ORF4 products have similar roles in chlamydiaphage infected RBs to A protein. In ϕX174-infected *E. coli*, the A gene product occurs in two forms termed A and A*. A* is expressed from a downstream in-frame initiation codon and is thus a truncated version of A (Linney *et al.*, 1972). The main role of A* protein is shut down of stage II replication and host chromosomal DNA synthesis (Ishiwa and Tessman, 1968). No A* protein equivalent has been identified in the chlamydiaphages

The amino-terminus of the protein encoded by chlamydiaphage ORF4 contains a variety of potential translational start codons. This has made it difficult to define the translational initiation codon (figure 3.10). The Chp2 ORF4 encodes the largest product with an extra 21 amino terminal amino acids compared to Chp3, ϕCPAR39 and ϕCPG1 which all share a primary initiation codon predicting a 315 amino acid product. However, in ϕCPG1 this is followed by an in-frame stop codon. The next potential translational start codon between the bacteriophages is at nt3447 in ϕCPG1 however, this is replaced by AAG (Lys) in Chp3. The next potential initiation codon is at nt3535 in Chp3 but this is replaced by ATC (Ile) in ϕCPG1. At nt3621 in ϕCPG1 all share a fourth potential translation initiation codon but in Chp3 this is AGG (Arg). Thus it is difficult to define a common translational initiation codon for the ORF4 product and it's A* equivalent because the initiation codons are not shared amongst the chlamydiaphages until nt3820 in Chp3. This would seem an unlikely candidate for the primary product as it would yield a protein of only 148 amino acids.

**Figure 3.9** Amino acid alignment of two regions in the Chp1, Chp2, Chp3, ϕCPG1 and ϕCPAR39 ORF4 proteins that shows homology with regions in the ϕX174 A protein. The amino acid residues that are conserved between the ϕX174 A protein and the chlamydiaphage ORF4 proteins are indicated in red type. Gaps between the sequences are indicated by green dashes.

a)

| | |
|---|---|
| **ϕX174** (248-302) | YFCVPEYGTANGRLHFHAVHFM-RTLPTGSVDNF--GR-RVRNRRQLNSLQNTWPYG |
| **Chp1** (181-237) | FLTVGEYGDKKGRMHWHMIVFGWKPKSEEQLEPYLGGKYRTDVRYRSRKLKELWKFG |
| **Chp2** (124-175) | YFGCGEYGSKLQRPHYHLLIYN-YDFPDKKL---LSKK-RGNPLFVSEKLMRLWPFG |
| **Chp3** (124-175) | YFGCGEYGSKLQRPHYHLLIYN-YDFPDKKL---LSKK-RGNPLFVSEKLMQLWPFG |
| **ϕCPAR39** (124-175) | YFGCGEYGSKLQRPHYHLLIYN-YDFPDKKL---LSKK-RGNPLFVSEKLMRLWPFG |
| **ϕCPG1** (72-123) | YFGCGEYGSKLQRPHYHLLIYN-YDFPDKKL---LSKK-RGNPLFVSEKLMRLWPFG |

b)

| | |
|---|---|
| **ϕX174** (342-351) | FYVAKYVNKK |
| **Chp1** (250-259) | FYVARYVQKK |
| **Chp2** (209-218) | GYVARYSLKK |
| **Chp3** (187-197) | GYVARYSLKK |
| **ϕCPAR39** (187-197) | GYVARYSLKK |
| **ϕCPG1** (136-145) | GYVARYSLKK |

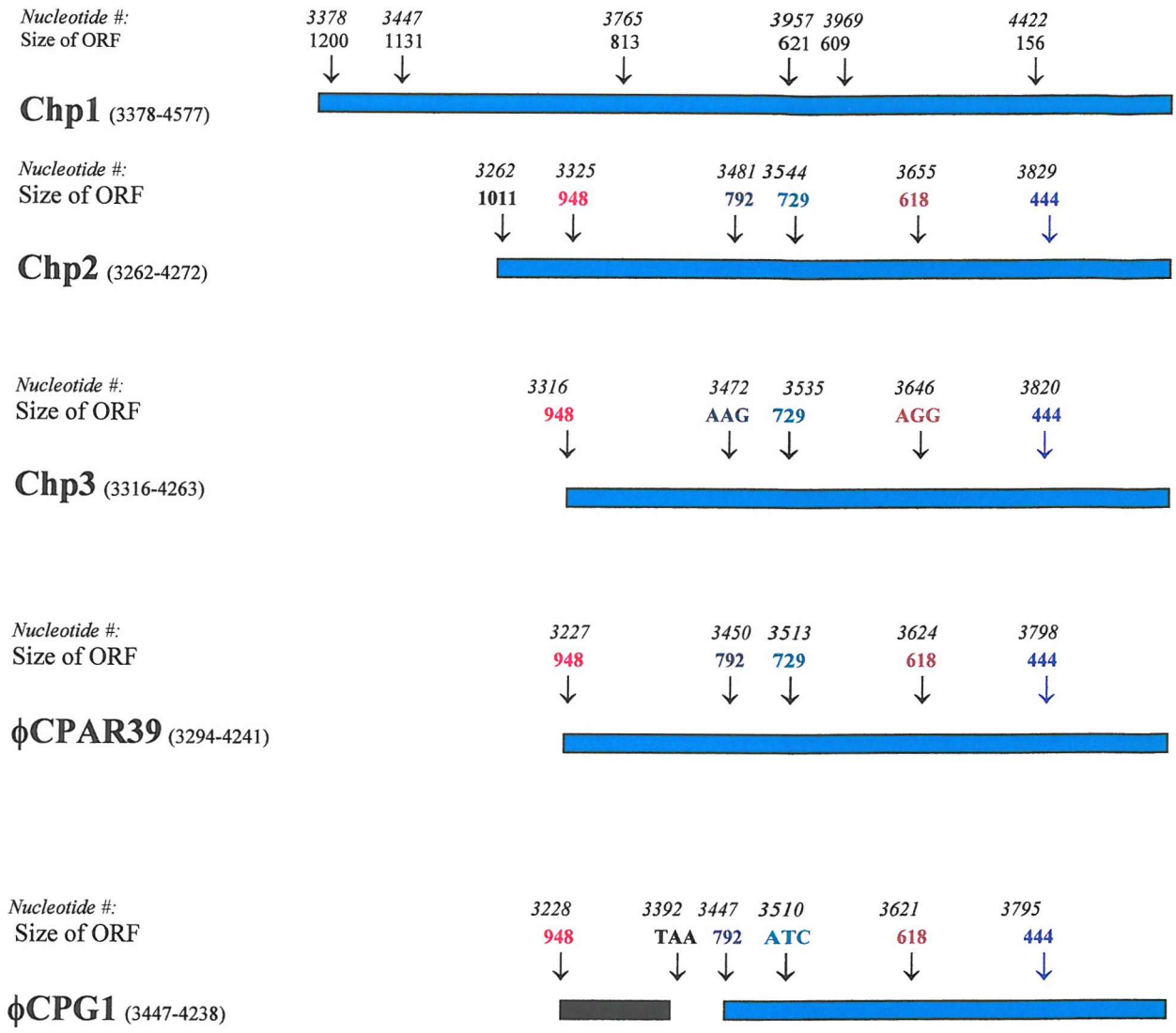



**Figure 3.10** A diagrammatic representation of the predicted translation initiation codons in the chlamydiaphage ORF4. All of the predicted ATG translation initiation codons upstream of the first conserved ATG codon and the size of the subsequent ORF are indicated.

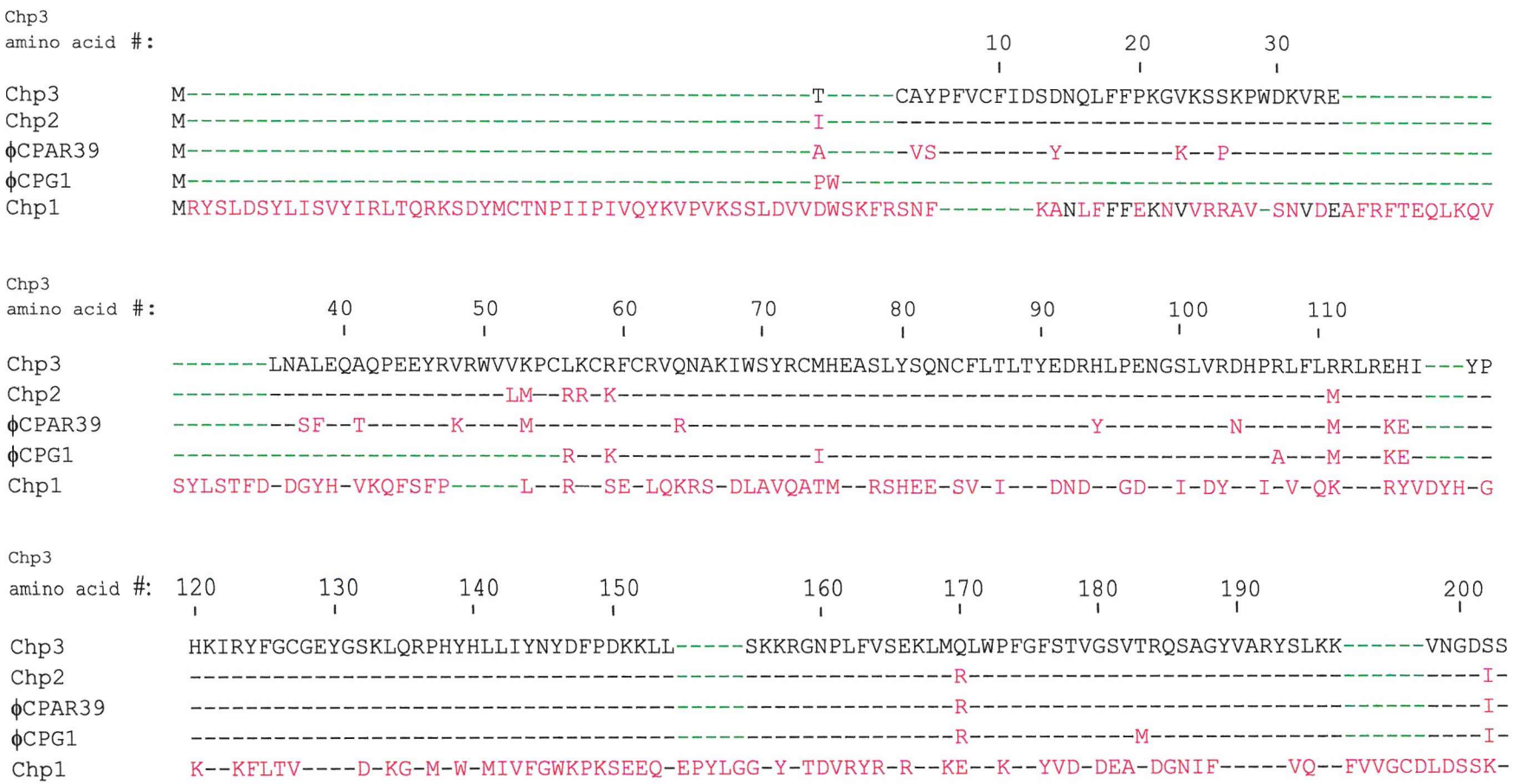
It is therefore likely that either translation of ORF4 protein does not initiate with a methionine residue in all four chlamydiaphages or that the size of the ORF4 protein is not conserved between these bacteriophages. For the task of carrying out comparisons between the amino acid sequences of the ORF4 protein I have assumed that the ORF4 protein of Chp2, ϕCPAR39 and Chp3 is a 315 amino acid residue protein and that the ORF4 protein of ϕCPG1 is 51 residues shorter.

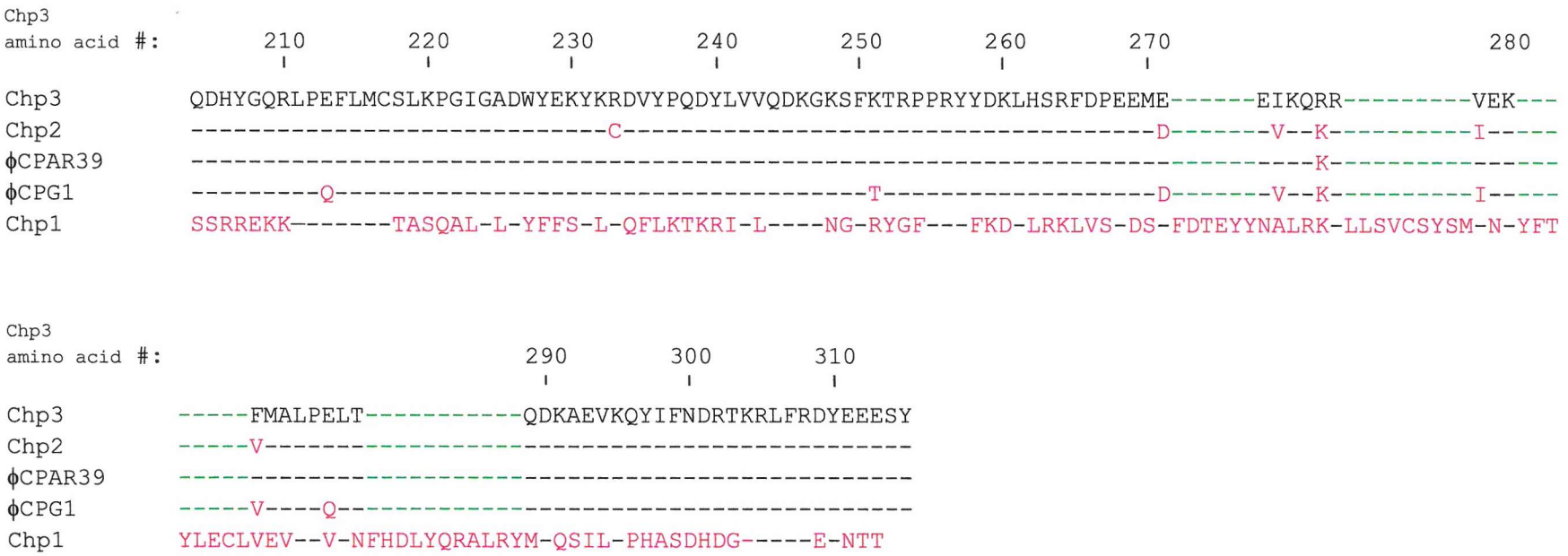
The highest amino acid sequence identity between the proteins encoded by ORF4 (as shown in table 3.9) is observed between the Chp2 and ϕCPG1 ORF4 proteins (95.8%) followed by the Chp2 and Chp3 ORF4 proteins (95.3%). The ORF4 protein of Chp1 has an amino terminus that is extended by approximately 80 amino acids and is highly divergent from the other chlamydiaphages. Figure 3.11 show the results of an alignment between the chlamydiaphage ORF4 proteins.

**Table 3.9 Amino acid sequence identity (%) between the ORF4 proteins of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 30.1 | 30.3 | 29.4 | 30.6 |
| **Chp2** | | 95.3 | 92.1 | 95.8 |
| **Chp3** | | | 93.4 | 93.2 |
| **ϕCPAR39** | | | | 94.3 |

**Figure 3.11** Amino acid sequence alignment of the chlamydiaphage ORF4 protein. Amino acids residues divergent to the Chp3 ORF4 protein are indicated. Conserved residues are indicated by black dashes. Gaps in the amino acid sequences are indicated by green dashes.




Chp3
amino acid #:
210
220
230
240
250
260
270
280
Chp3 QDHYGQRLPEFLMCSLKPGIGADWYEKYKRDVYPQDYLVVQDKGKSFKTRPPRYYDKLHSRFDPEEME------EIKQRR---------VEK---
Chp2 ----------C----------D-------V--K----------I-----
ϕCPAR39 ----------K----------
ϕCPG1 ---------Q----------T----------D-------V--K----------I-----
Chp1 SSRREKK-------TASQAL-L-YFFS-L-QFLKTKRI-L----NG-RYGF---FKD-LRKLVS-DS-FDTEYYNALRK-LLSVCSYSM-N-YFT
Chp3
amino acid #:
290
300
310
Chp3 -----FMALPELT----------QDKAEVKQYIFNDRTKRLFRDYEEESY
Chp2 -----V----------
ϕCPAR39 ----------
ϕCPG1 -----V----Q----------
Chp1 YLECLVEV--V-NFHDLYQRALRYM-QSIL-PHASDHDG-----E-NTT

### 3.5.2.5 ORF5 Protein

The function of the ORF5 protein is unknown and no significant amino acid homologies exist between the ORF5 protein and proteins encoded by the ϕX174 genome. Comparison with the ϕX174 proteins shows that the only protein of similar size, charge and composition is protein C (table 3.10). The gene encoding protein C is also situated in a similar genomic location to ORF5 (between ORF4/gpA and ORF1/gpF).

**Table 3.10 Comparison between the protein composition of the ϕX174 protein C and the Chp3 ORF5 protein.**

| | ϕX174 protein C | Chp3 ORF5 protein |
|---|---|---|
| Size (amino acid residues) | 86 | 84 |
| Strongly basic residues (+) | 14 | 7 |
| Strongly acidic residues (-) | 12 | 12 |
| Hydrophobic residues | 30 | 29 |
| Polar amino acids | 20 | 21 |

Protein C of ϕX174 is a ssDNA binding protein that is required for initiation of viral strand DNA synthesis (stage III DNA replication) (Fujisawa and Hayashi, 1977; Aoyama *et al.*, 1983). Stage III DNA replication involves the synthesis of the viral strand DNA directly into the procapsid and requires the present of protein C, protein A, procapsid and the host cell helicase (rep protein) (Aoyama and Hayashi, 1986).

The ORF5 proteins are highly conserved between the Chp2-like bacteriophages (97.6-100% amino acid sequence identity) as shown in table 3.11.

**Table 3.11 Amino acid sequence identity (%) between the ORF5 proteins of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 32.9 | 32.9 | 34.1 | 33.3 |
| **Chp2** | | 100 | 98.8 | 97.6 |
| **Chp3** | | | 98.8 | 97.6 |
| **ϕCPAR39** | | | | 98.8 |

The Chp2 and Chp3 ORF5 proteins are identical and there are only two amino acids changes between the ϕCPAR39 and ϕCPG1 ORF5 protein compared to the Chp3 ORF5 protein (figure 3.12). The first occurs in both the ϕCPAR39 and ϕCPG1 ORF5 proteins at residue 39, from threonine to serine and the second occurs in the ϕCPG1 ORF5 protein at residue 64, from threonine to asparagine. The lowest amino acid sequence identity 32.9-34.1% can be seen between the Chp1 ORF5 protein and the other chlamydiaphages.

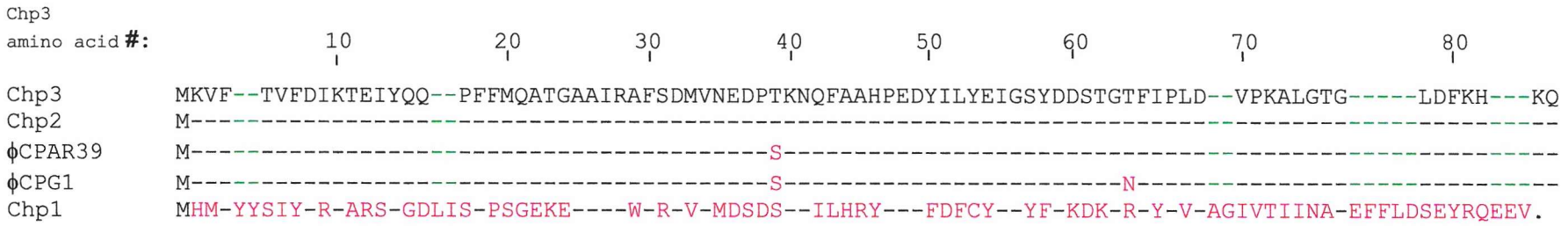



**Figure 3.12** Amino acid sequence alignment of the chlamydiaphages ORF5 protein. Amino acid residues divergent to the Chp3 ORF5 protein are indicated. Conserved residues are indicated by black dashes. Gaps between the amino acid sequences are shown by green dashes.

### 3.5.2.6 ORF6 Protein

ORF6 encodes identically sized polypeptides of 52 residues in Chp3, Chp2, ϕCPAR39 and ϕCPG1. There are no proteins with significant sequence similarity to ORF6 protein in ϕX174. Table 3.12 shows the results of comparisons between the amino acid sequences of the chlamydiaphage ORF6 protein. The amino acid comparisons revealed a high sequence identity between the ORF6 proteins of Chp3, ϕCPAR39, ϕCPG1 and Chp2 (88.7-92.5%) (figure 3.13). The Chp1 ORF6 protein shows the lowest amino acid identity (33.3-37.8%). However, this is expected as the Chp1 ORF6 is situated in a different genomic location than the Chp2-like bacteriophages ORF6 (the Chp1 ORF6 is situated at the 3' terminus of ORF2, and the Chp2-like bacteriophages ORF6 is situated at the 5' terminus of ORF2) and is larger is size.

**Table 3.12 Amino acid sequence identity (%) between the ORF6 proteins of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp1** | 37.8 | 33.3 | 33.3 | 37.0 |
| **Chp2** | | 88.7 | 92.5 | 92.5 |
| **Chp3** | | | 88.7 | 90.6 |
| **ϕCPAR39** | | | | 92.5 |

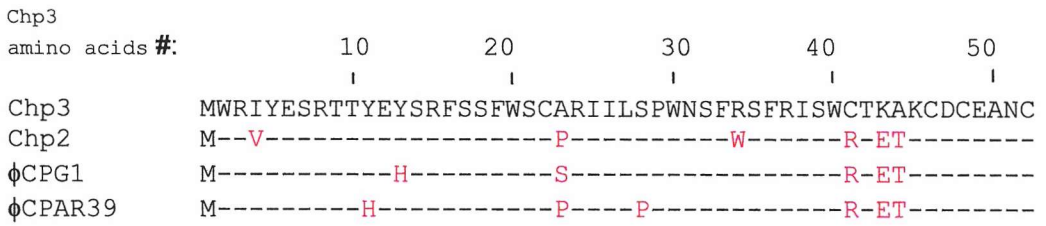



**Figure 3.13** Amino acid sequence alignment between the chlamydiaphages ORF6 protein. Amino acids residues divergent to the Chp3 ORF6 protein are indicated. Conserved residues are indicated by black dashes. The Chp1 ORF6 protein is highly divergent to the Chp2-like bacteriophages and was therefore not included in this alignment.

### 3.5.2.7 ORF7 Protein

The function of the protein encoded by ORF7 is unknown and no significant homologies exist between the ORF7 protein and proteins encoded by the ϕX174 genome. Secondary structure and hydrophobicity analyses predict a hydrophobic domain in the N-terminus of the ORF7 proteins from Chp1, ϕCPAR39 and ϕCPG1 that is absent from the Chp2 and Chp3 ORF7 proteins. This is because the amino-terminus of the proteins encoded by the ϕCPG1 and ϕCPAR39 ORF7 contain three additional potential start codons than the Chp2 and Chp3 ORF7 proteins (figure 3.14). This is due to a 6nt deletion in the nucleotide sequence of Chp2 and Chp3 situated in the predicted ATG translational start codon region of the ϕCPG1 and ϕCPAR39 ORF7.

The ϕX174 lysis protein (protein E) also shares this feature thus ORF7 protein might be the chlamydiaphage lysis protein. Protein E has a hydrophobic amino-terminal domain and a positively charged carboxyl terminal domain. The genes encoding microvirus lysis proteins are also embedded within a major morphogenesis gene using an alternative reading frame, similarly, ORF7 is located within gene encoding the major capsid protein VP1 in reading frame 2.

Table 3.13 shows the amino acid sequence identity between the chlamydiaphage ORF7 proteins. A high amino acid sequence identity can be observed between Chp2-like phages ORF7 protein (87.9-97.6%). The lowest amino acid sequence identity is seen between the Chp1 ORF7 protein and ORF7 proteins of the Chp2-like phages (45.5-57.1%).

**Table 3.13** **Amino acid sequence identity (%) between the ORF7 proteins of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | Chp1 | Chp2 | Chp3 | ϕCPAR39 | ϕCPG1 |
|---|---|---|---|---|---|
| **Chp1** | | 45.5 | 45.5 | 53.6 | 57.1 |
| **Chp2** | | | 97 | 87.9 | 90.9 |
| **Chp3** | | | | 90.9 | 93.9 |
| **ϕCPAR39** | | | | | 97.6 |

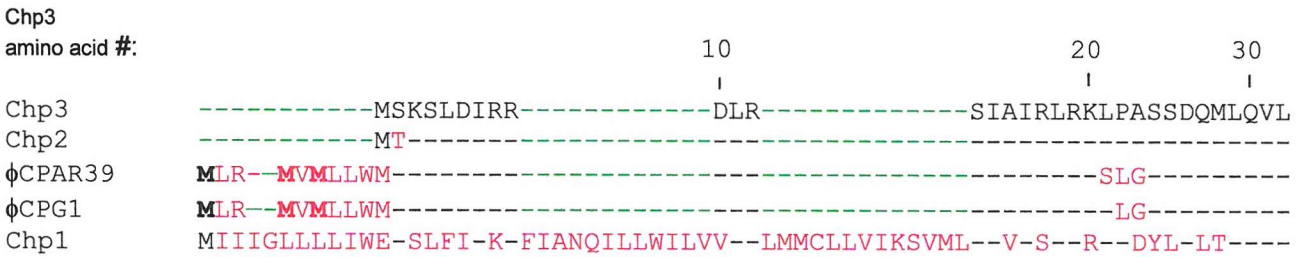



**Figure 3.14** Amino acid sequence alignment of the chlamydiaphages ORF7 protein. Amino acid residues divergent to the Chp3 ORF7 protein are indicated. Conserved residues are indicated by black dashes. Gaps between the amino acid sequences are shown by green dashes. The three additional potential start codons present in the ϕCPAR39 and ϕCPG1 ORF7 protein sequence are in bold.

### 3.5.2.8 ORF8 Protein

The ORF 8 protein is thought to be the DNA binding protein by analogy to the ϕX174 J protein (Liu *et al.*, 2000). Protein J is a small basic protein that is required for the condensation of the viral DNA before packaging into the prohead (Freymeyer *et al.*, 1977). The ORF8 protein is identical between Chp2, Chp3, ϕCPAR39 and ϕCPG1 (table 3.14). This suggests conservation of this protein, maybe due to it binding identical DNA regions in the bacteriophage genome. An amino acid sequence identity of 67.6% can be observed between the ORF8 protein of Chp1 compared to the ORF8 proteins of the other chlamydiaphages. Figure 3.15 shows an alignment between the Chp1 ORF8 protein and the Chp2-like bacteriophages ORF8 protein.

**Table 3.14 Amino acid sequence identity (%) between the ORF8 proteins of ϕCPG1, ϕCPAR39, Chp1, Chp2 and Chp3**

| | **Chp1** | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|---|
| **Chp1** | | 67.6 | 67.6 | 67.6 | 67.6 |
| **Chp2** | | | 100 | 100 | 100 |
| **Chp3** | | | | 100 | 100 |
| **ϕCPAR39** | | | | | 100 |

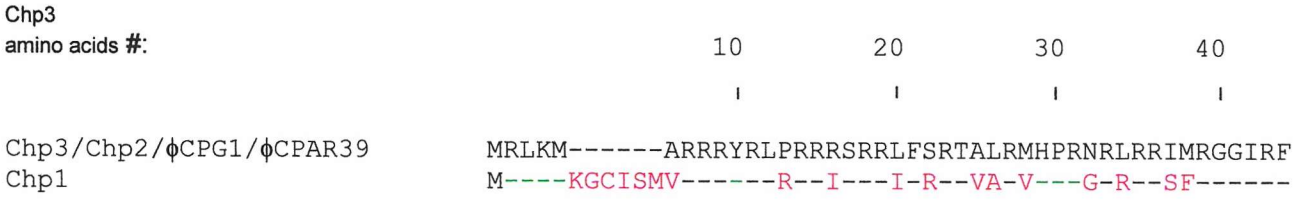



**Figure 3.15** Amino acid sequence alignment with the chlamydiaphages ORF8 protein. Amino acid residues divergent to the Chp3 ORF8 protein are indicated. Conserved residues are indicated by black dashes. Gaps between the sequences are indicated by green dashes.

### 3.5.2.9 ORF9 Protein

ORF9 encodes identically sized polypeptides of 50 residues in Chp3, Chp2, ϕCPAR39 and ϕCPG1. There are no proteins encoded in the ϕX174 genome with significant sequence similarity to the ORF9 protein. Table 3.15 shows the results of comparisons between the amino acid sequences of the ORF9 proteins. There is a high amino acid sequence identity 92.3-100% between the ORF9 proteins of Chp2, ϕCPAR39, ϕCPG1 and Chp2 (the ϕCPG1 and Chp2 ORF9 proteins are identical). There is no ORF situated in the same genomic location in the Chp1 genome that could potentially encode a Chp1 ORF9 protein. Figure 3.16 show an alignment between the chlamydiaphage ORF9 protein.

**Table 3.15 Amino acid sequence identity (%) between the ORF9 proteins of ϕCPG1, ϕCPAR39, Chp2 and Chp3**

| | **Chp2** | **Chp3** | **ϕCPAR39** | **ϕCPG1** |
|---|---|---|---|---|
| **Chp2** | | 92.3 | 94.2 | 100 |
| **Chp3** | | | 94.2 | 92.3 |
| **ϕCPAR39** | | | | 94.2 |

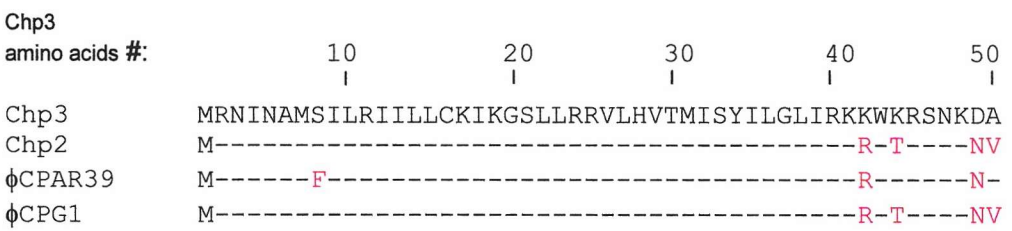



**Figure 3.16** Amino acid sequence alignment with the chlamydiaphages ORF9 protein. Amino acid residues divergent to the Chp3 ORF9 protein are indicated. Conserved residues are indicated by black dashes.

## 3.6 Host range of chlamydiaphages

The ability of a bacteriophage to infect a host is dependent on attachment and penetration into the host bacterium followed by replication and the production and release of active virions. It has been shown that chlamydiaphages bind to a protein receptor on the chlamydial surface (Everson *et al.*, 2002). This was determined by incubating EBs and RBs with phospholipase C, sodium periodate and proteinase K then assaying the binding of Chp2 to the EBs and RBs via ELISA using monoclonal antibodies 40 and 55. The results showed that Chp2 does not bind *C. abortus* after treatment with proteinase K indicating that the Chp2 receptor is composed of protein.

It has been predicted that two discrete regions situated in VP1 (IN5 and Ins) encodes the chlamydiaphage receptor-recognition domain. The IN5 loop is the largest of these two regions and it has been predicted that it forms mushroom-like protrusions on the capsid surface (Chipman *et al.*, 1998; Read *et al.*, 2000b). As these two regions are very similar between Chp2/Chp3 and ϕCPG1/ϕCPAR39 it has been hypothesised that the host ranges of Chp2 and Chp3 would be identical as would the host ranges of ϕCPAR39 and ϕCPG1, but that these two sets of chlamydiaphages would have different host ranges.

The host range of Chp3 was determined as follows: a standard inoculum of Chp3 that was capable of infecting >99% of *C. pecorum* by inclusion staining was prepared. Chlamydial EBs were incubated with Chp3, then cell monolayers were infected with the inoculum (Everson et al,. 2003). Thirty-nine different strains of chlamydiae were challenged with chlamydiaphage including a representative from all nine of the presently recognised species of the *Chlamydiaceae.* The strains challenged with Chp3 can be seen in table 3.16. After infection (48-72 hours) the cell monolayers infected with chlamydiae were screened by IF for the presence of bacteriophage using monoclonal 55 (specific to the chlamydiaphage VP1) (Everson *et al.*, 2002).

It was shown that Chp3 is able to infect *C. abortus*, *C. pecorum*, *C. felis* and *C. caviae*, but is unable to infect *C. pneumoniae* or any members of the *Chlamydia* genus of *Chlamydiaceae*.

Similarly, the host ranges of Chp2 (Everson *et al.*, 2002) and ϕCPAR39 (Everson *et al.*, 2002) have been determined by challenging the same thirty-nine strains of chlamydiae with each bacteriophage (table 3.16). It was shown that Chp2 is able to infect *C. abortus*, *C. felis* and *C. pecorum*. Chp2 has also recently been shown to infect *C. caviae* species. This was not observed in the initial host range studies, as the *C. caviae* VR813 strain used was highly sensitive to Chp2 infection and hence the Chp2 inoculum caused complete destruction of *C. caviae* before inclusions could develop. A lower titre of Chp2 inoculum was subsequently used to infected *C. caviae* and this showed that *C. caviae* could support the replication of Chp2. Chp2 does not infect *C. psittaci, C. pneumoniae*, or members of the *Chlamydia* genus (*C. trachomatis, C. suis* or *C. muridarum*).

ϕCPAR39, which was isolated from *C. pneumoniae* AR39 can infect *C. pecorum* and *C. abortus* but is unable to infect *C. felis*, *C. psittaci* or any members of the *Chlamydia* genus.

**Table 3.16** The species of chlamydiae that were challenged with Chp2, Chp3 and ϕCPAR39.

| Genus | Species | Strain | Host species | Source or Reference |
|---|---|---|---|---|
| *Chlamydophila* | *C. abortus* | A22 | Ovine | (Stamp *et al.*, 1950) |
| | | S26/3 | Ovine | (McClenaghan *et al.*, 1984) |
| | | S95/3 | Ovine | Jones |
| | | B577 | Ovine | VR-656 |
| | | IPA | Ovine | VR-629 |
| | | T35 | Ovine | (Griffiths *et al.*, 1992) |
| | | BA1 | Bovine | (Griffiths *et al.*, 1995) |
| | | Colo 4 | Bovine | Griffiths |
| | | B111 | Ovine | Griffiths |
| | | BS | Ovine | (Griffiths *et al.*, 1992) |
| | | EAE | Ovine | (Stamp *et al.*, 1950) |
| | *C. psittaci* | 6BC | Avian | VR-125 |
| | | Cal 10 | Avian | (Francis and Magill, 1938) |
| | *C. felis* | FP | Feline | (Cello, 1967) |
| | *C. caviae* | GPIC | Cavine | VR-813 |
| | *C. pecorum* | E58 | Bovine | VR-628 |
| | | BE53 | Bovine | Griffiths |
| | | Stra | Ovine | (Storz, 1963) |
| | | MO901 | Bovine | (Storz, 1963) |
| | | LW508 | Bovine | (Storz, 1963) |
| | | T52 | Ovine | (Griffiths *et al.*, 1992) |
| | *C. pneumoniae* | CWL029 | Human | VR-1310 |
| | | IOL 207 | Human | (Dywer *et al.*, 1972) |
| | | TW 183 | Human | VR-2282 |
| | | N16 | Human | (Wills *et al.*, 1990) |
| *Chlamydia* | *C. trachomatis* | L1/440/LN | Human | (Schachter and Meyer, 1969) |
| | | L2/434/BU | Human | VR-902B |
| | | IOL 1883 | Human | Treharne |
| | | IOL 238 | Human | Treharne |
| | | UW 36 | Human | VR-886 |
| | | IU 888 | Human | (Jones and Van der Pol, 1990) |
| | | TW-5 | Human | (Wang and Grayston, 1963) |
| | | UW-1 | Human | (Alexander *et al.*, 1967) |
| | | NL-D | Human | (Johansson *et al.*, 1997) |
| | | NI 1 | Human | (Tuffrey *et al.*, 1986) |
| | | Jali 5 | Human | Mabey |
| | | Jali 20 | Human | (Hayes *et al.*, 1990) |
| | *C. muridarum* | Mo Pn | Murine | VR-123 |
| | *C. suis* | DC 6 | Porcine | Sachse |

These results show that despite differences in the IN5 and Ins loops in VP1 these bacterophages can all infect *C. abortus*, *C. caviae* and *C. pecorum*. This indicates that these regions are not essential for determining host specificity. Chp2 and Chp3 have identical host ranges (*C. abortus*, *C. caviae*, *C. felis* and *C. pecorum*), however, ϕCPAR39 is unable to infect *C. felis* but can additionally infect *C. pneumoniae*. This suggests that Chp2/Chp3 and ϕCPAR39 have different host receptors, and that both receptors are only present on *C. abortus*, *C. caviae* and *C. pecorum*.

The inability of a chlamydiaphage to infect a host is due to either: the absence of a receptor on the hosts surface, the presence of a receptor but inability to penetrate the host, or the ability to attach and penetrate the host but a failure to replicate and produce active virions. To determine if the failure of Chp2 to infect a host was due to the absence of a receptor on the host surface a Chp2 binding assay to a susceptible host (*C. abortus*) and a non-susceptible host (*C. trachomatis*) was carried out via ELISA using monoclonal antibody 55 (Everson *et al.*, 2002). No significant binding between Chp2 and *C. trachomatis* could be observed. This indicates that an outer membrane receptor is require for attachment which is not present in chlamydiae resistant to chlamydiaphage infection (Everson *et al.*, 2002). Hence the inability of chlamydiaphages to infect a host is due to the absence or altered structure of the receptor.

## 3.7 ϕMH2K

During the sequencing of Chp3 a new bacteriophage ϕMH2K was described (Brentlinger *et al.*, 2002). ϕMH2K is a bacteriophage that infects *Bdellovibrio bacteriovorus*, a proteobacterium. ϕMH2K has a single-stranded DNA genome (4594bp long) similar in size and organisation to the chlamydiaphage genome except for the location of gene 5 (figure 3.17). The genome of ϕMH2K consists of eleven ORFs. ORFs 1,2,3,4,5 and 8 all have ribosome binding sites suggesting that they encode proteins. Other ORFs present in the bacteriophage genome are situated in overlapping reading frames (N, X, Y and Z).

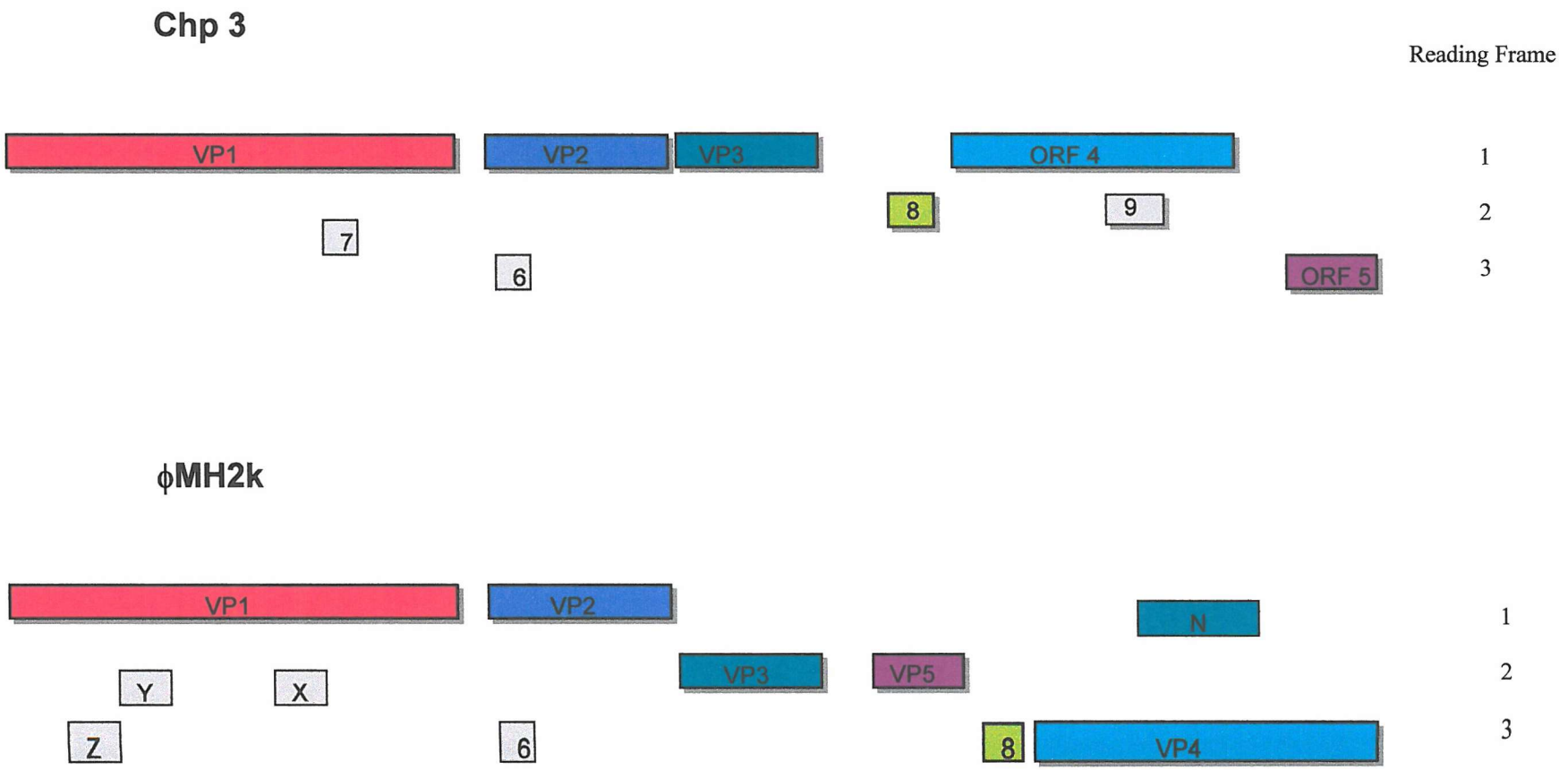



**Figure 3.17** A linear representation of the Chp3 and ϕMH2K genomes.

Table 3.17 shows amino acid sequence identity (%) between the putative ϕMH2K major structural proteins (VP1, VP2 and VP3) and the major structural proteins in the chlamydiaphages SpV4 and ϕX174. Interestingly, the ϕMH2K proteins show the highest percentage amino acid sequence identities with the chlamydiaphages than with SpV4 and ϕX174 (table 3.17). Even more surprising is the fact that the ϕMH2K proteins show similar percentage amino acid sequence identities with the Chp2-like bacteriophage proteins, as the Chp1 proteins do with the Chp2-like bacteriophage proteins. This is surprising due to the dissimilarity of their hosts and suggests that Chp1 like ϕMH2K forms a distinct group within the *Microviridae* family. Like the chlamydiaphages and SpV4, the genome of ϕMH2K is approximately 20% smaller than ϕX174. This is due to the absence of a major spike protein and external scaffolding proteins.

**Table 3.17 Amino acid sequence identity (%) between the ϕMH2K proteins and other members of the *Microviridae* family of bacteriophages.**

| | ϕMH2K vs Chp3 | ϕMH2K vs Chp1 | ϕMH2K vs ϕCPAR39 | ϕMH2K vs ϕCPG1 | ϕMH2K vs SpV4 | ϕMH2K vs ϕX174 |
|---|---|---|---|---|---|---|
| VP1 | 49.8 | 44 | 49.2 | 49.2 | 38 | 19 (F) |
| VP2 | 25.7 | 25.5 | 25.7 | 26 | 25 | 20 (H) |
| VP3 | 31.1 | 31.1 | 31.1 | 31.1 | 18.4 | 18 (B) |

## 3.8 Phylogeny of the *Microviridae* Family of Bacteriophages

To determine the relationship between members of the *Microviridae* family of bacteriophages an unrooted phylogenetic tree was constructed for the major capsid protein of the ten members that have been sequenced (figure 3.18). The *Microviridae* family of bacteriophages seems to fall in to two subfamilies (Brentlinger *et al.*, 2002). The bacteriophages that infect *Enterobacteriaceae* (this group includes ϕX174 and G4) and the bacteriophages that infect obligate intracellular bacteria, which include the chlamydiaphages and ϕMH2K. The main difference between these two subfamilies of *Microviridae* is that

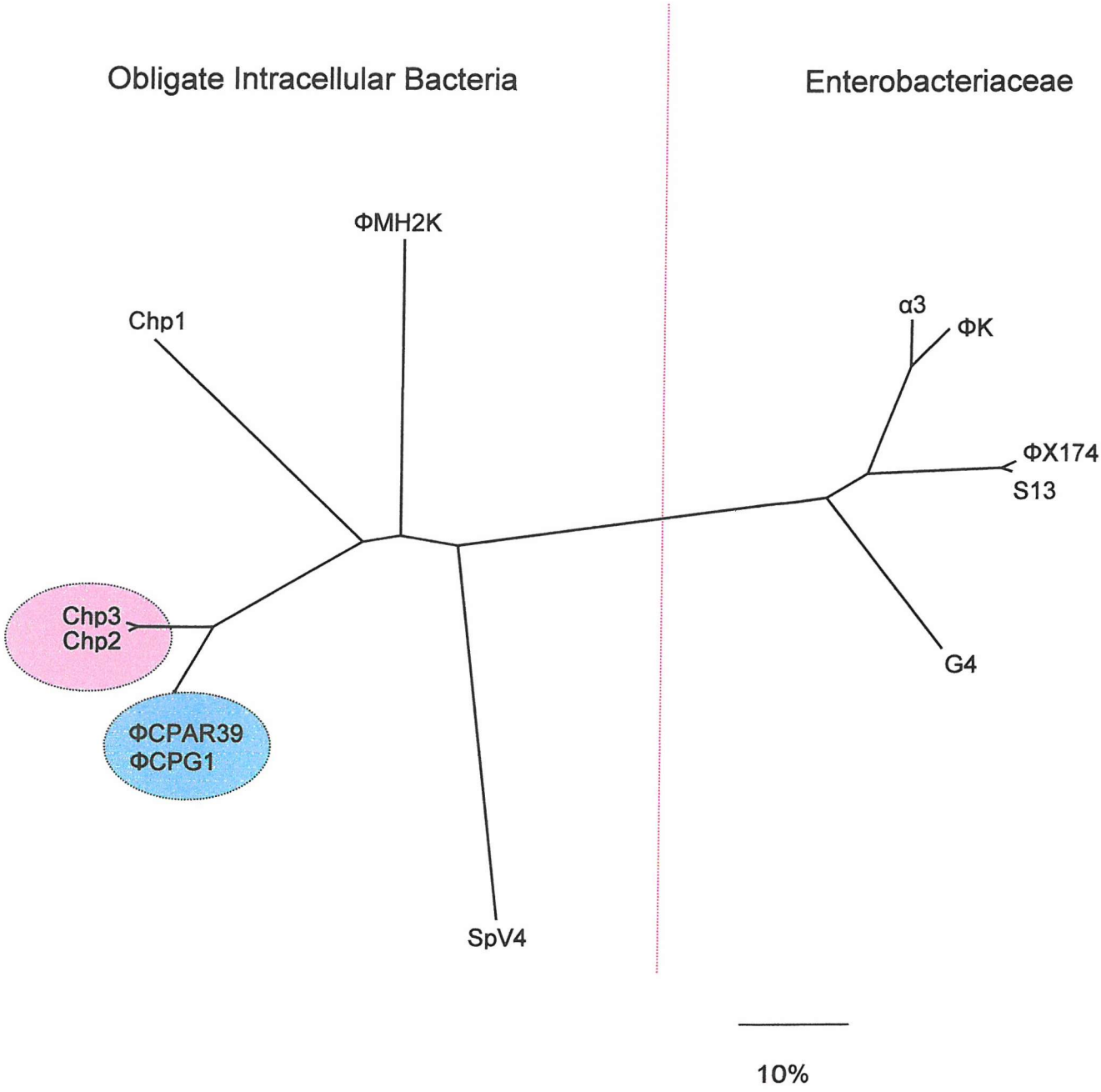



**Figure 3.18** An unrooted phylogenetic tree comparing the nucleotide sequences of the *Microviridae* VP1 structural proteins. This illustrate the relationship between the chlamydiaphages, ϕMH2k and ϕX174. Two distinct subfamilies can be seen within the *Microviridae* family, phages that infect intracellular bacteria and those that infect enterobacteria. Accession numbers (in parentheses) are as follows: Chp2 (AJ270057), ϕAR39 (AE002163), ϕCPG1 (U41758), Chp1 (D00624), MH2K (AF306496), SpV4 (M17988), α3 (X60322), ϕK (X60323), ϕX174 (J02482), S13 (M14428), G4 (J02454). The scale bar indicates percentage divergence.

the chlamydiaphages subfamily does not contain a major spike protein or external scaffolding proteins. The external scaffolding proteins function in ϕX174 by directing the placement of the major spike protein and stabilizing the procapsids (Dokland *et al.*, 1997; Dokland *et al.*, 1999). Members of the chlamydiaphage subfamily of *Microviridae* do not contain a spike protein and it is hypothesised that VP3 maybe the chlamydiaphage equivalent of the internal scaffolding protein (Liu *et al.*, 2000).

## 3.9 Discussion and Future Work

A new chlamydiaphage was described during this study (Chp3). Chp3 was isolated from a *C. pecorum* strain of chlamydiae and its discovery confirms the hypothesis that chlamydiaphages are widespread and that they seem to be confined to the *Chlamydophila* genus of *Chlamydiaceae*.

The Chp3 genome shares the highest nucleotide sequence identity (97.1%) with the Chp2 genome. There are only one hundred and forty-three nucleotide differences between the Chp2 and Chp3 genomes. Each of the Chp2/Chp3 ORFs are situated in the same genomic location and reading frame with the exception of ORF8. The Chp3 ORF8 is situated in a different reading frame than the Chp2 ORF8 due to an eleven-nucleotide deletion situated four nucleotides downstream of the Chp3 ORF3 terminator codon. A two-nucleotide insertion situated downstream of the Chp3 ORF8 puts the Chp3 ORF4 back in the same reading frame as the Chp2 ORF4. No significant secondary structures are predicted to be located within the eleven-nucleotide region of Chp2 that is deleted in Chp3.

The main differences between the genomes of Chp3, Chp2, ϕCPG1 and ϕCPAR39 are found in two discrete regions of the VP1 gene (the IN5 loop and Ins) (Read *et al.*, 2000b). It has been suggested that the IN5 loop is responsible for receptor recognition (Chipman *et al.*, 1998). The IN5 loop forms globular protusions on the surface of spiroplasma *Microviridae* (Chipman *et al.*, 1998). These protrusions are absent on the surface of coliphage *Microviridae* and may be a subsitute for the spike proteins (Fane *et*

*al.*, 2002). The IN5 loop of Chp2 and Chp3 is very similar and the IN5 region of ϕCPG1 and ϕCPAR39 is very similar. It is likely that this region is the chlamydiaphage receptor both Chp2 and Chp3 should have overlapping host ranges and similarly ϕCPG1 and ϕCPAR39 should have similar host ranges.

Chp2 and Chp3 have identical host ranges (*C. abortus*, *C. caviae*, *C. felis* and *C. pecorum*). ϕCPAR39 can also infect *C. abortus*, *C. caviae* and *C. pecorum* despite differences in the IN5 and Ins loops. Although, it is unable to infect *C. felis* and can additionally infect *C. pneumoniae*. This suggests that Chp2/Chp3 and ϕCPAR39 have different host receptors, and that both receptors are only present on *C. abortus*, *C. caviae* and *C. pecorum*. Based on these observations it is hypothesised that ϕCPG1 will have an identical host range to ϕCPAR39.

Future work could involve the identification of the host receptors for Chp2/Chp3 and ϕCPAR39/ϕCPG1 and the detection of other novel chlamydiaphages in different chlamydial hosts. The recent discovery of a bacteriophage (ϕMH2K) that is closely related to the chlamydiaphages in an unrelated bacterial host has shown that related bacteriophages are widespread. The isolation of ϕMH2K is a major discovery as mutants can be easily isolated in host-independent mutants of *Bdellovibrio*. This will allow prompt biochemical and genetic analysis of these bacteriophages which, can directly be applied to the chlamydiaphages (Brentlinger *et al.*, 2002).

As Chp2 has a higher infection rate in chlamydial cells than Chp3 it was subsequently used in all future studies investigating the molecular biology of chlamydiaphages.

# Chapter 4

# Studies on Chp2 Protein Expression

## 4.1 Introduction

Chlamydiaphage genomes encode eight major open reading frames (ORFs). It has been shown that the proteins encoded by ORF1, ORF2 and ORF3 are the structural proteins VP1, VP2 and VP3 respectively (Storey *et al.*, 1989). The functions of the proteins encoded by ORF4-8 are unknown, neither is it known which of these ORFs are expressed during the Chp2 replication cycle. A necessary step to study the biology of Chp2 is to identify which of the ORFs are expressed during bacteriophage infection. Initially to predict which of the ORFs encode proteins comparisons were made between the amino acid sequences of the proteins encoded by the Chp2 and ϕX174 genomes. The ϕX174 genome encodes 11 proteins (A, A*, B, C, D, E, F, G, H, J and K). The table below describes the functions of each of the proteins and shows the Chp2 proteins that are predicted to perform a similar role based on amino acid similarities and/or regions of homology.

**Table 4.1 Comparisons between the ϕX174 and Chp2 proteins**

| ϕX174 Proteins | Protein functions | Chp2 proteins predicted to perform a similar function |
|---|---|---|
| A | DNA replication<br>Stg II, StgIII | ORF4 protein |
| A* | Shut of host cell replication | / |
| B | Internal Scaffolding protein | VP3 |
| C | Single stranded DNA synthesis | ORF5 protein |
| D | Morphogenesis | / |
| E | Lysis protein | / |
| F | Major Capsid protein | VP1 |
| G | Major spike protein | / |
| H | DNA pilot protein | VP2 |
| J | Bacteriophage internal protein/<br>DNA Binding protein | ORF8 protein |
| K | Bacteriophage synthesis | / |

The ϕX174 proteins that are expressed during infection were identified by comparing the electrophoretic patterns of the proteins produced from both bacteriophage infected and uninfected *E. coli* (Mayol and Sinsheimer, 1970). Due to the large number of host proteins that are synthesised by *E. coli* during replication it was initially difficult to identify all of the bacteriophage specific proteins. Methods were subsequently devised to reduce the rate of host protein synthesis. One investigation, which utilised ultraviolet irradiation to decrease the number of host proteins that were synthesised during replication, allowed the identification of six bacteriophage specific proteins (Burgess and Denhardt, 1969). The genes encoding each of these proteins were subsequently mapped using ϕX174 mutants, based on the observation that infection with bacteriophages that contained mutations in genes encoding protein eliminates the expression of that protein (Burgess and Denhardt, 1969; Baker and Tessman, 1969; Benbow *et al.*, 1974). It was found that four of the six proteins were components of the bacteriophage coat (F, G, H and J). The other proteins were shown in *E. coli* to be responsible for lysis (protein E) as mutants were unable to cause efficient lysis of *E. coli* and a scaffolding protein (protein D) as D proteins are associated with procapsids but absent from mature phages (Burgess and Denhardt, 1969; Fujisawa and Hayashi, 1977).

Further studies allowed the identification of the proteins encoded by the genes A, A*, B, D and K. The gene A and gene A* products were identified using ssDNA cellulose chromatography due to their high affinity for ssDNA (Linney *et al.*, 1972). The function of protein A in RF replication was demonstrated by Lindqvist and Sinsheimer (1967) by studying infections with A mutants. No RF DNA synthesis was observed after infection by A mutants (Lindqvist and Sinsheimer, 1967). The gene B product was identified by comparing the total protein produced from a temperature sensitive gene B mutant grown under permissive and non-permissive conditions (Siden and Hayashi, 1974). The function of the B protein as a scaffolding protein was initially hypothesised during these investigations by examining the bacteriophage particles and intermediate particles that were produced during infection with B mutants. It was found that complexes between F protein aggregates and G protein aggregates were not formed indicating that the B protein stabilises the bonds between these complexes. The gene C product was identified by analysing the proteins

produced from UV-irradiated bacteriophage infected *E. coli* (Mayol *et al.,* 1970). It was shown that in C mutant infected cells, viral ssDNA synthesis is greatly reduced indicating that protein C is involved in ssDNA synthesis (Mayol and Sinsheimer, 1970). The gene encoding protein K was the last gene for which a function was discovered. It was identified by comparing the G4 bacteriophage gene K with regions of the ϕX174 genome (Shaw *et al.*, 1978). The expression of protein K was subsequently identified by locating a novel protein produced only by bacteriophage infected *E. coli* that was a similar size and possessed a similar leucine to lysine incorporation ratio. Protein K is an unessential protein but could be beneficial to bacteriophage growth as bacteriophages that contained mutations in gene K yielded a two-threefold decline in burst size (Tessman *et al.*, 1980).

Unlike ϕX174, Chp2 proteins cannot be purified *in vivo* in large quantities and Chp2 mutants have not been isolated. Therefore to determine which of the Chp2 ORFs are expressed during infection it would first be necessary to produce specific reagents to each protein so that they can be detected *in vivo*. The development of an assay to detect the expression of each of the proteins during infection will also enable studies to identify the function of each protein.

The aim of this investigation was to obtain mono-specific reagents to the proteins encoded by each of the major ORFs so that polyclonal antisera to these regions could be raised. This will allow studies into the expression of the individual gene products both *in vivo* and *in vitro.*

## 4.2 Construction of constructs containing ORF2-8

The Xpress™ (Invitrogen) system using the pRSET vectors was chosen to express ORF2-5 because the proteins encoded by ORF2-5 are predicted to be large and could easily be visualised by SDS-PAGE. Also in earlier projects VP1 from both Chp1 and Chp2 have been expressed using this system. ORF6, ORF7 and ORF8 would encode small proteins, which would make them difficult to purify using the pRSETA system. Instead it was planned for them to be expressed using the pMAL protein fusion and purification system (New England Biolabs) or the glutathione s-transferase (GST)

protein fusion and purification system. Both of these methods allow fusions with larger carrier proteins. This increases the size of the expressed product making it easier to detect by SDS PAGE and the carrier part of the molecule has particular binding properties to assist purification.

### 4.2.1 Construction of pRSETA clones containing ORF2-5

The pRSETA system is easy to use and allows a large amount of recombinant protein to be produced in *E. coli* cells then purified by column purification (2.9.1). DNA sequences that were ligated into the pRSETA vector are placed under the control of a T7 promoter and positioned in-frame with the sequence encoding an N-terminal fusion peptide. This includes a six histidine residue polypeptide which functions as a metal binding domain allowing the purification of the recombinant proteins. A diagrammatic representation of the pRSETA vector is shown in figure 4.1.

#### 4.2.1.1 Primer design

Primers to amplify each of the ORFs (2-5) were first designed (table 4.2). The primers were designed to incorporate a *Bam*H I (ORF5 -*Sac* I) site at the 5' termini of the PCR product and a stop codon and a *Hin*d III (ORF5 -*Kpn* I) site at the 3' termini of the product. This allowed the fragment to be ligated into the corresponding *Bam*H I/*Sac* I and *Hin*d III/*Kpn* I sites in the pRSETA vector, in-frame with the translation initiation codon. The primer sequences and the predicted sizes of the products they amplify are shown in table 4.2.

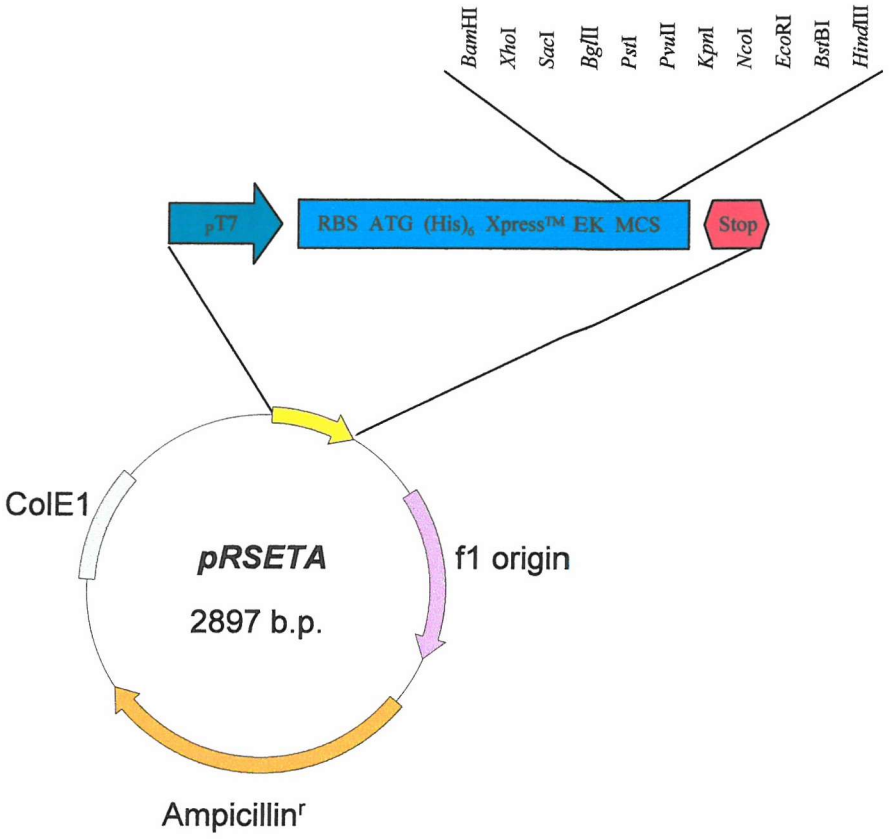



**Figure 4.1** A diagrammatic representation of the pRSETA vector used for the expression of ORF2-5. The PCR products of ORF2-4 were cloned into the *Bam*H I/*Hin*d III sites and ORF5 was cloned into the *Sac* I/*Kpn* I sites within the multiple cloning site in-frame with the T7 promoter.

**Table 4.2** **Primers used to amplify fragments of DNA for cloning into pRSETA vectors. The restriction sites, and 3' co-ordinates are shown in bold**

| Region Amplified | Primer name | Primer sequence 5'.....3' | Product size | Predict protein |
|---|---|---|---|---|
| ORF2 | VP2F<br>VP2R | CGTAAG**GGATCC**ATGAATCCCGAACAACTTACG[1824]<br>GCAGTA**AAGCTT**CTACCTTCCTTTTCTTGA[2347] | 558bp | 20.2kDa |
| ORF2 Region A | VP2RA | ATCACG**AAGCTT**CTAAGGATTATTCGGAGACCA[2083] | 483bp | 17.5kDa |
| ORF2 Region B | VP2FB | GCATAG**GGATCC**TCTTCTCCTGCAGGAGCG[2079] | 516bp | 18.7kDa |
| ORF3 | VP3F<br>VP3R | GCATAC**GGATCC**ATGTTTAAGTCGGCATATTCC[2385]<br>GAGCGT**AAGCTT**TCATTTTGGGCTAATCCAGG[2791] | 444bp | 16.8kDa |
| ORF4 | ORF4F<br>ORF4R | GCATAC**GGATCC**ATGTTGATTTTTAATGTCTTA[3282]<br>GAGCGT**AAGCTT**TTAGTAACTCTCCTCCTCATA[4253] | 1008bp | 40.2kDa |
| ORF4 Region B | ORF4FB | GCATAC**GGATCC**ATGTACGATTTTCCTGATAAGAAG[3780] | 552bp | 20kDa |
| ORF5 | ORF5F<br>ORF5R | GCATAC**GAGCTC**ATGAAAGTTTTTACAGTG[4319]<br>GAGCGT**GGTACC**CTACTGTTTGTGCTTAAAATC[4536] | 252bp | 9.5kDa |

F-Forward primer, R- Reverse primer,

### 4.2.1.2 Production of PCR products for cloning into pRSETA

To produce fragments of DNA for cloning into pRSETA, PCR reactions were carried out to amplify each of the ORFs as described in section 2.7 using Bio-x-Act DNA polymerase to minimise random mutations. Bio-x-Act consists of a mixture of polymerases that have 5'-3' DNA polymerase activity and 3'-5' proof-reading activity which prevents misincorporations of nucleotides during primer extension. Table 4.3 shows the reaction conditions for each primer pair. Following PCR, the products were analysed by gel electrophoresis to verify the size of the amplified fragments, which were then purified to remove any remaining template, nucleotides or primers by absorption to a silica membrane (2.7.3.1).

**Table 4.3 PCR conditions**

| *Region of genome* | *Denaturation* | *Annealing* | *Extension* |
|---|---|---|---|
| ORF2 | 94°C for 20 seconds | 50°C for 20 seconds | 72°C for 34 seconds |
| ORF2 Region A | 94°C for 20 seconds | 64 °C for 20 seconds | 72°C for 19 seconds |
| ORF2 Region B | 94°C for 20 seconds | 60 °C for 20 seconds | 72°C for 21 seconds |
| ORF3 | 94°C for 20 seconds | 61 °C for 20 seconds | 72°C for 27 seconds |
| ORF4 | 94°C for 20 seconds | 60 °C for 20 seconds | 72°C for 60 seconds |
| ORF4 region B | 94°C for 20 seconds | 60 °C for 20 seconds | 72 °C for 33 seconds |
| ORF5 | 94°C for 20 seconds | 50 °C for 20 seconds | 72°C for 30 seconds |

### 4.2.1.3 Cloning the PCR fragments into pRSETA

The DNA amplicons were digested with restriction enzymes (section 2.6.1.1) corresponding to the two sites incorporated into the 5' and 3' termini during PCR. Following restriction digestion the fragments of DNA were purified by absorption to a silica membrane (2.7.3.1) and ligated into a pRSETA vector (2.6.1.3) that had been digested with the corresponding restriction enzymes. The ligations were then transformed into competent *E. coli* JM101 cells (2.3.3). *E. coli* JM101 was used as the host strain in all transformations as it carries the F' pilus needed for M13/T7 bacteriophage infection.

A restriction digest of purified plasmid DNA was used to verify the correct construct in each case. The cloning junctions and inserts from clones that contained inserts of the correct size were then sequenced using both forward and reverse primers to confirm the inserted sequences were correct (2.8).

#### 4.2.1.4 *E. coli* T7 S30 system for circular DNA

To verify that the pRSETA clones containing ORF1-5 had been constructed correctly the recombinant proteins were expressed in an *in vitro* transcription/translation system. The *E. coli* S30 extract system (Promega) was used as it allows background proteins that occur in *in vivo* systems to be greatly reduced (2.14.2). It also allows the expressed proteins to have greater stability from protease degradation as the *E. coli* B strains used to prepare the S30 system are deficient in *lon* protease activity. This allows recombinant proteins that are expressed at low levels *in vivo* due to host-encoded repressor action to be expressed at higher levels.

The plasmid DNA from each construct was prepared by mini prep kit (Qiagen). Then all five recombinant plasmids (pRSETA containing ORF1-5) were used as template in *in vitro* transcription/translation reactions (2.14.2). The products from the reactions were analysed by SDS-PAGE followed by autoradiography (2.12.1,2.16). All five clones produced proteins of the correct size in this system (figure 4.2), indicating that they have been constructed correctly.

In previous work monoclonal antibodies (Mabs) have been generated to Chp2, however the specificity of these MAbs was unknown (Mabs 55, 60, 18, 106, 40). None of them showed any reactivity in an immunoblot with semi-purified Chp2 particles. Each of the MAbs was used in a RIPA with the five transcription/translation products (2.15). The results from this analysis showed that 55 and 18 reacted to the main capsid/coat protein VP1 (figure 4.3). No reactivity could be seen between any of the other Mabs and the translation products of ORF2-5. The Mabs (55 and 18) that reacted strongly and specifically to VP1 could be used in future investigations to study the expression of this protein. As no Mabs were identified that react strongly to any of the other proteins it was planned to raise polyclonal antisera to these antigens.

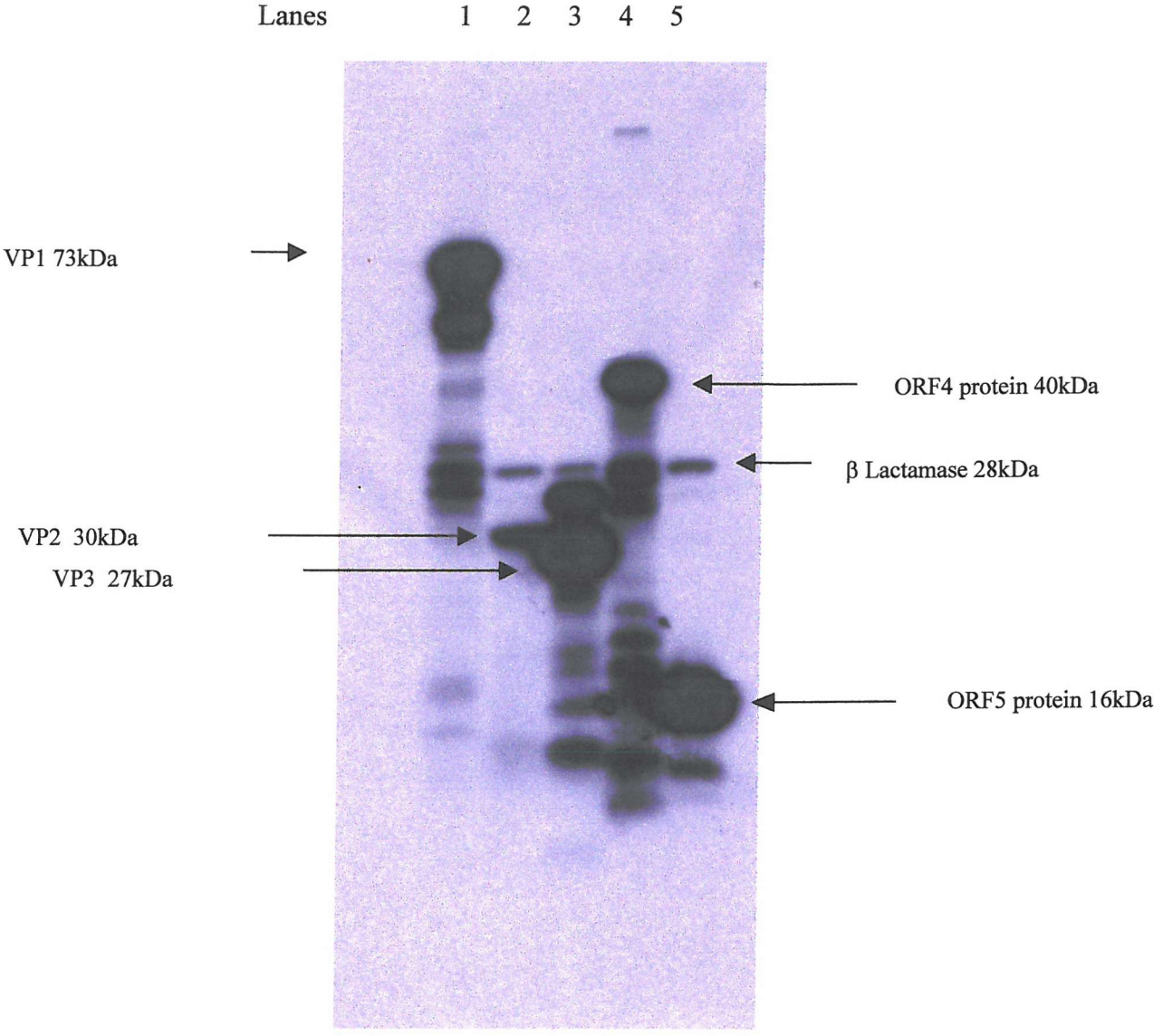



**Figure 4.2** Autoradiograph from a 12.5% SDS PAGE gel showing the proteins produced from an S30 *in vitro* transcription/translation reaction, for pRSETA containing ORF1-5. Lane 1. PRSETAORF1; Lane 2. pRSETAORF2; Lane 3. pRSETAORF3; Lane 4. pRSETAORF4; Lane 5. pRSETAORF5.

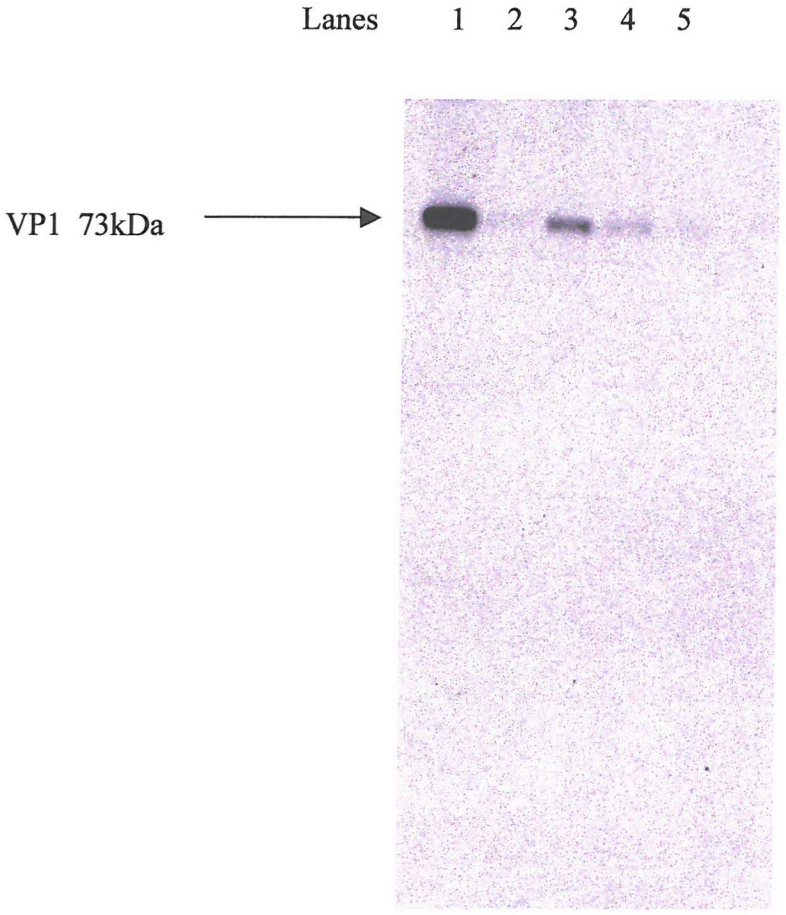



**Figure 4.3** Autoradiograph showing the proteins that immunoprecipitated with monoclonal antibodies raised to Chp2 particles from an $^{35}S$ labelled *in vitro* transcription and translation reaction of pRSETA containing ORF1. Lane 1. Monoclonal antibody 55; Lane 2. Monoclonal antibody 60; Lane 3. Monoclonal antibody 18; Lane 4. Monoclonal antibody 106; Lane 5. Monoclonal antibody 40.

#### 4.2.1.5 Expression of Chp2 recombinant proteins *in vivo*

The expression of each protein from the pRSETA constructs in *E. coli* containing ORF2-ORF5 was induced by the addition of M13/T7 bacteriophage and IPTG (2.9.1.4). The results were then analysed by SDS PAGE (2.12.1). Proteins were expressed from the pRSETA clones containing ORF3 and ORF5 corresponding to VP3 and ORF5 protein. However, the VP2 and ORF4 protein could not be expressed from the pRSETA clones containing ORF2 and ORF4, presumably due to VP2 and the ORF4 protein being toxic to the *E. coli* cells.

#### 4.2.1.6 Expression of clones in *E. coli* BL21 containing pLysS

As VP2 and ORF4 protein could not be expressed in *E. coli* JM101 cells, the pRSETA constructs that contained ORF2 and ORF4 were transformed into an *E. coli* strain that contained the plasmid pLysS. The presence of the plasmid pLysS gives target genes extra stability due to the presence of a T7 lysozyme that acts as an inhibitor of T7 RNA polymerase. This system allows genes to be expressed at a lower level and therefore increases the tolerance of the host bacteria to plasmids containing toxic inserts.

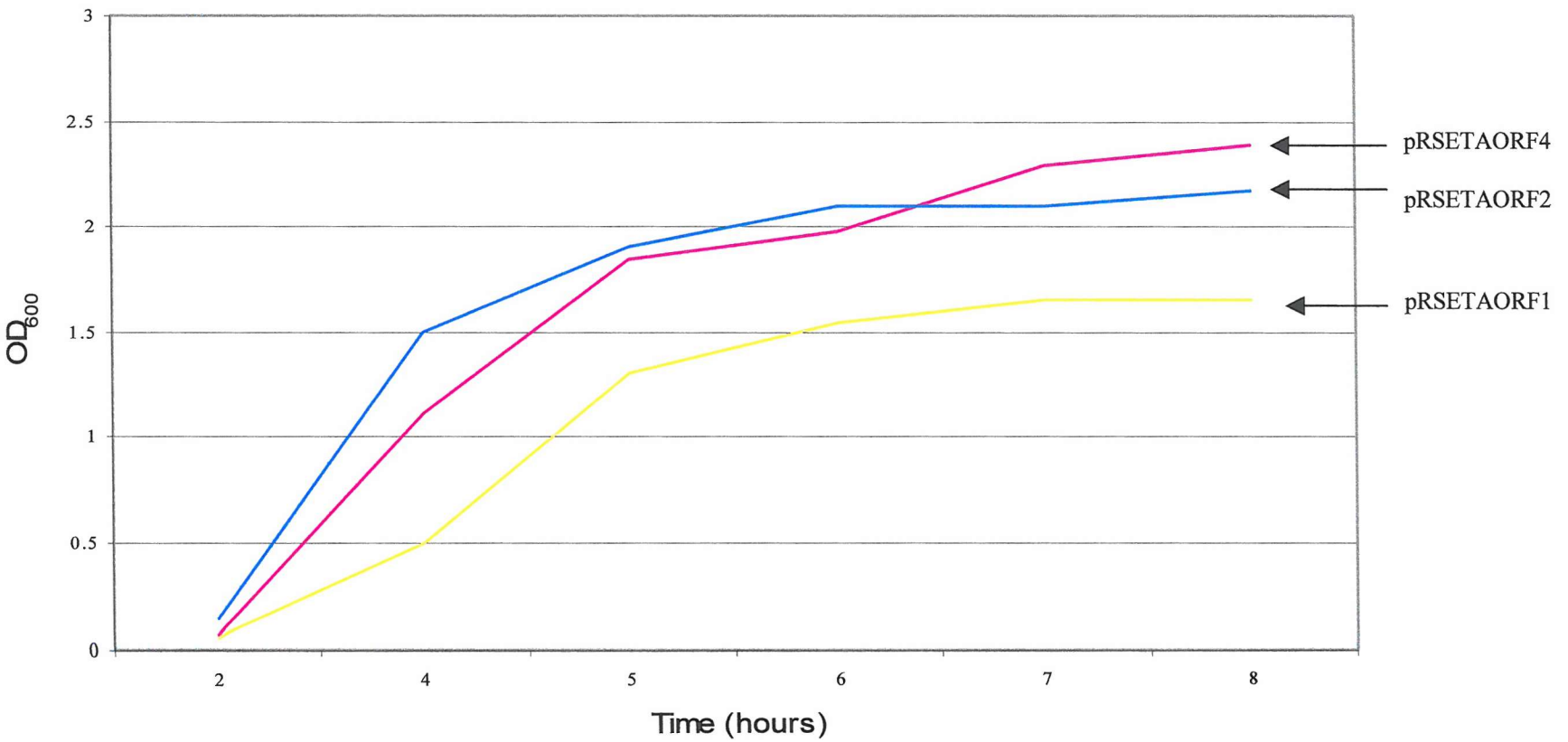
Plasmid DNA was prepared for the pRSETA constructs containing ORF2 and ORF4 using a mini prep kit (2.4.2.2) (Qiagen), and the plasmids were transformed into *E. coli* BL21 that contained the plasmid pLysS (2.3.3). The expression of each protein was then induced as described in section 2.9.1.5. The two proteins could not be expressed using this system indicating that they are still toxic to *E. coli* cells when expression was induced at a lower level.

#### 4.2.1.7 Effect of induction of VP2 and ORF4 protein expression on *E. coli* cells

From the previous investigation it was shown that VP2 and ORF4 protein could not be expressed *in vivo* in either *E. coli* JM101 or *E. coli* BL21 (pLysS) from pRSETA constructs containing ORF2 or ORF4. However, the expression of both these proteins

can be seen when the same constructs are used as template in an *in vitro* transcription/translation system indicating that the clones have been constructed correctly. The reason why these proteins could not be expressed *in vivo* may be due to the proteins been lethal to the host bacteria and causing either an inhibition of replication or lysis. To determine the effect of the induction of expression of these proteins on the host *E. coli*, the number of bacterial cells present before and after IPTG induction was measured and compared between three cultures of *E. coli* JM101. One culture that contained the plasmid pRSETAORF2, one that contained the plasmid pRSETAORF4 and one culture containing pRSETAORF1 that was included in the investigation to act as a positive control since VP1 has been expressed from this construct in a previous experiment. $OD_{600}$ readings were taken from each culture every hour after IPTG induction and compared. The results showed no significant inhibition or decrease in the number of bacterial cells present after IPTG induction from any of the cultures (figure 4.4).

**Figure 4.4** **A graph to show the OD600 of *E. coli* cells after induction of VP1, VP2 and ORF4 protein expression.**



### 4.2.2 Construction of pRSET clones containing parts of ORF2 and ORF4

As VP2 and ORF4 protein could not be expressed *in vivo* from pRSETA clones that contained ORF2 or ORF4, clones were constructed that contained parts of these ORFs to eliminate any sequences inhibiting the expression of these proteins.

#### 4.2.2.1 Primer design

Primers were designed to produce PCR products analogous to the 5′ (region A) and 3′ (region B) of ORF2 and to the 3' (region B) of ORF4 (see figure 4.5), which would correspond to the N- and C- terminus of the proteins. The primers were designed to incorporate a *Bam*H I site at the 5′ termini and a stop codon and *Hin*d III site at the 3′ termini of the PCR products. The primers used and product sizes are shown in table 4.4.

#### 4.2.2.2 Production of PCR products for cloning into pRSETA

DNA fragments were amplified for each region using the PCR conditions shown in table 4.3. The presence of a single PCR product was verified by agarose gel electrophoresis (2.5.2), which was then purified by absorption to a silica membrane (2.7.3.1) to remove any remaining primers and nucleotides. The DNA fragments were then digested with the restriction enzymes *Bam*H I and *Hin*d III and purified by absorption to a silica membrane (2.7.3.1).

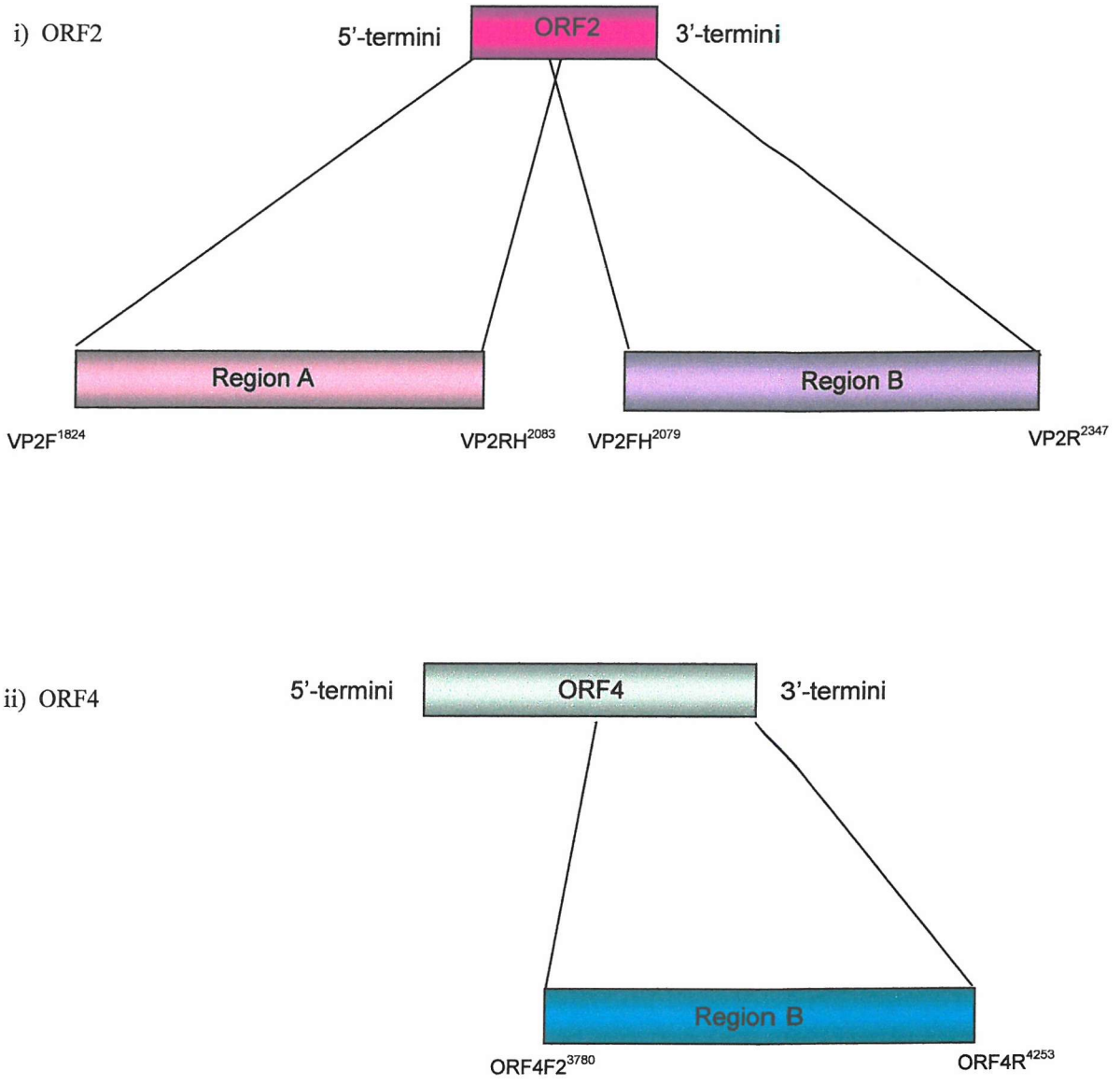



**Figure 4.5** A diagrammatic representation of the regions of ORF2 and ORF4 that were cloned into pRSETA. The PCR products of region A and region B of ORF2 and the region B of ORF4 were ligated into the *Bam*H I/*Hin*d III sites with-in the multiple cloning site of pRSETA. The primers used to amplify each region and the 3' co-ordinates of each primer are indicated.

#### 4.2.2.3 Cloning and expression of PCR fragments

Following purification the fragments of DNA were ligated in-frame with the nucleotide sequence of the N-terminal fusion peptide into *Bam*H I /*Hin*d III digested pRSETA. The ligations were then transformed into competent *E. coli* JM101 cells. To confirm the presence of the correct construct in each case, plasmid DNA was prepared (2.4.3.3) and a restriction digest was carried out to release the insert. The cloning junctions and inserts from constructs containing inserts of the correct size were then sequenced (using both forward and reverse primers), to confirm the inserted sequences were correct (2.8).

The expression of each protein was then induced as described in section 2.9.1.4. The VP2 fragments could not be expressed from the pRSETA constructs containing parts of ORF2. The ORF4 protein fragment was expressed from the pRSETA construct containing region B of ORF4.

As VP2 could not be expressed in *E. coli* cells a 19-residue peptide was produced (Cys-LTYERKKMQAELQNL[127]) (Sigma-Genosys, Ltd) based on a region predicted to have high antigenicity and high surface probability from a Jameson and Wolf plot (Jameson and Wolf, 1988) and Emini plot (Emini *et al.*, 1985) respectively (figure 4.6). As the peptide is a small molecule that would not be immunogenic alone it was bound to keyhole limpet hemacyanin (KLH) via an N-terminal cysteine residue.

### 4.2.3 Purification of the proteins expressed using the XPRESS™ protein expression system

The proteins expressed using the XPRESS™ protein expression system (VP3, ORF4 region B protein and ORF5 protein) were purified from the *E. coli* proteins by metal affinity chromatography using the six-histidine tag introduced by the N-terminal fusion peptide (2.9.1.7). The eluant from each purification was collected in fractions, then analysed by SDS-PAGE (2.12.1). The protein gel of VP3 is typical for these constructs (figure 4.7) therefore the protein purified from the other constructs is not shown.

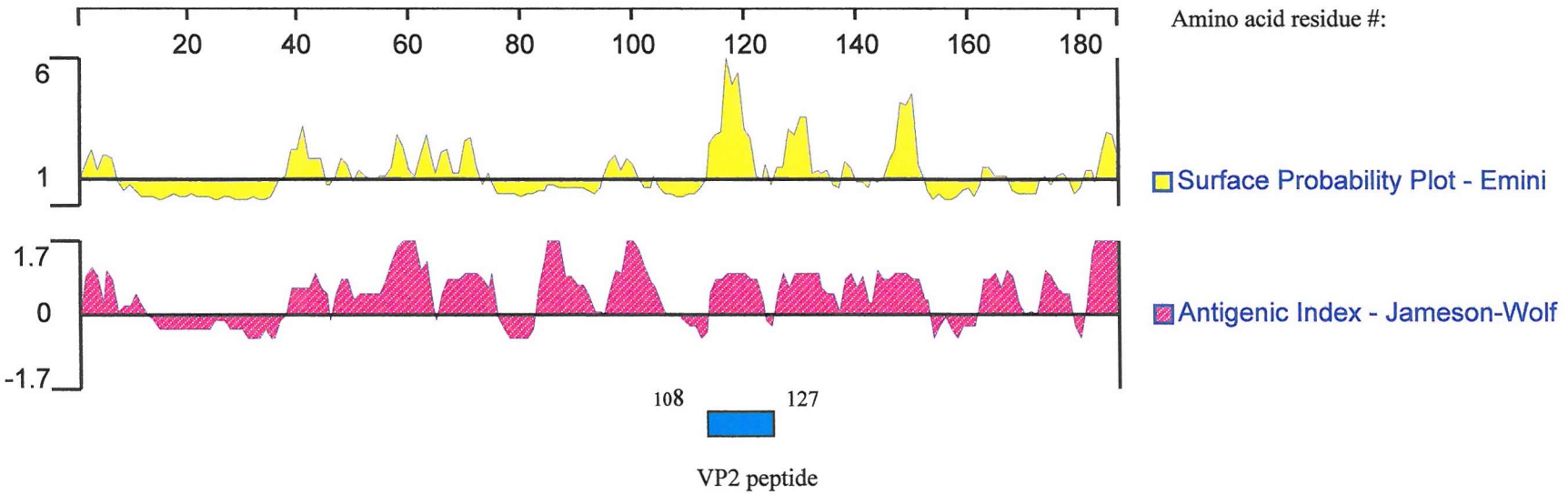



**Figure 4.6** A surface probability plot (Emini *et al* 1985) and antigenicity plot (Jameson and Wolf 1988) of VP2. The region that the VP2 peptide was designed on is indicated. This region is predicted to be surface exposed and have a high antigenicity.

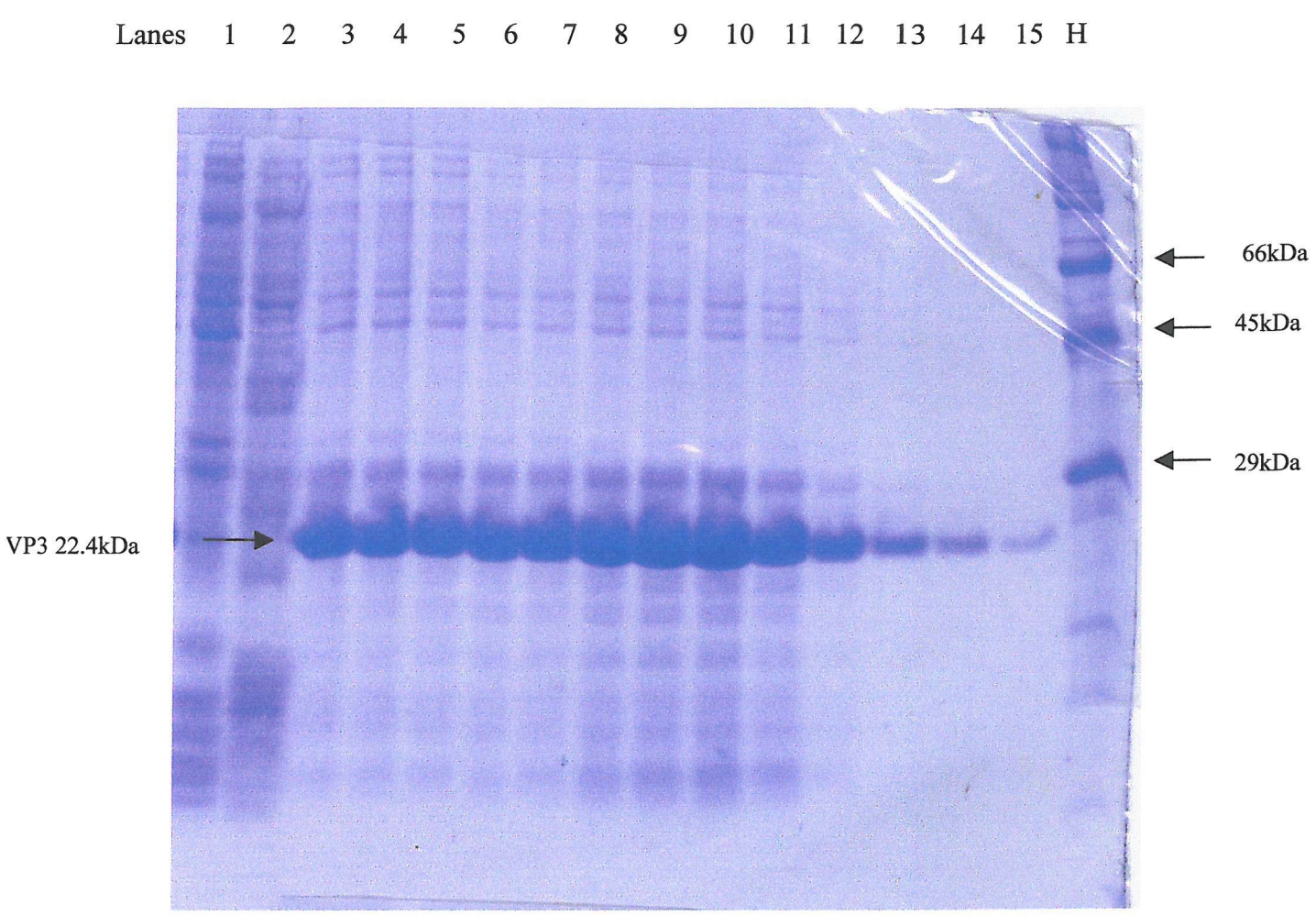



**Figure 4.7** A 12.5% SDS PAGE gel showing the 13 column purification fractions collected from the purified VP3 protein before dialysis. Lane 1. Wash 1; Lane 2. Wash 2; Lane 3-15. 20µl of the 500µl fractions eluted from the resin; Lane H. High Molecular Weight Markers.

#### 4.2.3.1 Dialysis of the purified proteins

During purification, the buffers that were used to wash away the *E. coli* proteins contained urea. Before using the protein as an immunogen it was necessary to remove the urea and to concentrate the protein in a smaller volume. The fractions of eluant that contained the expressed protein were combined in visking tubing and dialysed against PBS (2.9.1.7.3). Figure 4.8 shows the concentrated VP3 and ORF5 protein.

Interestingly, two bands can be observed for the ORF5 protein. Figure 4.9 shows the concentrated ORF4 protein fragment.

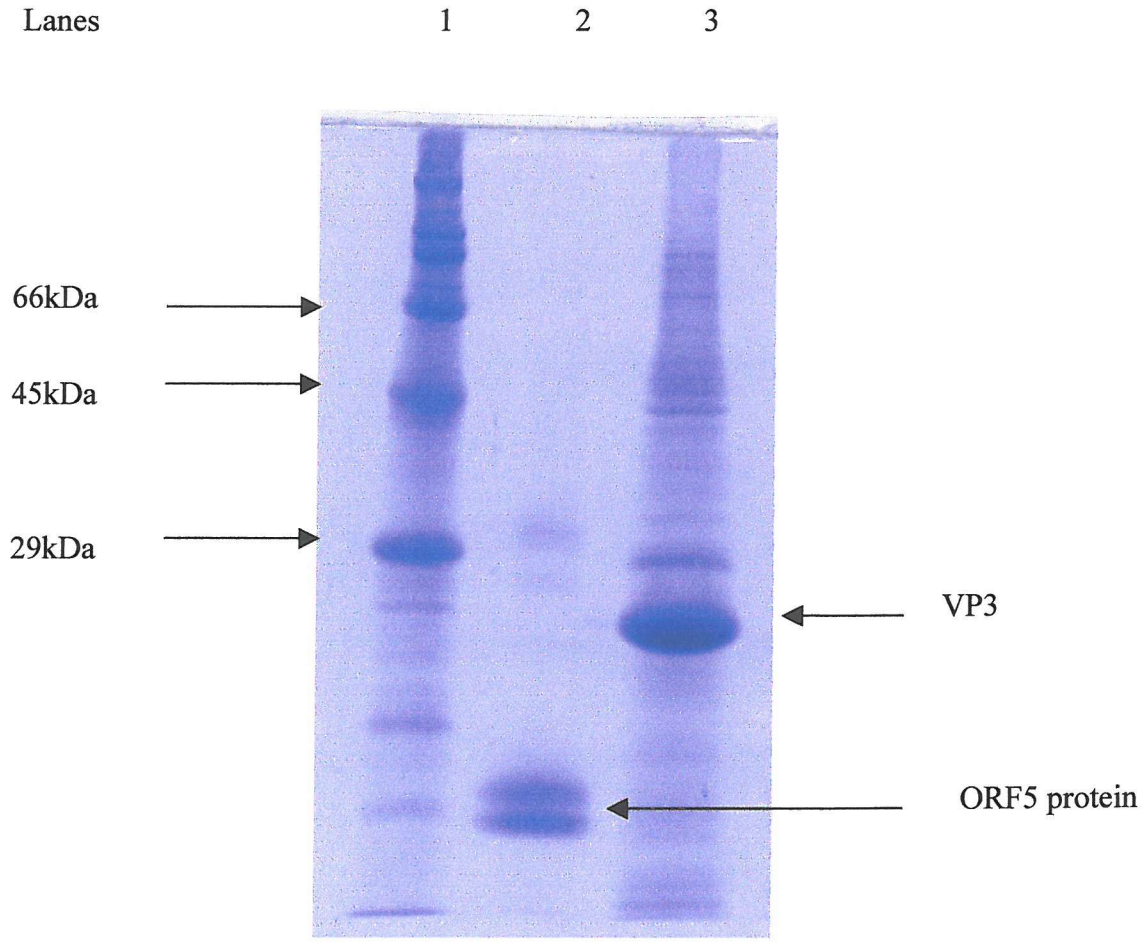



Figure 4.8 A 12.5% SDS-PAGE gel showing the final dialysed proteins used for immunisation. Lane 1. High molecular weight markers; Lane 2. ORF5 protein; Lane 3. VP3.

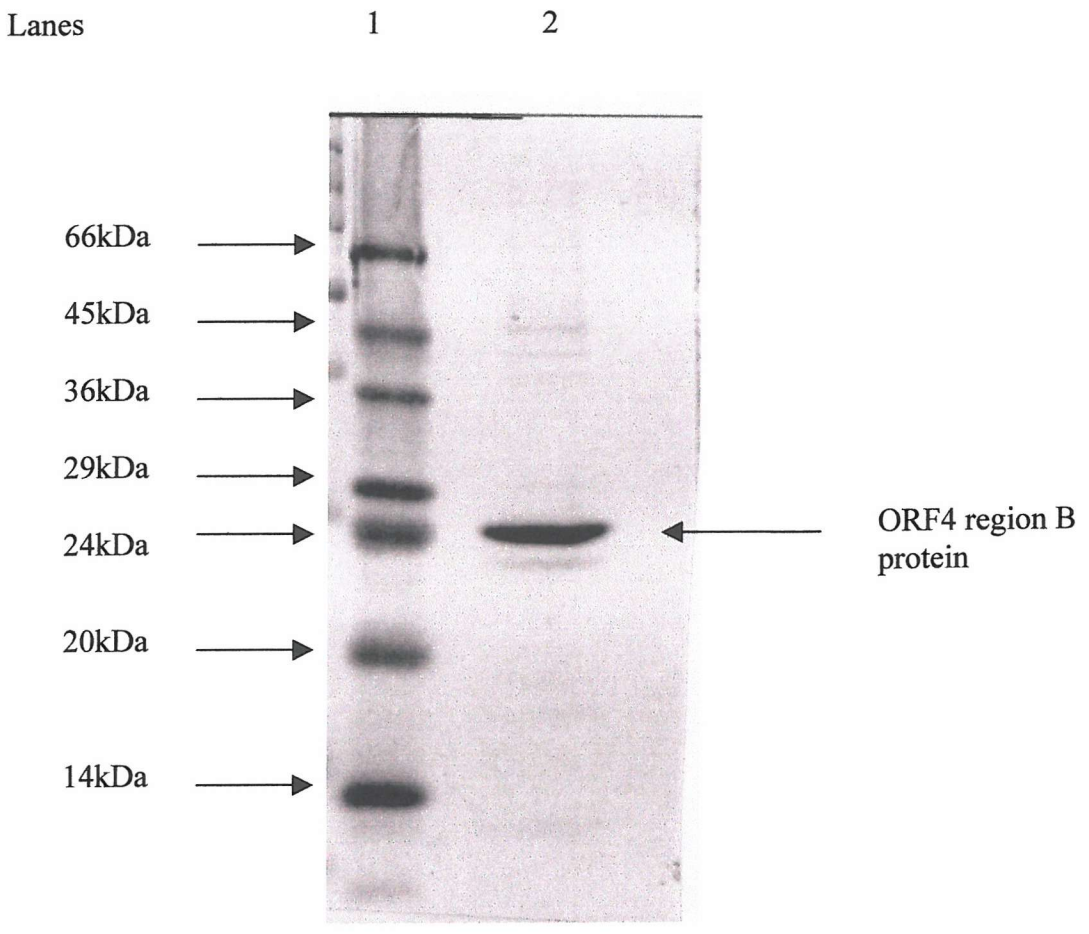



**Figure 4.9** A 12.5% SDS-PAGE gel showing the final dialysed ORF4 region B protein used for immunisation. Lane 1. Low molecular weight marker; Lane 2. ORF4 region B protein.

### 4.2.4 Construction of pMAL-c2 clones containing ORF6/ORF7 and ORF8

The pMAL protein fusion system allows the fusion of proteins to a maltose binding protein (MBP) therefore increases the size of the expressed product (2.9.3). This system uses the *mal*E translation initiation signals plus an inducible tac promoter to allow a high level of expression. The expressed proteins can subsequently be purified using an affinity purification system specific for MBP (2.9.3). A diagrammatic representation of pMAL-c2 is shown in figure 4.10.

#### 4.2.4.1 Primer design

Primers to amplify ORF6-8 were designed to incorporate a *Bam*H I site at the 5' termini of the PCR product and a stop codon and a *Hin*d III site at the 3' termini of the product. This allowed the fragment to be ligated into the corresponding *Bam*H I and *Hin*d III sites in the pMAL-c2 vector, in-frame with the 3' end of *malE*. The primer sequences and the predicted size of the products they amplify can be seen in table 4.4.

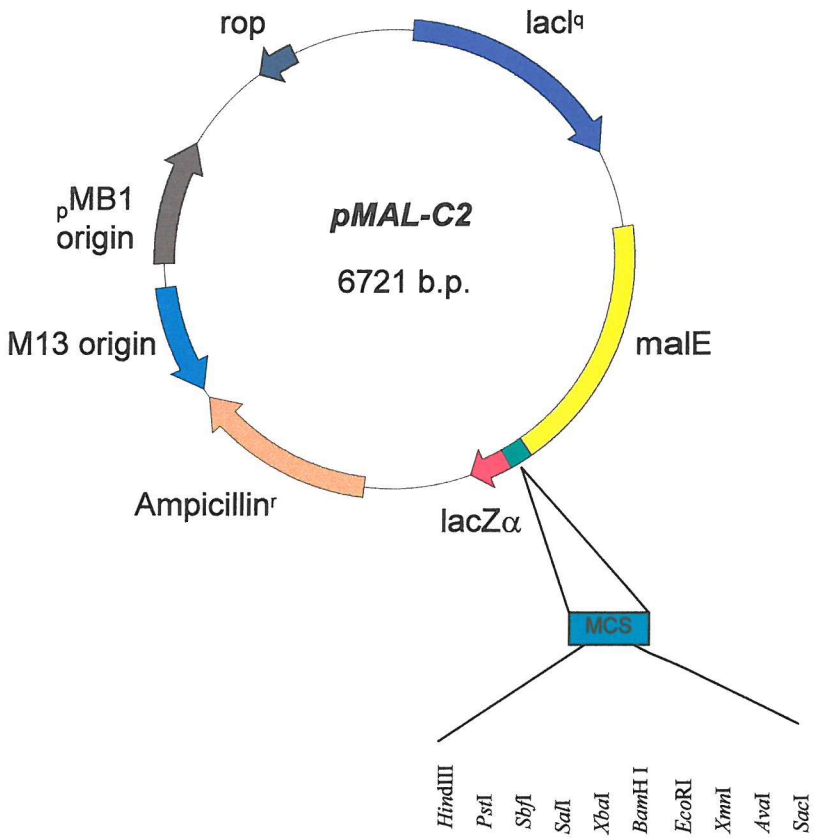



**Figure 4.10** A diagrammatic representation of the pMAL-c2 vector used for the expression of ORF6, ORF7 and ORF8. The PCR products of ORF6-8 were cloned into the *Bam*H I/*Hin*d III sites within the multiple cloning site in-frame with malE.

### 4.2.4.2 Production of PCR products for cloning into pMAL-c2

The PCR reactions were carried out using the primers in table 4.4 with Bio-x-Act DNA polymerase as described in section 2.7. Table 4.5 shows the reaction conditions for each amplicon. Following PCR, the products were analysed by gel electrophoresis and purified by absorption to a silica membrane to remove any remaining template, nucleotides or primers (2.7.3.2).

**Table 4.4 Primers used to amplify fragments of DNA for cloning into pMAL-C2 vectors. The restriction sites are shown in italics and the start or stop codon and the 3' co-ordinates are shown in bold.**

| Region amplified | Primer Name | Primer sequence 5'.....3' | Product size |
|---|---|---|---|
| ORF6 | Chp2ORF6F | GCA TAC *GGA TCC* **ATG** TGG AGA GTT TAT GAA$^{1808}$ | 156bp |
| | Chp2ORF6R | GAG CGT *AAG CTT* **CTA** GCA ATT TGC TTC$^{1957}$ | |
| ORF7 | Chp2ORF7F | GCA TAC *GGA TCC* **ATG** ACC AAG TCT TTG$^{1420}$ | 96bp |
| | Chp2ORF7R | GCA GTA *AAG CTT* **CTA** AAG AAC TTG TAG CAT$^{1515}$ | |
| ORF8 | Chp2ORF8F | GCA TAC *GGA TCC* **ATG** AGG TTA AAA ATG$^{3024}$ | 132bp |
| | Chp2ORF8R | GCA GTA *AAG CTT* **CTA** GAA CCT AAT GCC$^{3152}$ | |

**Table 4.5 PCR conditions**

| Region of genome | Denaturation | Annealing | Extension |
|---|---|---|---|
| ORF6 | 94°C for 20 sec | 55°C for 20 sec | 72°C for 9 sec |
| ORF7 | 94°C for 20 sec | 55°C for 30 sec | 72°C for 6 sec |
| ORF8 | 94°C for 20 sec | 53°C for 20sec | 72°C for 8 sec |

### 4.2.4.3 Cloning PCR fragments

The amplified fragments of DNA were then digested with restriction enzymes (section 2.6.1.1) corresponding to the two sites incorporated into the 5' and 3' termini during PCR. Following restriction digestion the DNA fragments were purified by absorption to a silica membrane (2.7.3.2) and ligated into a pMAL-c2 vector (2.9.3) digested with the corresponding restriction enzymes. The ligations were transformed into competent *E. coli* JM101 cells (2.3.3). To confirm the presence of the correct insert, plasmid DNA was prepared (2.4.2.2), and a *Bam*H I/*Hin*d III restriction digestion was carried out to release the insert. All constructs containing inserts of the correct size were sequenced (2.8) with forward and reverse primers to verify that the inserted sequences were correct.

The expression of ORF6 protein, ORF7 protein and ORF8 protein was then induced as described in section 2.9.3.

#### 4.2.4.4 Purification of the expressed proteins (ORF6 protein, ORF7 protein and ORF8 protein)

After expression each protein was purified from the *E. coli* proteins by affinity purification specific for MBP as described in section 2.9.3.3.

### 4.2.5 Dialysis of the purified proteins

During purification the MBP fusion proteins were eluted from the column using maltose. Before using the protein as an immunogen it was necessary to remove the maltose and to concentrate the protein in a smaller volume. The fractions that contained the expressed protein were combined in visking tubing and dialysed against PBS. Figure 4.11 shows the final dialysed products for ORF6 protein, ORF7 protein and ORF8 protein present after dialysis. The ORF6 fusion protein was purified in large amounts using this method and was subsequently used to produce polyclonal antisera to ORF6 protein. The concentration of the ORF7 fusion protein that was expressed was low even after the protein was concentrated in a smaller volume. The ORF8 fusion protein was purified as 4 bands of protein, probably due to degradation of the protein by *E. coli* proteases that were not removed from the sample during purification. The expression and purification of the ORF7 and ORF8 fusion proteins were repeated, but similar results were observed.

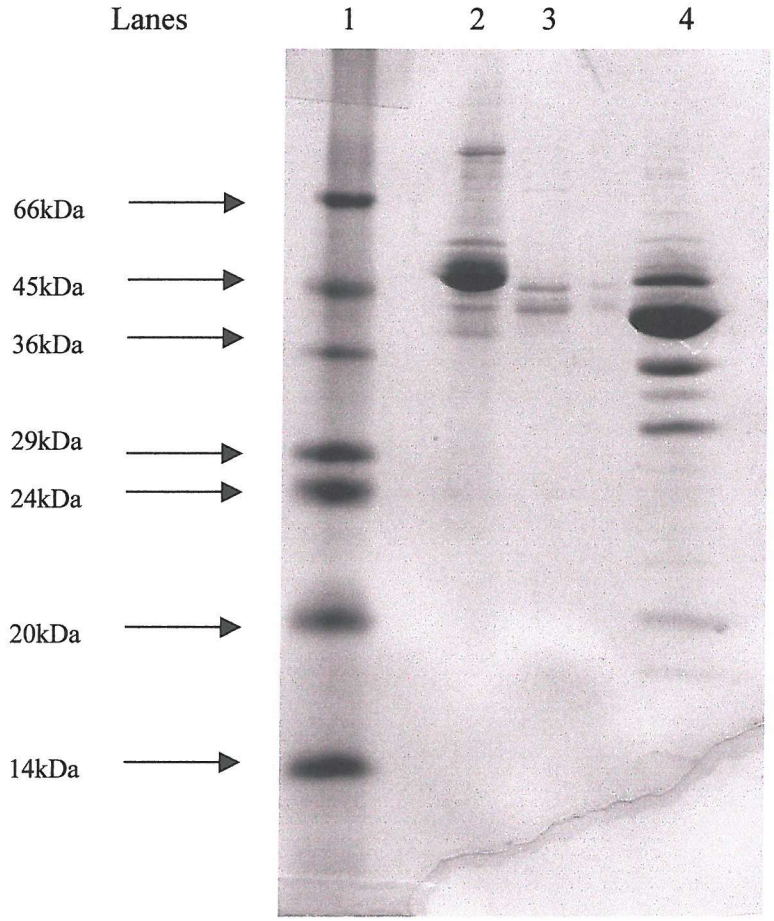



**Figure 4.11** A 12.5% SDS-PAGE gel showing the final dialysed ORF6 MBP fusion protein used for immunusation and the ORF7 MBP fusion protein and ORF8 MBP fusion protein. Lane 1. Low molecular weight marker; Lane 2. ORF6 MBP fusion protein; Lane 3. ORF7 MBP fusion protein; Lane 4. ORF8 MBP fusion protein.

### 4.2.6 Construction of pGEX-4T1 clones containing ORF7 and ORF8

As the ORF7 and ORF8 proteins could not be expressed and purified in large enough quantities to be used as an immunogen using the MBP system, these proteins were instead expressed using the glutathione s-transferase (GST) fusion protein and purification system (2.9.2). This system has been designed for the high-level expression of protein fusions with GST in any *E. coli* host. The pGEX-4T1 vectors contain a tac promoter and the pBR322 origin of replication. A diagrammatic representation of pGEX-4T1 is shown in figure 4.12.

#### 4.2.6.1 Primer design

Primers were designed to produce PCR products that contained a *Bam*H I site at the 5′ termini and a stop codon and *Xho* I site at the 3′ termini, which would allow the insertion of the PCR products in-frame with the translation initiation codon of GST into pGEX-4T1 (2.9.2). Table 4.6 shows the primers designed to amplify ORF7 and ORF8.

**Table 4.6 Primers used to amplify fragments of DNA for cloning into pGEX-4T-1 vectors. The restriction sites are shown in italics and the start or stop codon and the 3' co-ordinates are shown in bold**

| Region amplified | Primer Name | Primer sequence 5'.....3' | Product size |
|---|---|---|---|
| ORF7 | Chp2ORF7F | GCA TAC *GGA TCC* **ATG** ACC AAG TCT TTG$^{1420}$ | 96bp |
| | Chp2ORF7R GST | GCA GTA *CTC GAG* **CTA** AAG AAC TTG TAG$^{1512}$ | |
| ORF8 | Chp2ORF8F | GCA TAC *GGA TCC* **ATG** AGG TTA AAA ATG$^{3024}$ | 132bp |
| | Chp2ORF8R | GCA GTA *CTC GAG* **CTA** GAA CCT AAT GCC$^{3152}$ | |

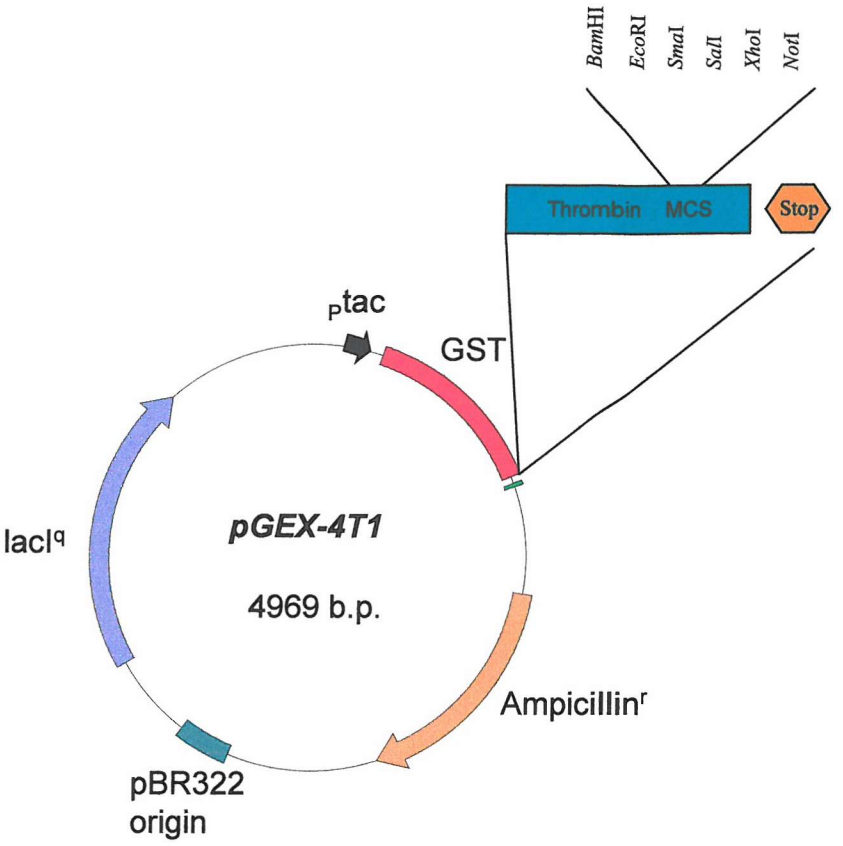



**Figure 4.12** A diagrammatic representation of the pGEX-4T1 vector used for the expression of ORF7 and ORF8. The PCR products of ORF7 and ORF8 were cloned into the *Bam*H I/*Xho* I sites within the multiple cloning site in-frame with GST.

### 4.2.6.2 Production of PCR products for cloning into pGEX-4T1

The DNA fragments were amplified using the PCR reaction conditions in table 4.7. Following PCR, the DNA fragments were analysed by gel electrophoresis to verify the size of the amplified products and purified by absorption to a silica membrane to remove any remaining template, nucleotides or primers (2.7.3.2).

## 4.7 PCR conditions

| Region of genome | Denaturation | Annealing | Extension |
|---|---|---|---|
| ORF7 | 94°C for 20 sec | 53°C for 30 sec | 72°C for 6 sec |
| ORF8 | 94°C for 20 sec | 58°C for 20sec | 72°C for 8 sec |

### 4.2.6.3 Cloning PCR fragments

The PCR products were digested with the restriction enzymes *Bam*H I and *Xho* I and purified by absorption to a silica membrane to remove any remaining nucleotides and restriction enzymes (2.7.3.1). The DNA fragments were ligated into *Bam*H I/*Xho* I digested pGEX-4T-1 vector and the ligations were transformed into competent *E. coli* JM101. To confirm that the correct construct has been produced in each case a *Bam*H I/*Xho* I restriction digest was carried to release the insert. Constructs that contained inserts of the correct size were sequenced using forward and reverse primers to verify that the inserted sequences were correct (2.8).

The expression of each protein was then induced, expressed proteins following purification are shown in figure 4.13/4.14.

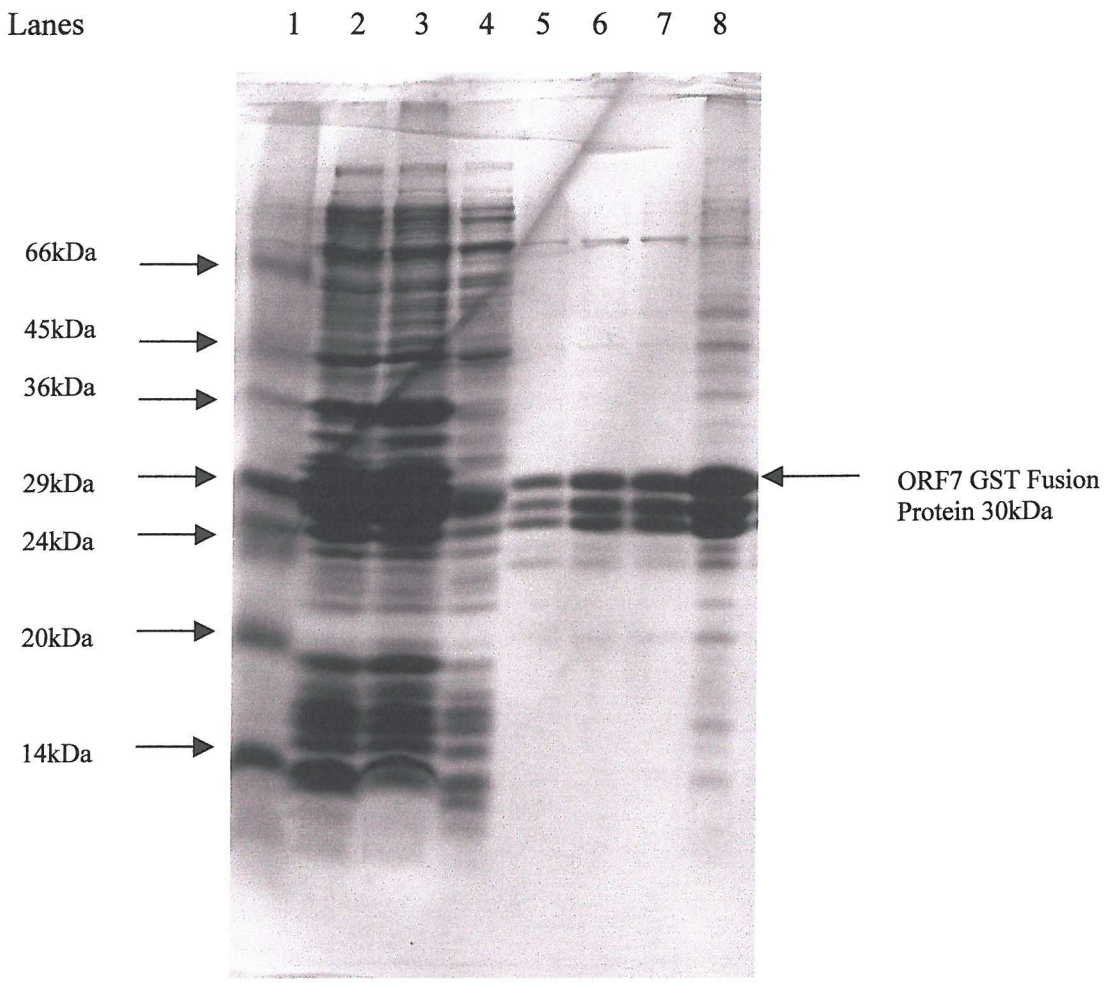



**Figure 4.13** A 12.5% SDS PAGE gel showing the column fractions collected from the purified ORF7 protein. Lane 1. Low molecular weight markers; Lane 2. Sample flow through; Lane 3. Wash 1; Lane 4. Wash 2; Lane 5-7. 20µl of the 500µl fractions eluted from the beads; Lane 8. 20µl of the bead suspension.

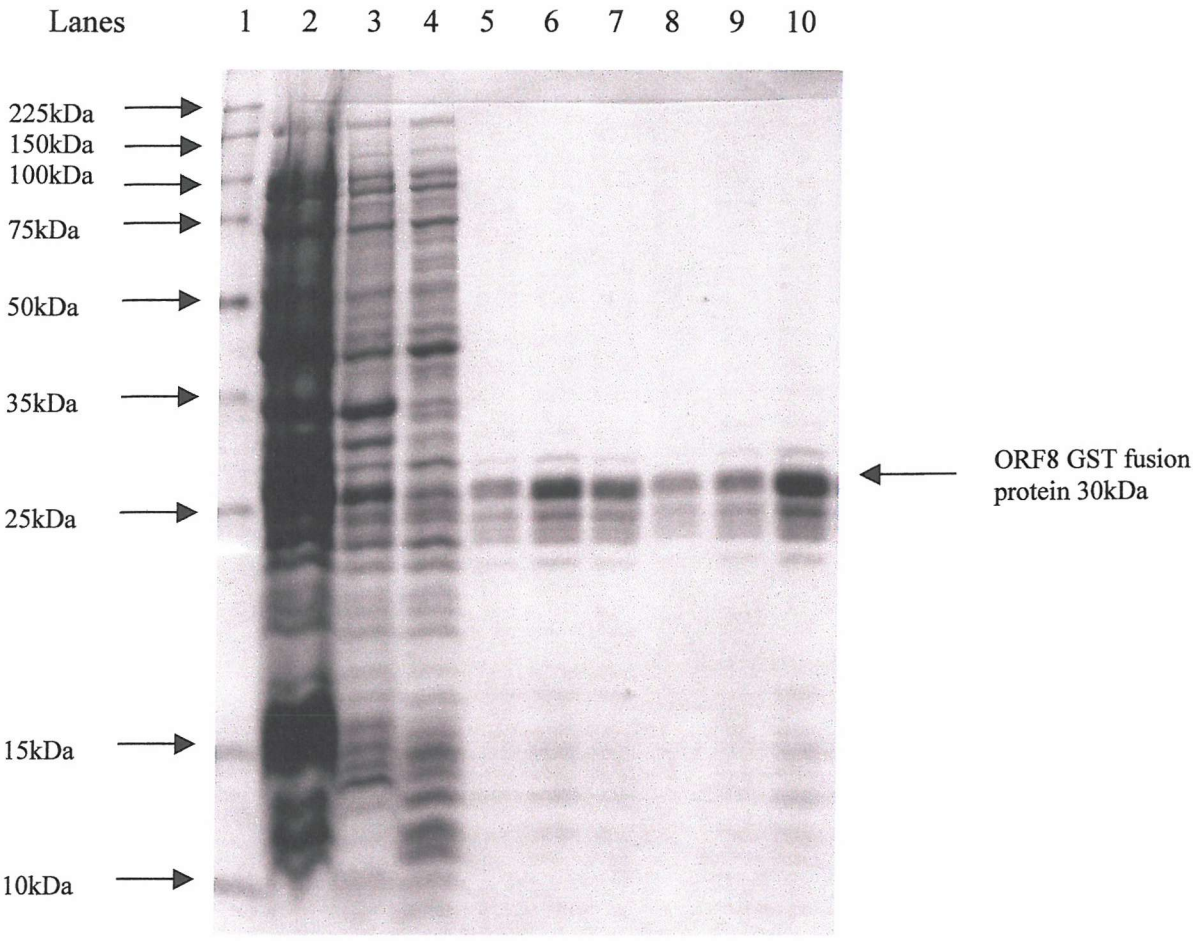



**Figure 4.14** A 12.5% SDS PAGE gel showing the column fractions collected from the purified ORF8 GST fusion protein. Lane 1. Broad range protein molecular weight marker; Lane 2. Sample flow through; Lane 3. Wash 1; Lane 4. Wash 2; Lane 5-9. 20µl of the 500µl fractions eluted from the beads; Lane 10. 20µl of the bead suspension.

#### 4.2.6.4 Purification of the expressed proteins

After expression the ORF7 and ORF8 proteins were purified from the *E. coli* proteins using affinity chromatography for GST as described in section 2.9.2. The ORF7 and ORF8 fusion proteins were subsequently used to produce polyclonal antisera.

## 4.3 Production of antisera to the purified proteins

The purified proteins from the *E. coli* expressions of ORF3-8 and the VP2 peptide were used to raise polyclonal antisera as described in section 2.10. Polyclonal antisera to these proteins needed to be raised so that the expression of each of the individual gene products could be investigated. Rats were initially chosen to produce polyclonal antisera to these proteins because they give a good yield of antisera and are a convenient animal to use for the amounts of sera needed. Pre-immune sera was obtained from each rat prior to immunisation and tested by immunoflourescence to check that it did not react with Chp2 infected *C. abortus* inclusions (2.19). No immunoflouresence was observed between Chp2 infected *C. abortus* inclusions and any of the pre-immune sera. The pre-immune sera were also immunoblotted against each of the proteins to ensure there was no pre-immune reaction. The pre-immune sera did not show any pre-immune reactivity with VP3, ORF4 protein or the ORF5 protein. Two rats were subsequently immunised per protein with VP3, ORF4 region B protein and ORF5 protein (2.10).

Pre-immune sera taken from three different rats showed pre-immune reactivity with the ORF6, ORF7 and ORF8 fusion proteins. Consequently pre-immune sera was obtained from a different species (mouse) to test against these proteins. A large number of mice had to be screened to find pre-immune mouse sera that did not react to Chp2 infected *C. abortus* inclusions by immunoflourescence and by immunoblot against the ORF6, ORF7 and ORF8 fusion proteins (2.19). Eventually animals fulfilling these criteria were found where the pre-immune sera did not show any pre-immune reactivity to either Chp2 infected *C. abortus* inclusions or the ORF6, ORF7 and ORF8 fusion proteins. One mouse was subsequently immunized per protein with the ORF6, ORF7 and ORF8 fusion proteins. Two mice were also immunised with the VP2 peptide as two mice were available that showed no reactivity to Chp2 infected *C.*

*abortus* inclusions. The amount of each recombinant protein used for immunisation can be seen in table 4.8.

**Table 4.8 Approximate protein concentration (determined by SDS-PAGE) and volumes to be used per immunisation.**

| Protein | Concentration | Volume per immunisation |
|---|---|---|
| VP2 | 40μg /20μl | 50μl |
| VP3 | 25μg /20μl | 80μl |
| ORF4 protein | 20μg /20μl | 100μl |
| ORF5 protein | 100μg /20μl | 20μl |
| ORF6 protein | 40μg /20μl | 50μl |
| ORF7 protein | 30μg /20μl | 60 μl |
| ORF8 protein | 40μg /20μl | 50 μl |

The hyper-immume polyclonal antisera from each rat/mouse were then analysed by immunoblotting against the relevant purified protein. The polyclonal antisera raised reacted strongly with the recombinant protein it was raised against. Although some of the antisera also cross-reacted with *E. coli* proteins that had apparently not been removed during purification (figure 4.15/4.16).

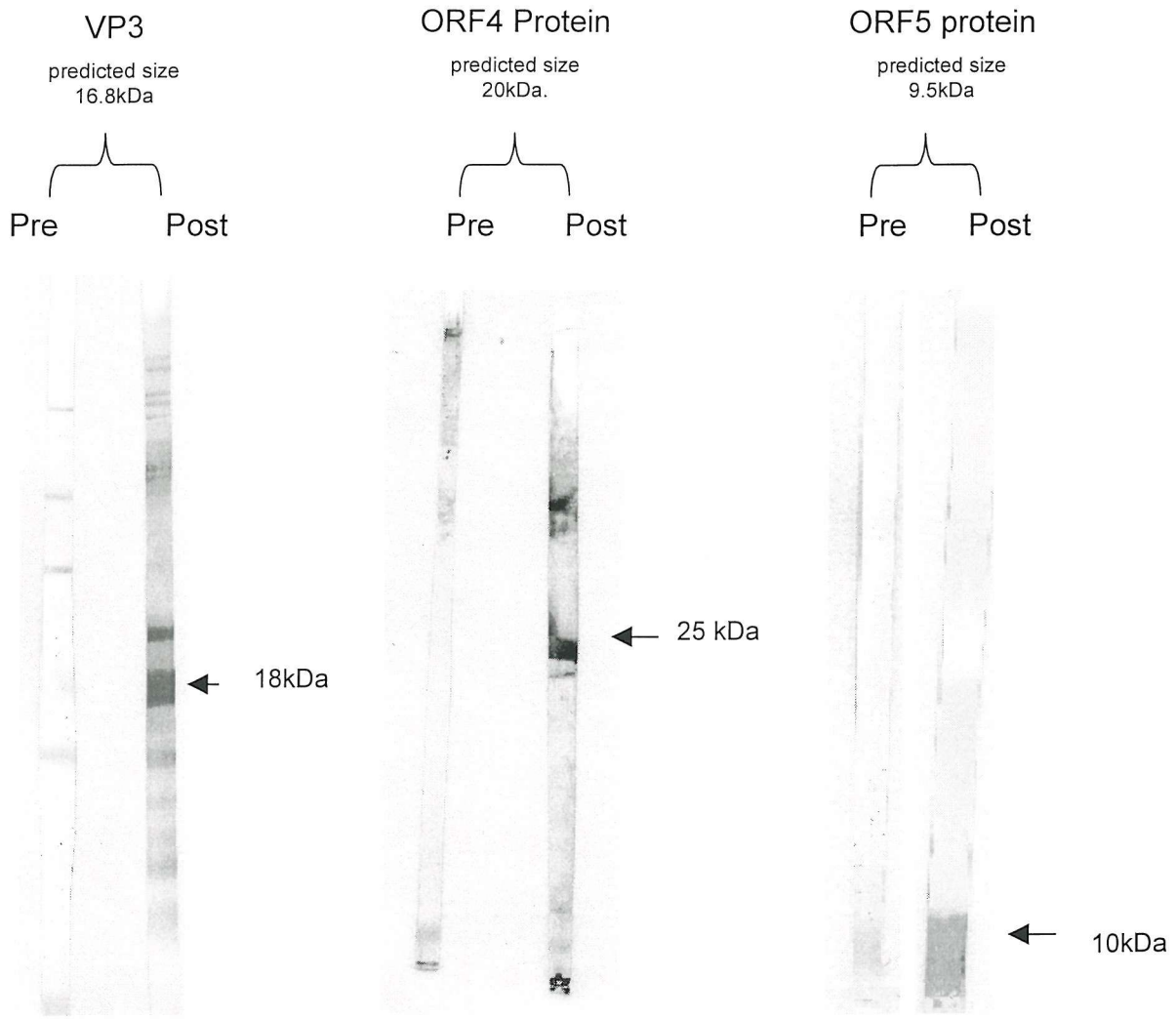



**Figure 4.15** Pre and hyperimmune immunoblots with rat polyclonal antisera (1/1000 dilution) to VP3, ORF4 protein and ORF5 protein. The antisera were blotted against the respective immunising agent run on a 12.5% SDS-PAGE gel. The positions of the recombinant proteins are indicated.

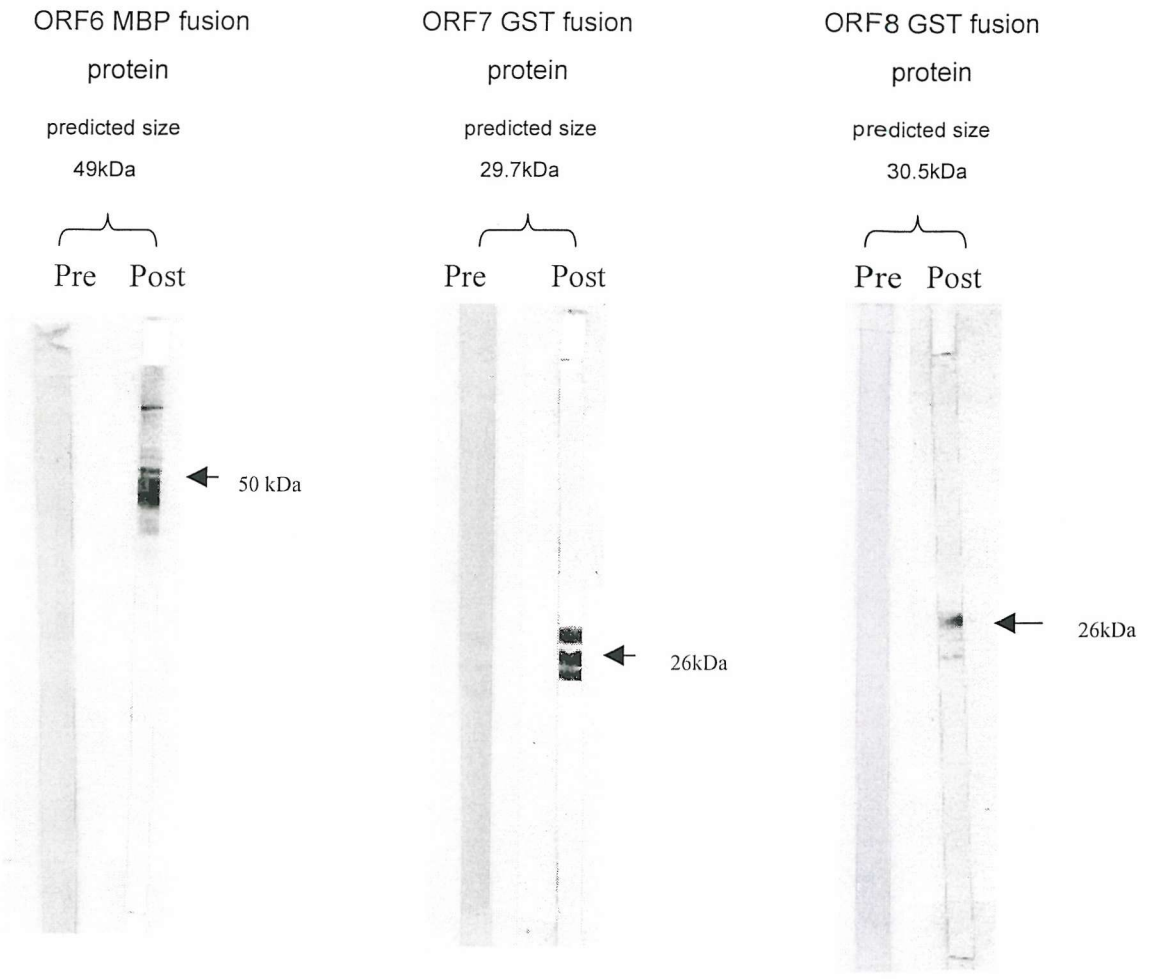



**Figure 4.16** Pre and hyperimune immunoblots with mouse polyclonal antisera to ORF6 MBP fusion protein, ORF7 GST fusion protein and ORF8 GST fusion protein. The polyclonal antisera (1/1000 dilution) were blotted against the respective immunising agent run on a 12.5% SDS-PAGE gel. The position of the recombinant proteins are indicated.

## 4.4 Radio-immune precipitation assay

It is known from a previous investigation that monoclonal antibody 55 can be used to detect the *in vitro* transcription/translation product of ORF1 using a RIPA (section 4.2.1.4). The polyclonal antisera produced to the proteins encoded by ORF2-5s were similarly used to determine whether each of the proteins could be detected from the *in vitro* transcription/translation products of ORF2-5 using a RIPA (2.15). The RIPA products were analysed by SDS-PAGE followed by autoradiography (2.16). The antisera to VP3 reacted with a protein of 16kDa produced by the ORF3 construct, the antisera to ORF4 protein reacted to a protein of 40kDa produced by the ORF4 construct and the ORF5 protein antisera reacted with a protein of 10kDa produced by the ORF5 construct. This confirmed that the antisera were reacting with the correct proteins. (figure 4.17). VP2 could not be detected using the antisera raised to the VP2 peptide. VP2 is estimated to be 21kDa, a protein of this size is produced by the ORF2 construct (section 4.2.1.4). It is possible that the VP2-peptide antiserum does not recognise the folded intact protein.

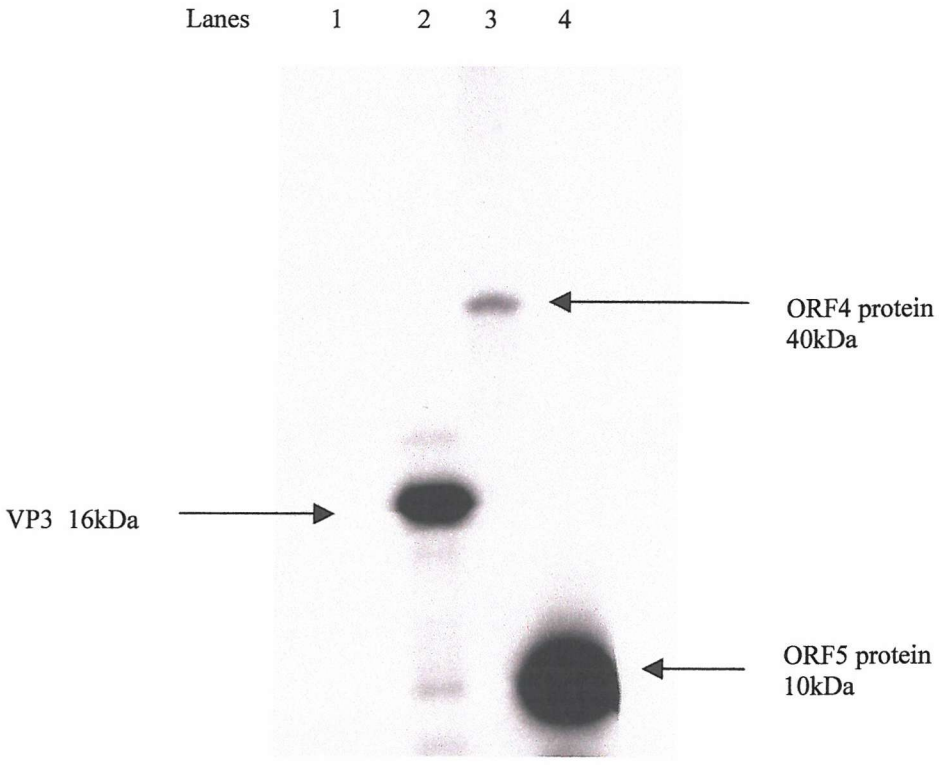


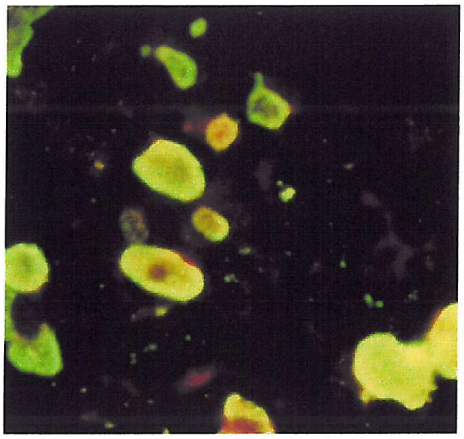

**Figure 4.17** Autoradiograph showing the proteins that immunoprecipitated with polyclonal antisera raised to VP2, VP3, ORF4 protein and ORF5 protein from S30 *in vitro* transcription/translation reactions for pRSETA containing ORF2-ORF5. Lane 1. pRSETAORF2; Lane 2. pRSETAORF3; Lane 3. pRSETAORF4; Lane 4. pRSETAORF5.

## 4.5 Immunoflourescence of Chp2 infected *C. abortus* inclusions with the polyclonal antisera produced to the proteins encoded by ORF1-8.

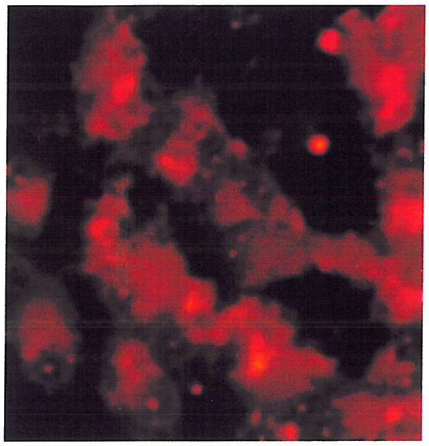
The polyclonal antisera produced to the Chp2 proteins were used to analyse which of the ORFs are expressed during infection of *C. abortus* inclusions with Chp2 as assessed by immunoflourescence (2.19). Bacteriophage-infected inclusions were screened 66 hours p.i with each polyclonal antisera and the monoclonal antibody specific to VP1 (55). The results from the analysis showed that VP1, VP3, ORF5 protein and ORF7 protein are expressed in Chp2 infected *C. abortus* inclusions. The expression of VP2, ORF4 protein, ORF6 protein and ORF8 protein could not be detected (figure 4.18). This indicates that either, these proteins are not expressed during infection or that the expression of each protein is low.

The areas of fluorescence, which corresponded to the location of each protein varied between the different proteins. VP1 and VP3 were concentrated primarily within the chlamydial inclusions with speckles of the proteins detected around the inclusions. ORF5 protein and ORF7 protein were contained within the chlamydial inclusions and no protein was detected in the surrounding areas. Interestingly, ORF7 protein seemed to be mainly located around the inclusion membrane. These observations are consistent with the predicted functions of the proteins for example the two structural proteins (VP1 and VP3) are located within the inclusions and as speckles around the inclusions, which probably corresponds to mature Chp2 virions. Similarly ORF5 protein, which is not present within the mature Chp2 virion, is only seen within inclusions.

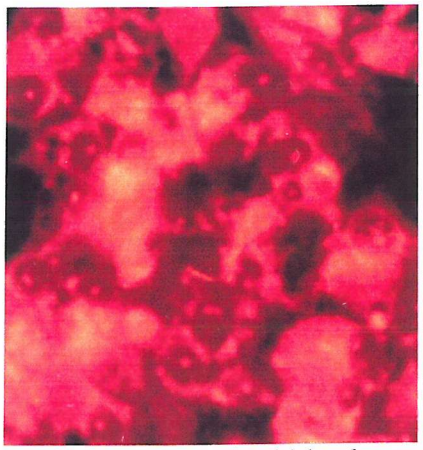
**Figure 4.18** Fluoresecent micrograph of Chp2 infected *C. abortus* inclusions within BGMK cells.



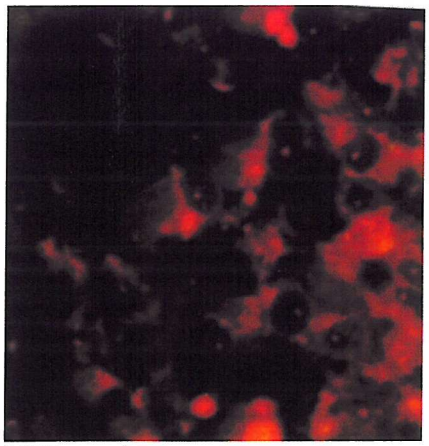
VP1 (mAb 55E11)



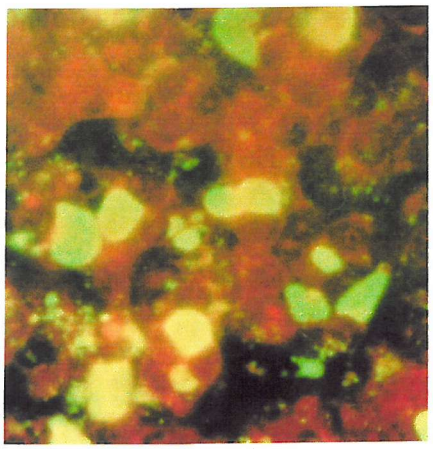
VP2 peptide pre-immune sera



VP2 peptide terminal bleed sera

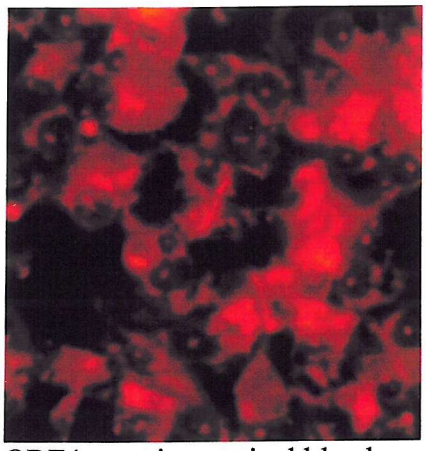


VP3 pre-immune sera

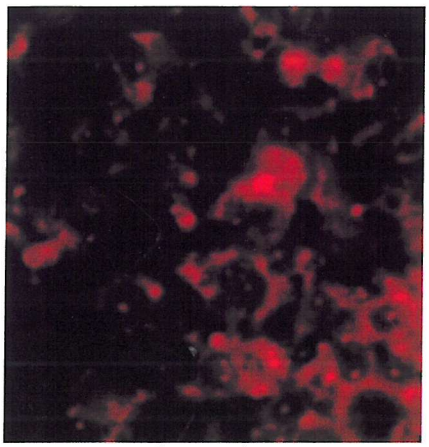


VP3 terminal bleed sera

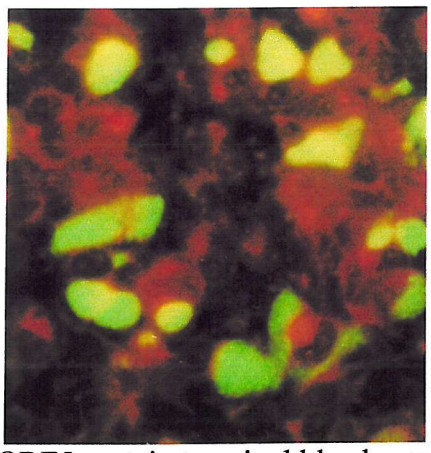
ORF4 protein pre-immune sera



ORF4 protein terminal bleed sera

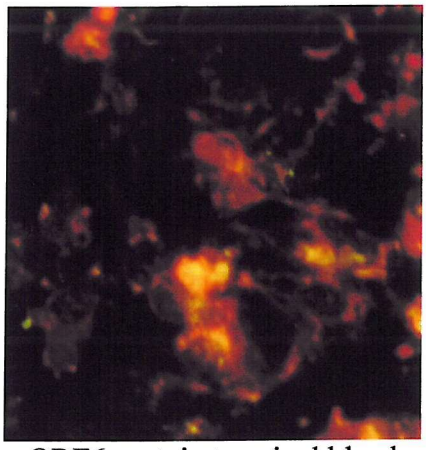


ORF5 protein pre-immune sera

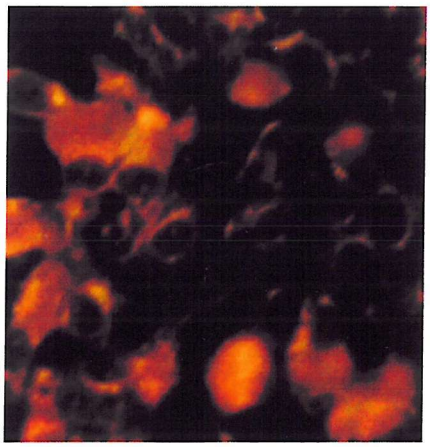


ORF5 protein terminal bleed sera

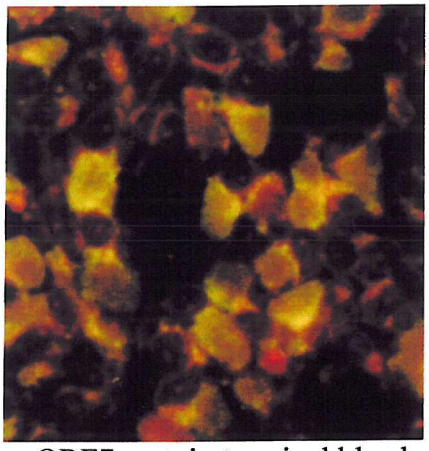
ORF6 protein pre-immune sera



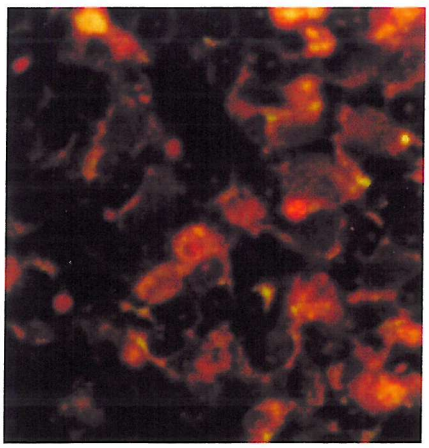
ORF6 protein terminal bleed sera



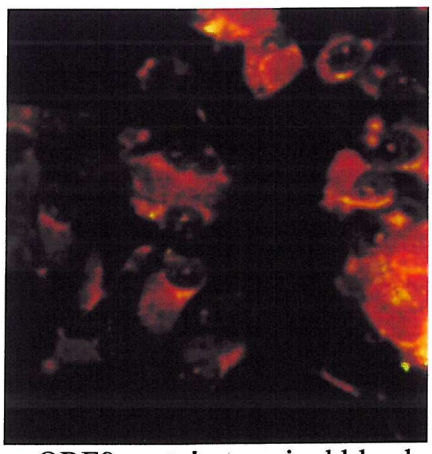
ORF7 protein pre-immune sera



ORF7 protein terminal bleed sera



ORF8 protein pre-immune sera



ORF8 protein terminal bleed sera

## 4.6 Western blot of semi purified Chp2

To identify the structural proteins present in the Chp2 virion, the polyclonal antisera produced to the proteins encoded by ORF2-ORF8 and the VP1 monoclonal antibody were immunoblotted against semi-purified Chp2 particles (2.12.5). It was predicted prior to this investigation that the immuno-specific reagents to VP1, VP2, VP3 and ORF8 protein would react with the semi-purified Chp2 particles. This is because VP1, VP2 and VP3 have been shown to be present in the chlamydiaphage virion (Storey *et al.*, 1989) and ORF8 protein is a similar protein to the ϕX174 DNA binding protein (protein J), which is present in the ϕX174 virion. The results from this investigation showed that only VP3 and ORF4 protein could be detected by the antisera in the semi-purified Chp2 particles, as the VP3 polyclonal antisera reacted strongly to a protein of 16 kDa and the ORF4 protein polyclonal antisera reacted to a protein of 38kDa (figure 4.19). The presence of ORF4 protein was surprising but is probably due to this protein co-purifying with the Chp2 particles. No reactivity between any of the other antisera or the VP1 monoclonal antibody could be seen. The reason VP1 could not be detected may be due to the VP1 monoclonal antibody not recognising this protein in its denatured form, perhaps due to it reacting with an epitope that is only present when this protein is folded. As VP2 could not be detected from the *in vitro* transcription and translation products of the ORF2 construct (4.2.1.4) it seems likely that the VP2 peptide antisera is unable to detect intact VP2. ORF8 protein was not detected possibly due to its small size or low rate of translation.

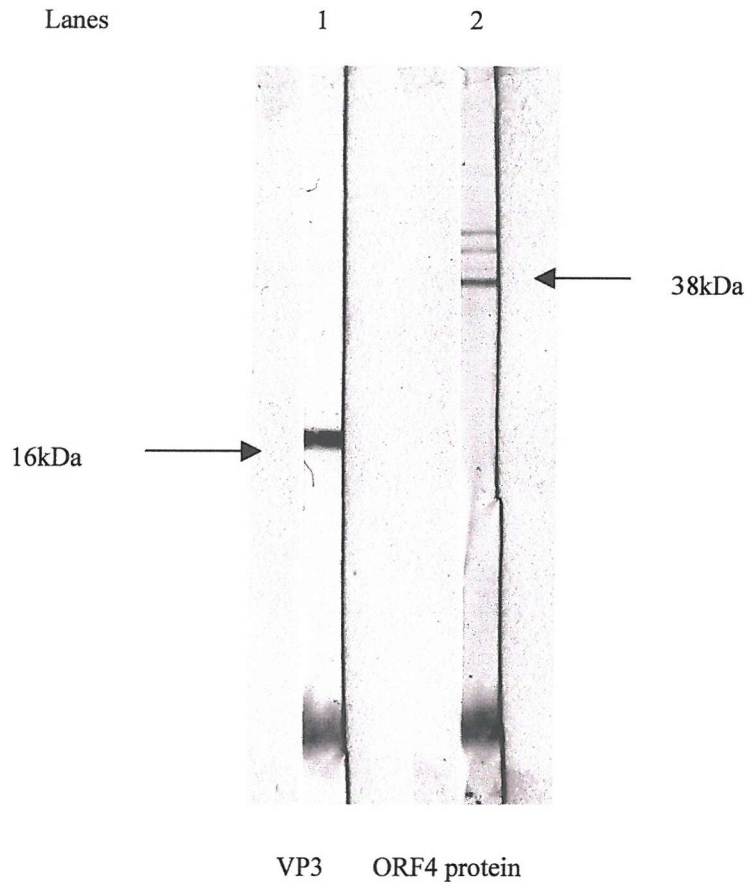



**Figure 4.19** Immunoblots with rat polyclonal antisera (1/1000) to VP3 and ORF4 protein. The antisera were blotted against semi-purified Chp2 particles run on a 12.5% SDS PAGE gel. The positions of the Chp2 proteins are indicated.

## 4.7 Discussion and Future work

The production of immuno-specific reagents to the proteins encoded by ORFs1-8 during this investigation represented an important step in making the tools necessary to identify which of these proteins are expressed during infection.

The immuno-specific reagents produced to each protein were used to screen Chp2 infected *C. abortus* inclusions and semi-purified Chp2 particles. It was found VP1, VP3 and ORF4, ORF5 and ORF7 protein are expressed during infection. The expression of VP2, ORF6 protein and ORF8 protein were not detected. As VP2 of Chp1 is known to be present in the Chp1 virion (Storey *et al.*, 1989) it is unlikely that the Chp2 VP2 is not expressed. It is possible that the VP2-peptide did not raise a strong immune response in the host. This could be due to the hosts immune system being compromised throughout the immunisations by a bacterial or viral infection or due to insufficient amounts of antigen being used to raise a good antibody response. As there are no proteins with significant sequence similarity to ORF6 protein in ϕX174 it is likely that this ORF does not encode protein. The protein encoded by ORF8 would be a similar protein to the ϕX174 J protein. It is possible that due to the smaller size of the Chp2 genome a DNA binding protein is not required. However, due to the important role protein J performs during ϕX174 genome packaging it would be surprising if a similar protein was not utilised by Chp2.

It was demonstrated during this investigation that VP1, VP3 and the ORF5, ORF6, ORF7 and ORF8 proteins could be expressed in *E. coli*. The expression of VP2 and ORF4 protein could not be demonstrated. The reason why VP2 could not be expressed may relate to the function of this protein. VP2 is thought to be the chlamydiaphage DNA pilot protein. The ϕX174 pilot protein (protein H) is enclosed within the mature bacteriophage capsid and is essential for DNA injection into the host. It is possible that the toxicity of this protein stems from its ability to enter the host during DNA injection, perhaps due to it interfering with bacterial membranes. However, no inhibitory effect or decrease in the number of bacterial cells present after IPTG induction was observed in *E. coli* containing the ORF2 recombinant plasmid. The ϕX174 protein H could not be expressed from a plasmid containing complete

gene H (Van Der Avoort *et al.*, 1983). However, the plasmid containing the complete gene H also contained the region between gene H and gene A, which is the ϕX174 incompatibility sequence. The incompatibility sequence is located between nucleotides 3904 and 3989 and includes the intercistronic region between gene H and gene A, plus a small portion of each of these genes (Van Der Avoort *et al.*, 1982). Its function is to bind the RF DNA to a site where replication enzymes are present. ORF2 is situated in a different genomic location than gene H. Therefore it seems unlikely that the reason VP2 was not expressed was due to a Chp2 incompatibility region.

The ORF4 protein shares regions of homology to the ϕX174 A protein. A protein nicks RF DNA at the origin allowing viral DNA synthesis. A protein could not be expressed in *E. coli* (Van Der Avoort *et al.*, 1983). A smaller in frame ORF, gene A* is also present within gene A. It was thought that the A protein was not expressed due to the consequential expression of A* protein. A* protein causes host cell shut down by introducing nicks in single-stranded regions of the *E. coli* chromosomal DNA causing arrest of host cell DNA synthesis (Eisenberg and Ascarelli, 1981). No A* protein homologue has been identified in Chp2. Gene A* is situated within the 3' region of gene A. As the 3' termini of ORF4 has been expressed (ORF4 region B protein) it seems unlikely that expression of a Chp2 A* homologue is responsible for the inhibition of complete ORF4 protein expression (unless complete A* protein was not present in the ORF4 region B protein).

This section will now offer future research ideas to assist in identifying the functions of previously uncharacterised proteins (ORF4 protein, ORF5 protein, ORF6 protein, ORF7 protein and ORF8 protein) based on comparisons with the ϕX174 proteins. The ORF4 protein shows two regions of homology with the C-terminus of the ϕX174 A protein. A protein nicks the ϕX174 RF DNA at the origin, which allows the initiation of rolling circle replication. ORF4 protein was not detected in Chp2-infected *C. abortus* inclusions during these investigations, probably due to a low rate of transcription and consequently a low translation rate of ORF4 protein. This corresponds to the low number of gene A transcripts produced in ϕX174 infected cells (Hayashi *et al.*, 1976). However, it was detected in semi-purified Chp2 particles. To

show that ORF4 protein has endonuclease activity and functions as the initiator of DNA synthesis, it would first be necessary to show that ORF4 protein binds to RF DNA. ORF4 protein could first be expressed using an *in vitro* transcription/translation system and then incubated both with and without the vector pUC18Chp2 (which contains the full-length Chp2 genome ligated into the vector pUC18). The results from the incubation could then be analysed by agarose gel electrophoresis. Any difference between the mobility of the DNA that was incubated with ORF4 protein would be due to ORF4 protein covalently attaching to the RF origin of replication. The origin of replication could similarly be determined by ligating fragments of the Chp2 genome into plasmid vectors then incubating the DNA both with, and without ORF4 protein. Any difference in mobility of the DNA between the two reactions would again be due to ORF4 protein covalently attaching to the DNA at the origin, and hence indicate the presence of the origin of replication.

These suggestions are based on the observation that the mobility of plasmid vectors containing the ϕX174 origin of replication through an agarose gel was affected by incubation with the A protein (Brown *et al.*, 1983). It was also shown that ss DNA was produced when plasmid DNA containing the ϕX174 origin of replication was incubated with A protein in a RF→SS DNA replication system. It would be interesting to carry out similar studies with the Chp2 origin of replication and ORF4 protein.

The function of ORF5 protein is unknown; comparison with ϕX174 proteins suggests that the only other protein of similar size, charge and location is protein C (table 3.10). Protein C is a ss DNA binding protein that is required for the initiation of stage III DNA synthesis (Aoyama *et al.*, 1983). The function of protein C was shown by studying the binding of protein C to $^{14}C$ radiolabeled single-stranded and double-stranded DNA molecules using a Millipore filler assay. It was shown that protein C bound non-specifically to single-stranded DNA and had a lower affinity to double-stranded DNA. Similar studies could be carried out on purified ORF5 protein.

ORF6 encodes a polypeptide of 52 residues in Chp3, Chp2, ϕAR39 and ϕCPG1. This protein could not be detected during Chp2 infection of *C. abortus* inclusions and there are no proteins with significant sequence similarity to ORF6 protein in ϕX174.

The function of the protein encoded by ORF7 is also unknown and no significant homologies exist between the ORF7 protein and proteins encoded by ϕX174 genes. As suggested in section 3.5.2.7, due to the presence of a N-terminal hydrophobic domain in ϕCPG1 and ϕAR39, ORF7 protein could be the chlamydiaphage lysis protein (similar to the ϕX174 protein E). Protein E causes lysis of *E. coli* cells by inhibiting translocase I catalysed reactions (Bernhardt *et al.*, 2001). As the Chp2 ORF7 protein could be expressed in *E. coli* cells, it is possible that it is not functional as a lysis protein in *E. coli*. The ϕMH2K lysis protein, protein N is a hydrophobic protein and gene N is situated within the ϕMH2K ORF4 gene. This protein is 109 residues in size (Brentlinger *et al.*, 2002). There is a small ORF situated in this region in the chlamydiaphages (ORF9), but the protein encoded by ORF9 does not show any significant sequence identity to protein N and is half the size of protein N (50 residues). However, if this ORF does encode the Chp2 lysis protein it could explain why the expression of complete ORF4 protein was not observed *in vivo*. This could be investigated further by producing a mono-specific reagent to the protein encoded by ORF9.

The ORF 8 protein is 44 amino acids in length and contains 19 basic residues and no acidic residues. It is thought to be a DNA binding protein by analogy to the ϕX174 J protein. Protein J is a small basic protein of 37 amino acids and contains 12 basic residues and no acidic residues. Protein J is thought to participate in packaging the viral DNA into the prohead, by binding DNA using simple charge interactions between the negatively charged DNA backbone and basic amino acid residues (Jennings and Fane, 1997). The function of the ORF8 protein could be investigated by expressing this protein *in vitro* then incubating it both with and without RF DNA. The reactions could then be analysed by agarose electrophoresis, any difference in mobility between the two reactions would be due to ORF8 protein binding to the DNA.

# Chapter 5

# Chp2 Promoter Regions

## 5.1 Introduction

A typical prokaryotic DNA transcriptional assemblage contains a promoter region where the RNA polymerase binds to begin transcription, the sequence to be transcribed plus a terminator or hairpin structure where transcription ends. The characteristic eubacterial promoter region contains four major components: the point where transcription begins; two consensus zones which are located upstream of the start codon (+1) typically positioned at nucleotides -10 (Pribnow, 1975a; Pribnow, 1975b) and –35 (Schaller *et al.*, 1975) and the region located between the two consensus sequences.

The identification of chlamydial promoters by sequence comparison has been a difficult process due to the lack of overall sequence identity with the characteristic eubacterial promoter. The chlamydial RNA polymerase is able to recognize promoter regions that are both similar and divergent from the eubacterial $\sigma^{70}$ consensus (Mathews and Sriprakash, 1994). However, no overall chlamydial consensus promoter has been identified, although a late stage promoter consensus has been proposed. Many chlamydial genes have been shown to contain multiple promoter regions. It is thought that multiple promoters are utilised as a mechanism for the regulation of gene expression (Timms and Mathews, 2002).

It would be useful to identify regions which act as promoters in Chp2, as this would help to understand how gene expression is regulated. However, no promoter regions have been identified in the Chp2 genome that show homology to the characteristic eubacterial promoter. As nothing is known about the transcriptional mechanisms of the chlamydiaphages, and no chlamydial consensus promoter has been determined, an alternative approach is to make comparisons between chlamydiaphages and other single stranded DNA viruses (such as ϕX174 and SpV4) where comprehensive transcriptional studies have been made.

In ϕX174, transcription occurs throughout the bacteriophage infection cycle. No switch between different classes of genes has been observed. Transcription of mRNA occurs directly from the complementary strand of DNA and promoters are situated before genes A, B and D. Four terminator sites are also present in the ϕX174 genome, after genes F, G, H and J, two of which are active only in the presence of the rho termination factor (figure 5.1) (Hayashi *et al.*, 1976).

The promoter and terminator sites were identified by Hayashi *et al.,* 1976. In this work the ϕX174 mRNAs were isolated and separated by gel electrophoresis then visualised by hybridization to RF DNA restriction endonuclease cleaved fragments. Gene A mRNA was difficult to detect therefore thought to be unstable. The amount of each species of mRNA could be estimated, as the molecular weight and cistrons present in each RNA species were known. These studies showed that ϕX174 does not have any elaborate mechanism of gene regulation, although it does have a primitive attenuation system. This system is based on the positions of the promoters and terminators in the bacteriophage genome. Genes coding for the most abundantly expressed proteins are situated close to promoter regions whereas genes coding for proteins that are required in smaller amounts are situated at the distal end of operons (Hayashi *et al.*, 1988).

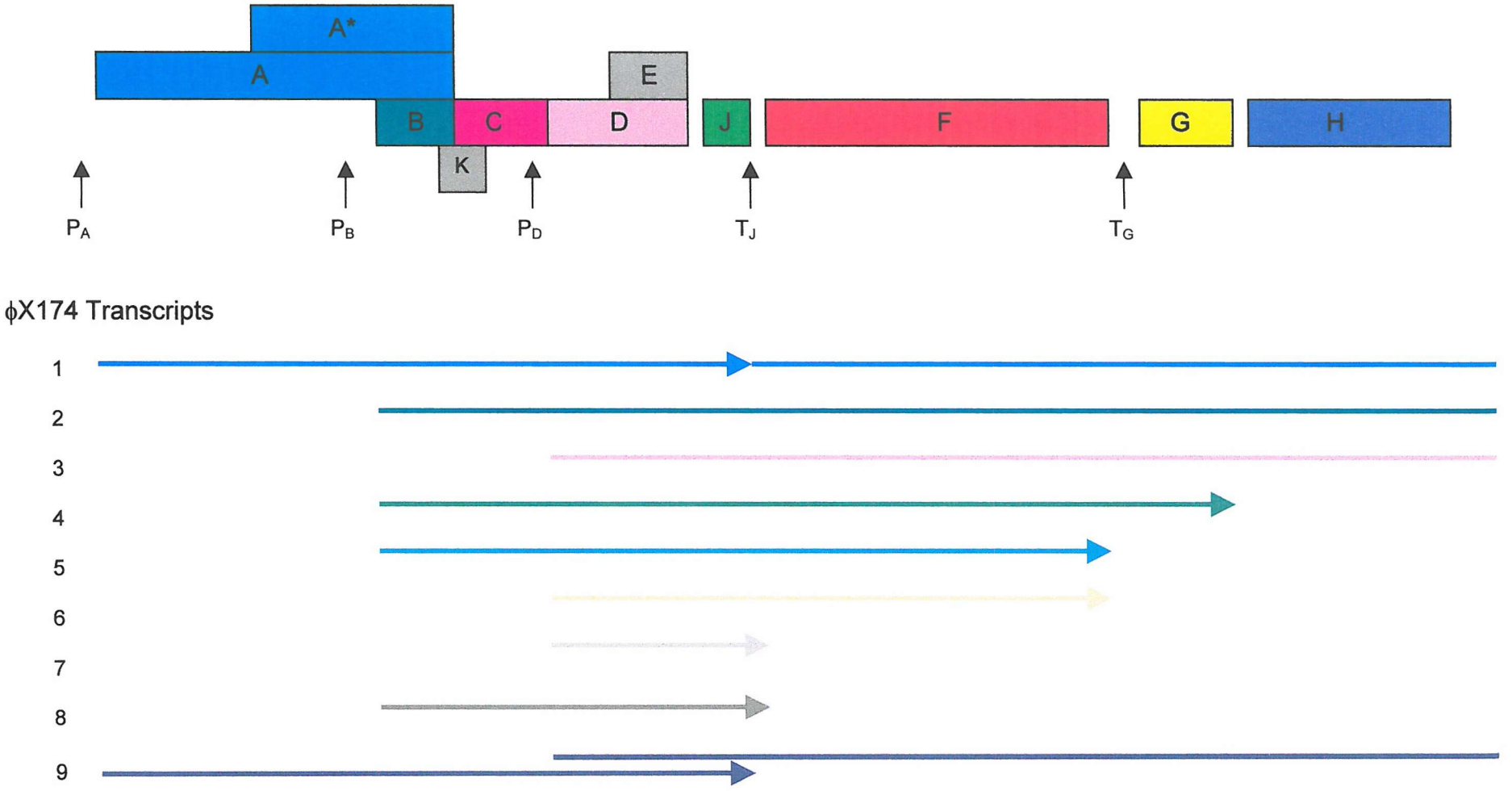



**Figure 5.1** A linear representation of the ϕX174 genome. The main RNA species in infected cells are indicated.

In SpV4 four promoters have been identified, P3 located before the SpV4 equivalent of the Chp2 ORF4 (ORF2), P1 and P0 located after the SpV4 ORF2 and P2 situated before the SpV4 ORF3 (Stamburski *et al.*, 1990). All promoter regions show homology to the eubacterial promoter consensus sequences TATAAT (-10) and TTGACA (-35). Only one termination sequence has been identified in the SpV4 genome located after the SpV4 ORF4. The main transcripts produced by SpV4 *in vivo* were mapped by Northern blot analysis. Total RNA was extracted from SpV4 infected and uninfected spiroplasma using radiolabeled probes to SpV4 DNA. Four main transcripts were identified at 2.7kb (transcript 1), 3.4kb (transcript 3), 4.4kb (transcript 2) and 7.8kb (transcript 4) see figure 5.2. All of the transcripts terminated at a single shared position. The 5' end of the 3.4kb transcript was mapped to P1 or P0. The transcription initiation start points of the other transcripts were consistent with the location of P2 and P3. The 2.7kb transcript begins at P2 and the 4.4kb transcript is initiated at P3. The 7.8kb transcript is initiated at P1 or P0 and occurs when the RNA polymerase by passes the terminator and undergoes a second complete circuit around the SpV4 genome before termination (figure 5.2) (Stamburski *et al.*, 1990).

Like ϕX174, transcription control in SpV4 is thought to be due to termination events this assumption is based on the arrangements of the promoters and terminator regions (Renaudin and Bove, 1994).

The P0 and P1 SpV4 promoter regions are active in *E. coli* cells (Stamburski *et al.*, 1990). This was identified by ligating DNA regions containing the promoters into the promoter selection vector pKK 232-8 then transforming the vectors into *E.coli* cells. The pKK 232-8 vector contains a CAT gene minus its promoter, hence the gene can only be expressed if a fragment of DNA, containing a promoter sequence that is recognised by the *E. coli* RNA polymerase, is inserted before the CAT gene. Clones that contained active promoters were identified by their resistance to chloramphenicol. Positive clones were then further examined by Southern blot hybridization using SpV4 probes to confirm the presence of the bacteriophage promoters (Stamburski *et al.*, 1990).

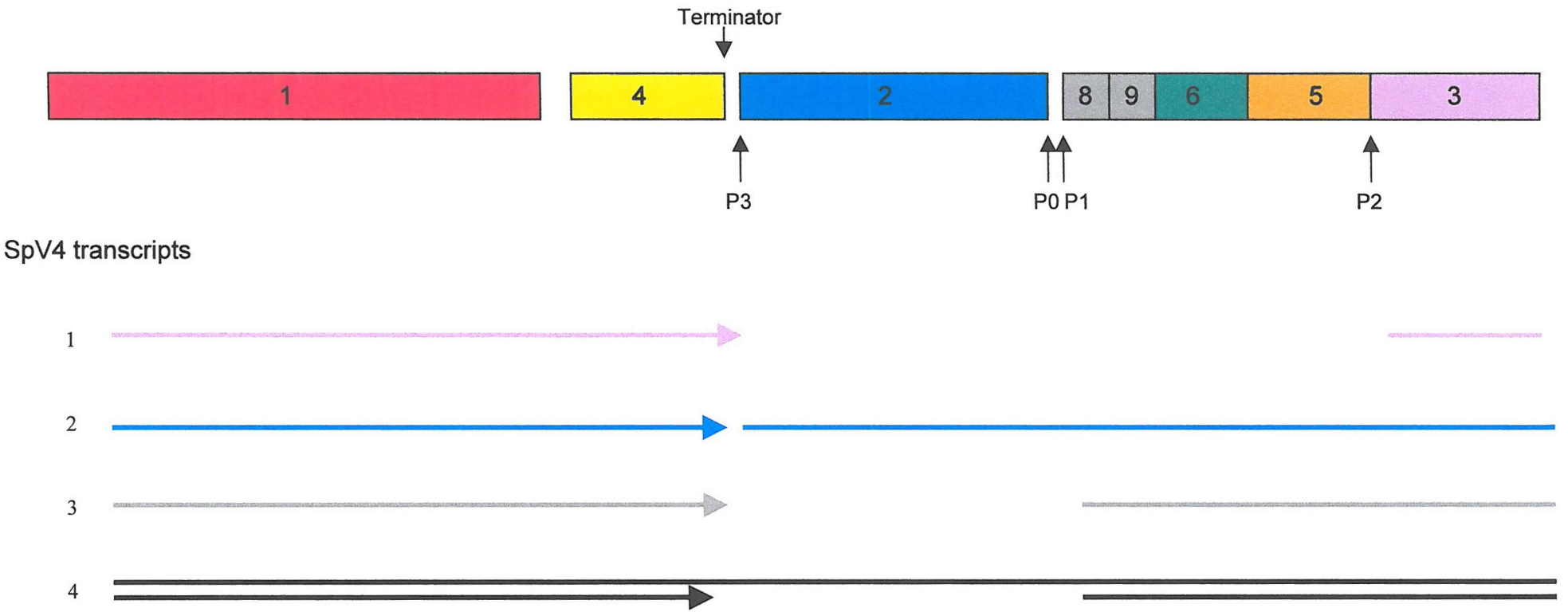



**Figure 5.2** A linear representation of the SpV4 genome. The main RNA species in infected cells are indicated.

The overall objective of my study was to investigate the transcriptional properties of Chp2; primarily to identify the Chp2 transcripts and to determine if any of the chlamydiaphage promoters are active in *E. coli* or an *in vitro* transcription-translation coupled system.

## 5.2 Northern hybridisation analysis of cells containing Chp2 infected *C. abortus*

The aim of this study was to identify the major transcripts produced by Chp2 in *C. abortus* infected cells and then map each transcript to its 5' initiation point and 3' termination signal. From a previous Northern blot hybridisation time course experiment it is known that Chp2 transcripts are seen at 48 and 66 hours after infection (figure 5.3). Based on this observation RNA was initially extracted from Chp2 infected *C. abortus* cells, 66 hours after infection.

RNA was extracted from Chp2 infected, chlamydial infected cells, using a SV total RNA isolation system (Promega) as described in section (2.4.3). After the RNA had been extracted from the cell lysate, 10µl was run on a denaturing formalydehyde gel as described in section (2.17.1), and then blotted onto a nylon membrane (2.17.3). The membranes were stored at –20°C until required and any unused RNA was stored in nuclease-free water in 20µl volumes at –80°C.

### 5.2.1 Probing RNA for the presence of Chp2 transcripts

A hybridisation was performed on the membrane bound RNA as described in section (2.18) with $^{32}$P dCTP radiolabelled probes to ORF1, ORF2, ORF3, ORF4 and ORF5 generated using a Prime-a-Gene kit (Promega). Hybridisation was performed under standard conditions and hybridisation signals detected by autoradiography. The results observed from each probe were identical; a major genome sized transcript (4500bp) and an intense smear that seemed to be masking several minor transcripts. The intensity of the smear and the transcripts varied slightly from sample to sample.

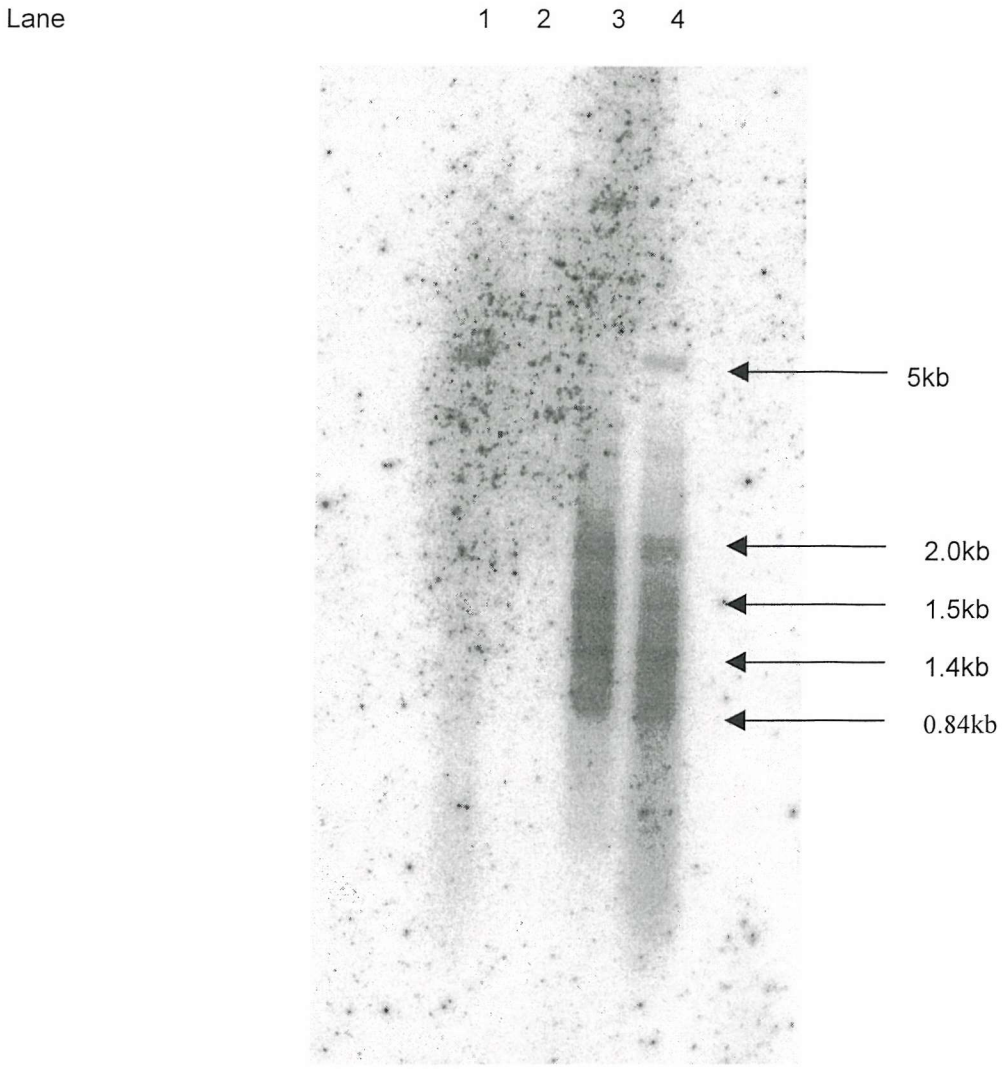



**Figure 5.3** An auto-radiograph of a Northern blot containing RNA from Chp2 infected *C. abortus* probed with $^{32}P$ randomly labelled Chp2 DNA. **Lane 1.** 0 hours post infection (p. i); **Lane 2**. 24 hours p.i; **Lane 3**. 48 hours p.i; **Lane 4**. 66 hours p.i.

Initially it was assumed that Chp2 produced a single genome-sized transcript, plus several minor transcripts. However, as it was predicted prior to the investigation that the detection level of Chp2 transcripts should be low, because the yield of Chp2 infected chlamydiae is low, the extracted RNA was subjected to further treatments prior to electrophoresis. To determine if the major transcript was RNA or contaminating ssDNA genome.

10µl of the extracted RNA sample was subjected to a further DNase I incubation step to eliminate any Chp2 ssDNA carry over, and 10µl was treated with NaOH to hydrolyse all of the RNA in the sample. A sample of Chp2 single-stranded DNA was also included to ascertain how ssDNA migrated on a denaturing formalydehyde gel.

These three samples plus untreated RNA were run on a formalydehyde gel as described in section (2.17.1) then blotted onto a nylon membrane as before. Hybridisation was then carried out with a $^{32}P$ Radiolabelled probe of ORF5. The results from the hybridisation were analysed by autoradiography (figure 5.4).

The major genome sized 'transcript' and the intense smear were present in lane 2 (which corresponds to alkaline treated RNA). Indicating that the major 'transcript' was not RNA. This was confirmed by the fact that the major genome sized 'transcript' was not observed in the lane that corresponded to RNA that had undergone a second DNase I treatment. These results showed that the major product observed in the previous Northern hybridisations is Chp2 ssDNA. This was further verified by the fact that Chp2 ssDNA ran at the same size on the denaturing gel (figure 5.4).

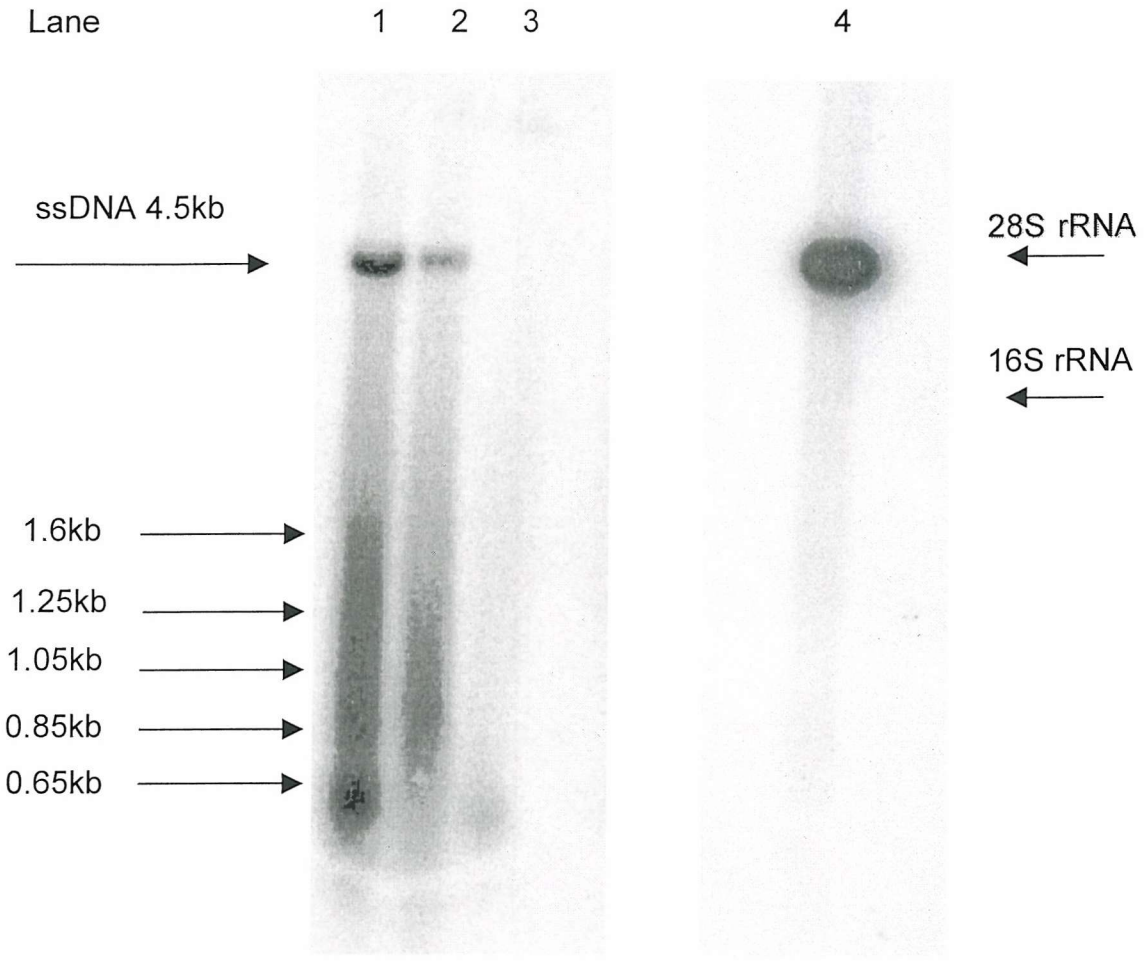



**Figure 5.4** An auto-radiograph of a Northern blot containing RNA from Chp2 infected *C.abortus* at 66 hours p.i probed with $^{32}P$ radio-labelled Chp2 ORF5 DNA. **Lane 1**. Untreated RNA extract; **Lane 2**. Alkaline treated RNA extract; **Lane 3.** DNase I treated RNA extract; **Lane 4**. Chp2 ssDNA.

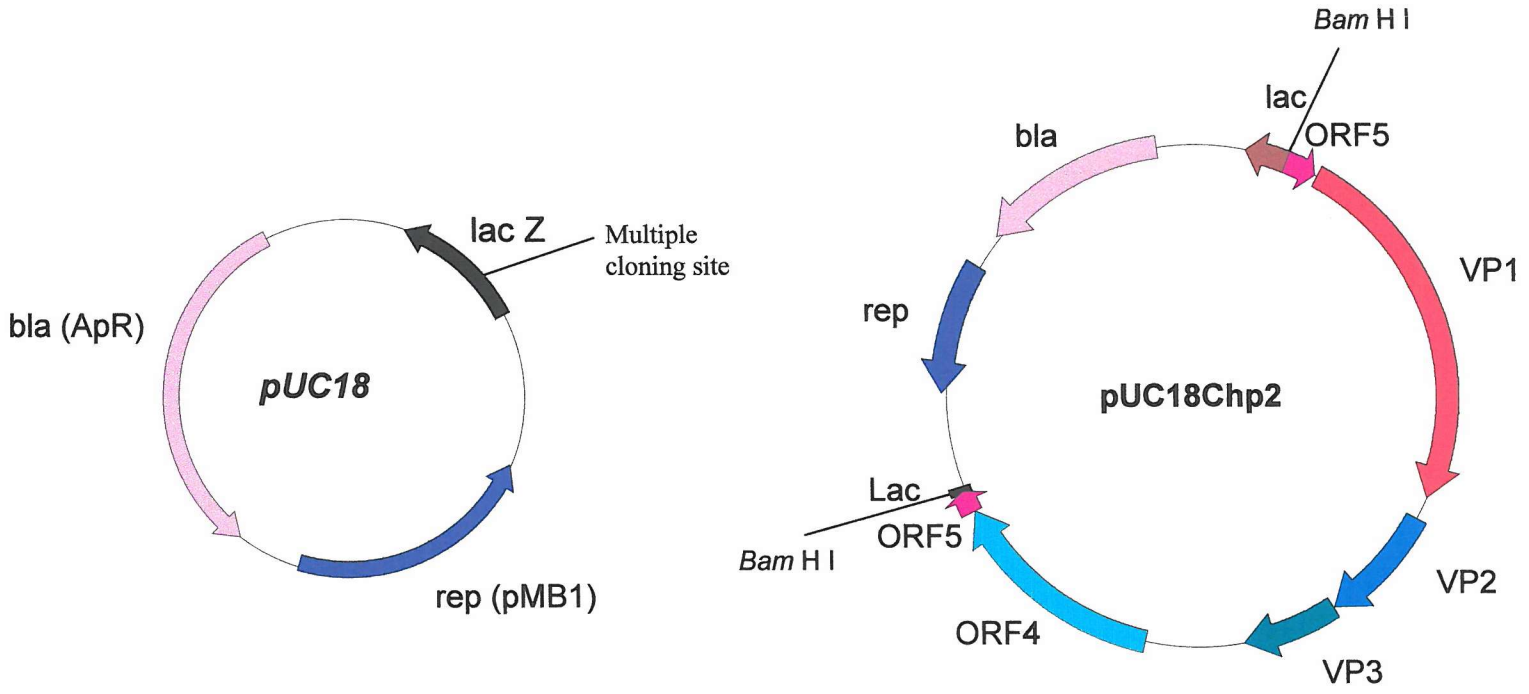
Faint 'transcripts' masked by the intense smear of Chp2 ssDNA, have also been observed during these investigations. There are two explanations for the presence of these 'transcripts': that they are actual Chp2 transcripts or that they are contaminating ssDNA that is running at different mobility due to supercoiling or secondary structures. Up to nine putative Chp2 'transcripts' were identified when a Northern blot was performed with $^{32}P$ ORF5 radiolabelled probe. However, the 'transcripts' were also observed after the alkaline lysis treatment of the extracted RNA and were not present after the second DNase I incubation step indicating that they are likely to be DNA. Bands can also be seen faintly on the autoradiograph in the lane corresponding to ssDNA. These observations indicate that the faint putative Chp2 'transcripts'are artefacts.

## 5.3 Bacteriophage promoters

As the Chp2 transcripts are in low abundance and could not be accurately defined, an experiment was devised to identify Chp2 promoters *in vitro*. Previous work has shown that SpV4 P1 and P0 promoters were active in *E. coli*, whilst this was not surprising as there were significant similarities between the SpV4 promoters and the *E. coli* consensus, it seemed possible that Chp2 promoters might also be functional in *E. coli*.

To determine if any of the Chp2 promoters were functional in *E. coli*, a coupled commercially available S30 *in vitro* transcription/translation system was used (2.14.2). The expression of genes in this system requires that the sequence contain a good *E.coli* promoter e.g. lacUV5, tac, trc lambda $P_R$ and Lambda $P_L$, or a T7 promoter.

For the initial investigation pUC18 vector containing the full length Chp2 genome ligated into the multiple cloning region via a *Bam* HI site situated in the Chp2 ORF5 was used as a template in the transcription/translation reaction (figure 5.5). The pUC18 plasmid contains: a bla gene, that codes for beta-lactamase (which is responsible for ampicillin resistance); a lac Z gene that encodes the N-terminal fragment of beta-galactosidase; a CAP protein binding site; a lac repressor binding site and the promoter Plac, all from regions of the *E.coli* genome (figure 5.5).




**i) pUC18 vector** **ii) pUC18Chp2**

**Figure 5.5** Gene organisation of the vector pUC18 and pUC18Chp2.

This clone and pUC18 vector alone were used as templates in the *in vitro* transciption/translation *E.coli* S30 system. It was hypothesised that any extra proteins expressed in the reaction containing pUC18Chp2 would be due to the expression of a Chp2 protein and may indicate Chp2 promoter activity.

### 5.3.1 *in vitro* transcription/translation system

The plasmid pUC18 and pUC18Chp2 were prepared by mini prep (Qiagen) and the sizes were verified and quantified by agarose gel electrophoresis (2.5.2). The plasmids were used as templates (500ng DNA) in S30 reactions as described in section 2.14.2. The results from this investigation were then analysed by SDS-PAGE followed by autoradiography as described in section (2.16). The results (figure 5.6) showed an extra translation product in the reaction for pUC18 containing the full-length Chp2 genome. The extra translation product was approximately 21kDa in size. Two of the bacteriophage proteins are estimated to be around this size, VP2 and VP3, which are 20.6kDa and 16.4kDa in size respectively.

There are three possible explanations for the presence of this extra translation product. The most likely is that it is due to the plasmid lacZ promoter transcribing the bacteriophage genes. This is more likely for genes situated at the beginning of the genome closest to the promoter i.e. ORF1. Another explanation is that one of the bacteriophage promoters is active in the *E.coli* system and is either transcribing a bacteriophage gene or transcribing the end of lac Z. This explanation would be more likely for genes situated in the middle/end of the bacteriophage genome, which are situated further downstream from the promoter, i.e. ORF2/ ORF3. As ORF2 is situated after a small non-coding region in the bacteriophage genome it was hypothesized that the extra translation product was VP2.

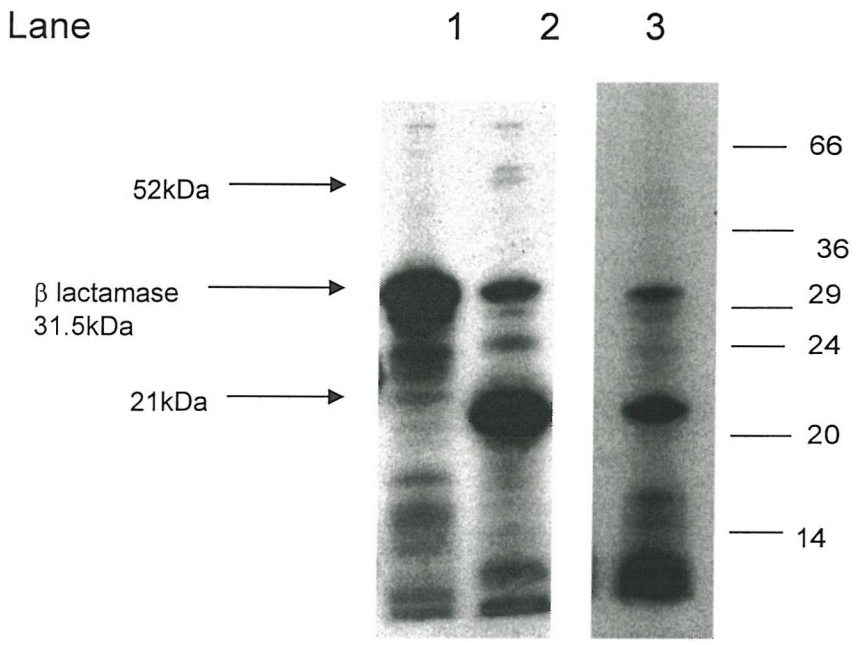


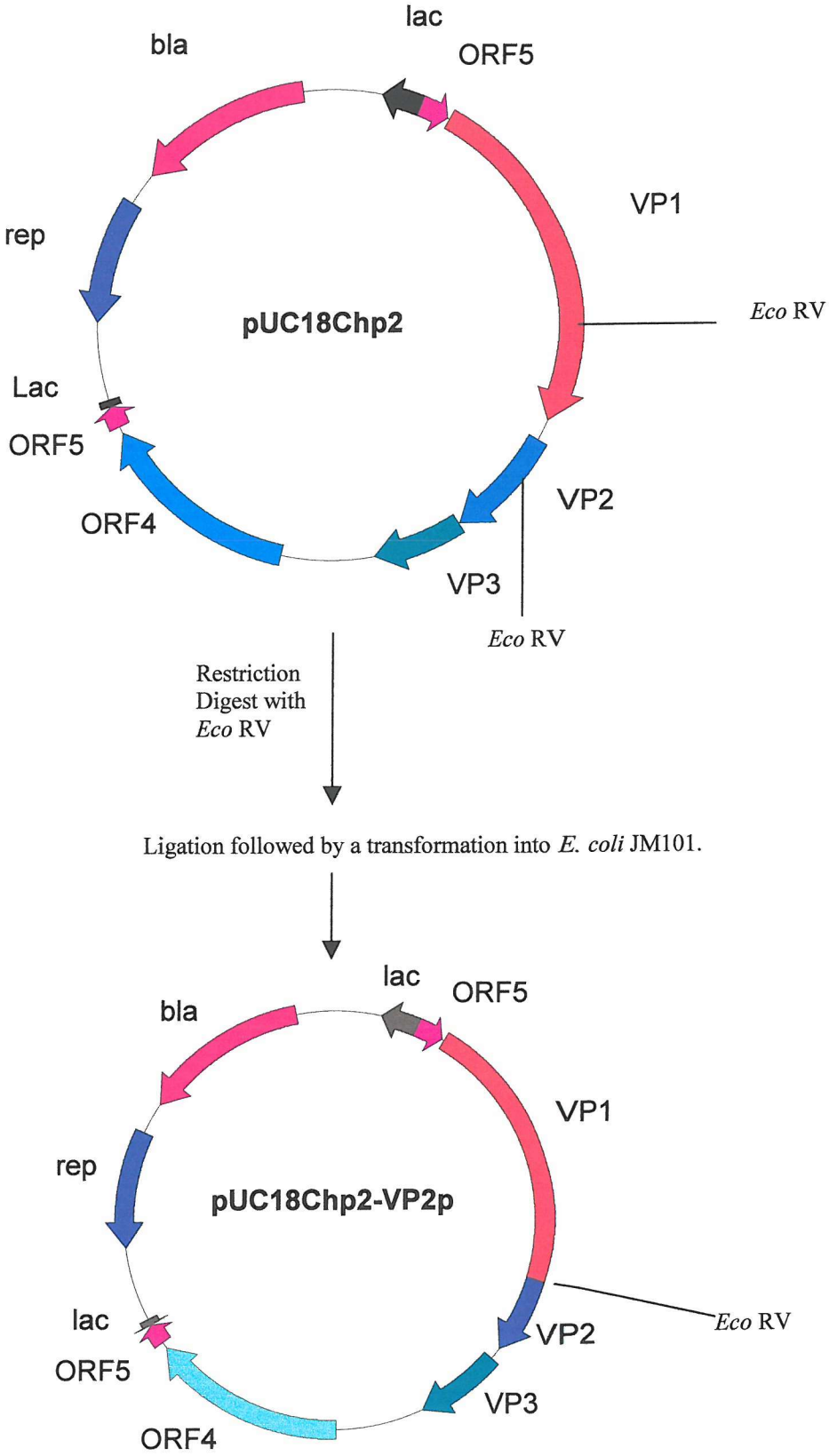

**Figure 5.6** An auto-radiograph of a 12.5% SDS PAGE gel containing the products of a S30 *in vitro* transcription/translation. The molecular weight markers are indicated. **Lane 1**. pUC18; **Lane 2**. pUC18Chp2; **Lane 3**. pUC18Chp2-p2.

### 5.3.2 Removal of the putative ORF2 promoter region.

The aim of this experiment was to remove a fragment of DNA upstream of ORF2 in the pUC18Chp2 construct. This was carried out to determine if the extra translation product observed in the S30 extract system was due to an active promoter situated upstream of the ORF2 translational initiation codon.

The plasmid pUC18Chp2 was prepared by mini prep (Qiagen) and a restriction digestion was carried out on the plasmid with *Eco* RV. This enzyme specifically cuts the Chp2 genome in two regions, once in ORF1 at position 1423bp and once in ORF2 at position 1903bp. It does not cut the pUC18 vector. The digestion removed a 480bp fragment of DNA from the construct consisting of the end of ORF1, the non-coding region between ORF1 and ORF2 and the first part of ORF2. This fragment contains the ORF2 start codon (ATG) plus the region upstream that could contain a Chp2 promoter. After the restriction digest the plasmid was analysed by gel electrophoresis. Two DNA bands were observed on the agarose gel, a band of approximately 500bp (the end of ORF1 to the first half of ORF2) and a band 6.7 kb in size (the plasmid minus 500bp). The DNA band that corresponded to the 6.7kb plasmid pUC18Chp2 minus the ORF2promoter region (pUC18Chp2-p2) was extracted from the gel using a surgical blade and purified using a QIAquick gel extraction kit (2.7.3.2). The plasmid vector pUC18Chp2-p2 was then re-ligated (2.6.1.3), and transformed into *E. coli* DH5α competent cells (2.3.3). Clones containing the plasmid were analysed by STET (2.4.2.3) to determine that the plasmid insert was the correct size followed by sequence analysis as described in section 2.8 using the primers VP1F (ATTTCTTCACTTGGATCCATGGTTAGGA$^{10}$) and VP2R (ATCACGAAGCTTCTAAGGATTATTCGGAGACCA$^{2083}$), which confirmed the 500bp deletion (figure 5.7).

**Figure 5.7** Construction of pUC18Chp2-VP2



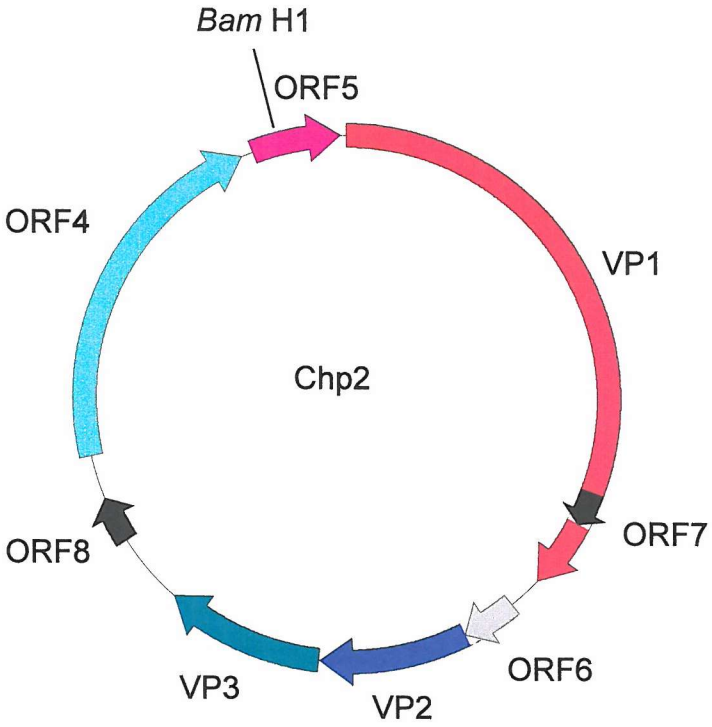
### 5.3.3 S30 *in vitro* transcription/translation of pUC18-Chp2 -500bp

The plasmid DNA of pUC18Chp2-p2 was isolated using a mini prep kit (2.4.2.1) and the size verified by gel electrophoresis. The plasmid templates pUC18, pUC18Chp2, and pUC18Chp2-p2 were then used as templates in S30 reactions (section 2.14.2), analysed by SDS-PAGE followed by autoradiography as described in section (2.12, 2.16). The results for pUC18-Chp2 and pUC18-Chp2-p2 were identical, indicating that the extra translation product observed in the autoradiograph for pUC18Chp2 was not due to an active promoter situated upstream of the ORF2 translational initiation codon (figure 5.6).

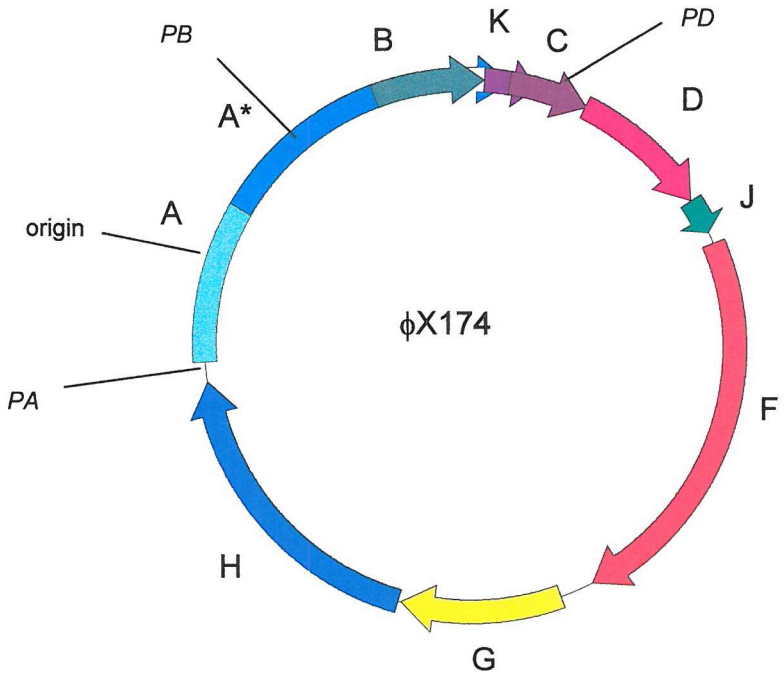
### 5.3.4 Identification of bacteriophage promoters active in E. coli

Based on the observation in section 5.2.3, that a Chp2 promoter may be active in an *E. coli in vitro* transcription/translation system, it was assumed that the promoter would also be active *in vivo* in *E.coli* cells. To determine if Chp2 promoters are active in *E.coli* cells the cloning vectors pNM480/81/82 were utilised. These vectors have been constructed for the evaluation of DNA sequences with potential promoter activity. The vectors allow fragments of DNA with suspected promoter activity to be fused to β-galactosidase. There are three vectors that allow the fragments to be cloned in all three reading frames. Due to the high copy number of the plasmid an increased sensitivity in detection of genes with promoter activity is seen. Fragments of DNA which have promoter activity can be detected by the ability of the cells containing the plasmid to hydrolyse BCIG to a dark blue precipitate resulting in the colonies looking blue (Minton.N.P, 1984).

Potential promoter regions in the Chp2 genome were identified by comparing the genome organisation of Chp2 to ϕX174 (figure 5.8) and SpV4, then identifying regions which contain promoters in ϕX174 and SpV4 and the corresponding region in the Chp2 genome. Two regions of the Chp2 genome were chosen that could contain promoters, the region upstream of the ORF5 translation-initiation codon that is situated in a region that corresponds to $P_D$ in ϕX174 and $P_2$ in SpV4. The region




i) Chp2




ii) ϕX174

**Figure 5.8** A diagrammatic representation of the Chp2 and ϕX174 genome organisation.

upstream of the ORF4 translation-initiation codon was also chosen as it corresponds to $P_A$ in the φX174 and P3 in SpV4. Primers were then designed to amplify a region approximately 200 bp long upstream of the ORF4 (figure 5.9) and ORF5 (figure 5.10) translation-initiation codons. The regions were then ligated in frame with the β-galactosidase gene into pNM480/1/2.

The recombinant plasmids were transformed into *E. coli* JM109 cells and plated out onto Xgal plates as described in section 2.2.9.

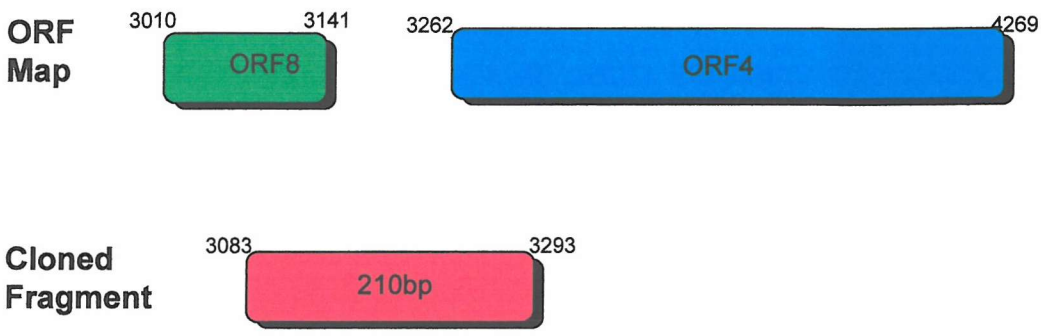
The plasmid clones containing the region upstream of the ORF5 translation-initiation codon gained the ability to hydrolyse BCIG resulting in blue colonies on Xgal plates. This result indicates that a bacteriophage transcription-initiation signal recognised by the *E. coli* RNA polymerase is present on the DNA molecule in this region ($P_5$). The clone containing the region upstream of the ORF4 translation-initiation codon gave paler blue colonies suggesting that a weak *E. coli* transcription-initiation signal is present ($P_4$). A vector only transformation was also carried out as a negative control. This gave white colonies when plated out onto Xgal plates.

Positive clones were screened by STET followed by sequence analysis using primers ORF5F/ORF5R and ORF4F/ORF4R to verify that the inserted sequences were correct (2.8).

The two regions inserted into the promoter cloning vectors were both approximately 200bp in size. To find the minimal functional component of $P_5$ and investigate if the upstream flanking DNA sequence had an effect on the activity of the promoter, an experiment was designed to reduce the size of the inserted fragment.

Two further forward primers were synthesised to amplify fragments of DNA 134bp and 67bp in size upstream of the ORF5 translation-initiation codon. These fragments of DNA were then ligated in-frame with the β-galactosidase gene in pNM480. The recombinant plasmids were transformed into *E. coli* JM109 cells and plated out onto Xgal plates as described in section 2.2.9.

**Figure 5.9** A diagrammatic representation of the Chp2 putative promoter region upstream of ORF4.




**Fragment sequence**

*Bam* H1

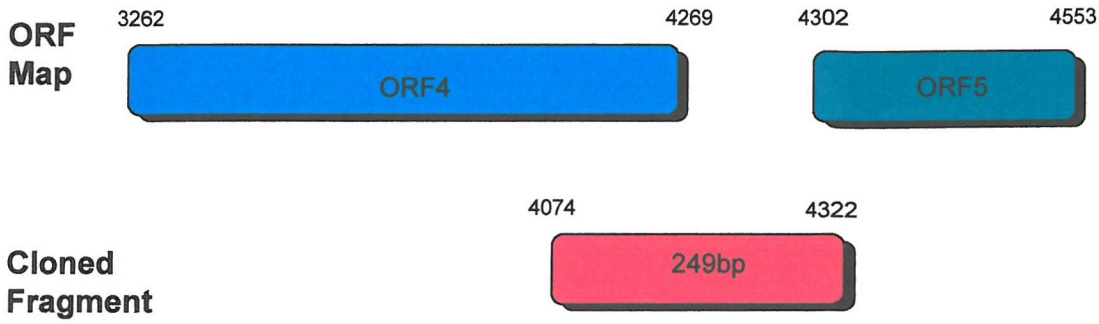
**CATTGAGGAT CCATCCAAGA AATAGGC**TTC GAAGAATTAT GCGTGGCGGC ATTAGGTTCT 3143
AGTTTTGGAC GTTAAGGAAA TCTTTAAGGT TATGCTAAAT TAGCTGCTAT GTATAATTTG 3203
GCTCGTGACG AATGTCATAT TCGCACCATT TAAATTGTAC ACAGCAGTTG AAGGCTTAGA 3263
**TGTTGATTTT TAATGTCGAC GCCTTCATTT TT**

*Sal* I

Translational initiation codon

Primers are shown in bold.
Restriction sites are indicated in blue.

**Figure 5.10** A diagrammatic representation of the Chp2 putative promoter region situated upstream of ORF5.




**Fragment sequence**

*Bam* H1

**TAAGACGGAT CCTCCACG**TT ACTATGATAA GTTACATTCT CGGTTTGATC CGGAAGAGAT 4134
GGACGAGGTC AAACAAAAAC GTATAGAGAA AGTTATGGCT TTGCCTGAGC TAACTCAGGA 4194
TAAGGCTGAG GTGAAGCAAT ATATTTTCAA TGACCGTACG AAGAGACTCT TTAGAGACTA 4254
TGAGGAGGAG AGTTACTAAA CTTTTTTAAA AAATAGGAGC TTTTTTCA**AT GAAAGCTTTT** 4314
**ACAGTGTTT**

*Hind* III

Translational initiation codon

Primers are shown in bold.
Restriction sites are indicated in blue.

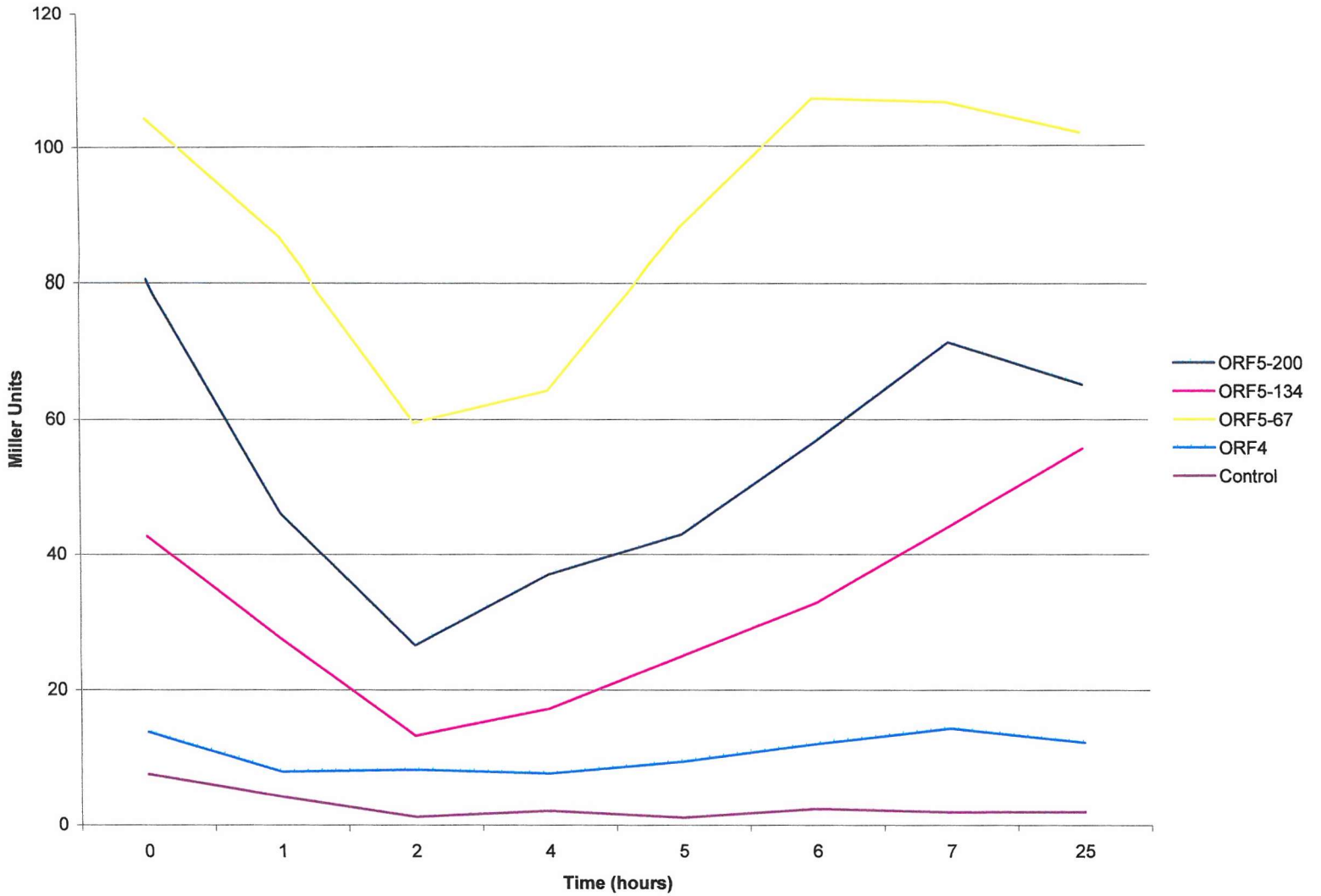
The colonies containing the plasmid with the 134bp insert and the colonies containing the plasmid with the 67bp insert both gained the ability to hydrolyse BCIG into a dark blue precipitate. This indicates that $P_5$ does not require the upstream region for activity in *E. coli*, however, the flanking region may effect promoter activity.

### 5.3.5 β-galactosidase activity of promoter clones

To determine the relative strengths of each promoter fragment the β- galactosidase enzyme activities of the *E. coli* JM109 cells containing the $P_4$ and $P_5$ promoter clones were measured. The enzyme assays were performed on intact cells, using the method described in section 2.13. The enzyme activity was expressed as units per ml of culture (U/ml) (Miller, 1972). Figure 5.11 shows the results from this assay. The 67bp $P_5$ promoter clone was the most effective promoter showing a peak β-galactosidase activity of 101.9 U/ml. The $P_5$ 200bp promoter clone gave a peak enzyme activity of 65.1 U/ml and the $P_5$ 134bp promoter clone gave a peak enzyme activity of 55.7 U/ml. The $P_4$ promoter clone was the least effective promoter showing a peak enzyme activity of 12.2U/ml. The vector only control showed a 1.95 U/ml peak enzyme activity. The results indicated that the 5' flanking sequence of the $P_5$ promoter influences the overall effectiveness of the promoter in *E. coli.*

The enzyme activities start off high then drop in activity before rising to the initial levels again. The initial drop in enzyme activity is probably due to the bacteria being in their Lag growth phase and adapting to the new media.

**Figure 5.11** A graph to show β-Galactosidase enzyme activities of *E. coli* JM109 containing the Chp2 ORF4 and ORF5 promoter regions cloned into plasmids pNM482and pNM480 respectively. Enzyme activities are expressed as units per ml of culture (U/ml) (Miller 1972).



## 5.4 Construction of clone P15CATChp2.

The aim of this investigation was to determine if the $P_5$ promoter would allow the transcription and consequently translation of ORF5 and the subsequent Chp2 ORFs (ORF, ORF2, ORF3) *in vitro*. Resulting in the production of ORF5 protein, VP1, VP2 and VP3.

The full-length Chp2 genome clone utilized in the *in vitro* experiment in section 5.3 (pUC18Chp2) contains the Chp2 genome ligated into pUC18 via a *Bam* HI site in ORF5. The $P_5$ promoter may be responsible for the extra translation product observed during the *in vitro* transcription/translation studies in section 5.3. This translation product could consist of the end of ORF5 protein up to a fortuitous translational stop codon situated in pUC18. A predicted ORF is present in this region when an ORF map of the nucleotide sequence of pUC18Chp2 is generated (MapDraw, DNASTAR.). This ORF consists of the end of ORF5 up to a translational stop codon situated at position nt2798 in pUC18 between lacZ and rep(pMB1) generating a 369nt ORF. However, the protein encoded by this ORF is predicted to be smaller (13kDa) than the extra translation product produced by pUC18Chp2 (21kDa), and the 21kDa product cannot be detected in a RIPA assay using polyclonal antisera to ORF5 protein.

It is anticipated that when the new clone is used as template in an *in vitro* S30 transcription/translation, the *E. coli* RNA polymerase will bind to $P_5$ then transcribe the whole of ORF5 then go on to transcribe the subsequent cistrons e.g. ORF1, ORF2 and ORF3. This polycistronic mRNA could then be translated by the *E. coli* translational machinery present in the S30 system, resulting in the production of ORF5 protein, VP1, VP2 and VP3.

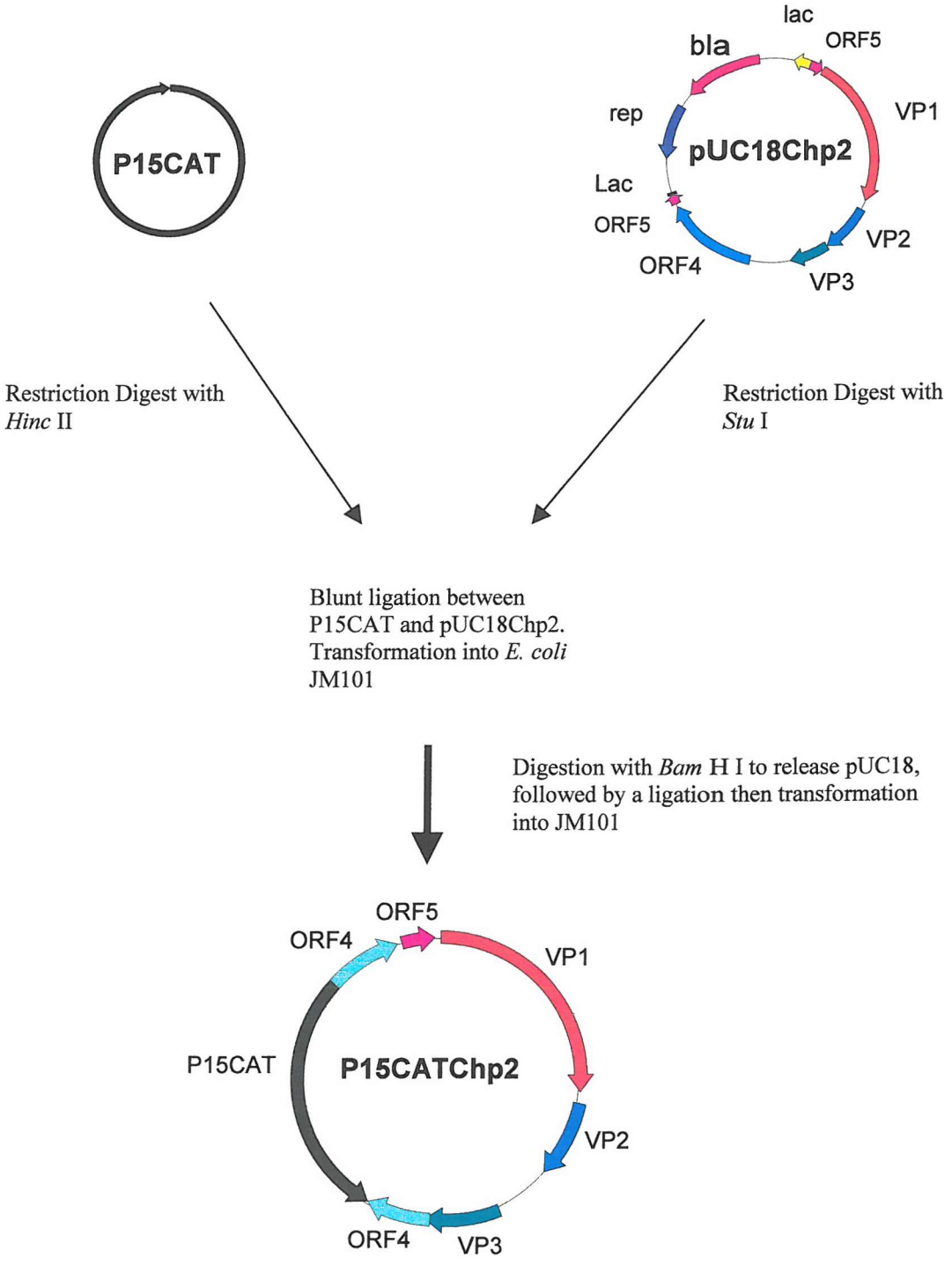
To change the vector location in pUC18Chp2 from ORF5 to ORF4, an in house vector P15CAT vector was utilised. This vector was chosen because it does not contain a *Bam* HI restriction endonucleace cleavage site. The vector P15CAT contains an *E. coli* P15 origin of replication and a CAT gene.

The in house vector was first digested with *Hinc* II then ligated into the pUC18Chp2 construct via a *Stu* I site located in the ORF4 (figure 5.12). The construct was transformed into *E. coli* JM101 cells and positive clones were identified, by selecting for both ampicillin and choramphenicol resistance. The transformation results were analysed by STETS followed by gel electrophoresis and positive colonies were confirmed by restriction digest analysis. A clone that contained pUC18Chp2-P15CAT, was then digested with *Bam* HI to release the pUC18 vector. The digest yielded two products a 2686bp product corresponding to the vector pUC18 and a 6376bp DNA product corresponding to the p15CATChp2 construct. The 6376bp DNA band was then extracted from the gel using a surgical blade and purified using a QIAquick gel extraction kit (2.7.3.2). The construct was next re-ligated and transformed into *E. coli* JM101. The results from the transformation were examined by STETs followed by restriction digest and sequence analysis. To determine the orientation of the P15CAT vector and verify the sequence of the junctions between P15CAT and the Chp2 genome.

### 5.4.1 S30 *in vitro* transcription/translation of p15CATChp2

Two plasmids p15CAT and p15CAT containing the full-length Chp2 genome, were then used as templates in the *E. coli* S30 *in vitro* transcription/translation system (2.14.2), and the results analysed by SDS PAGE followed by auto-radiography (2.16). In the autoradiograph, four extra translation products were observed in the lane that corresponded to p15CAT containing the full-length Chp2 genome. One at 66kDa which could only be VP1 (63.5kDa), the second at 24kDa which could be VP2 (20.2kDa) the third which was 20kDa and could correspond to VP3 (16.8kDa) and a fourth product 8.5kDa in size which is hypothesized to be ORF5 protein (9.5kDa) (figure 5.13). The 27kDa product, which corresponds to CAT, is present in both P15CATChp2 and the vector only control.

**Figure 5.12** Construction of P15CATChp2.



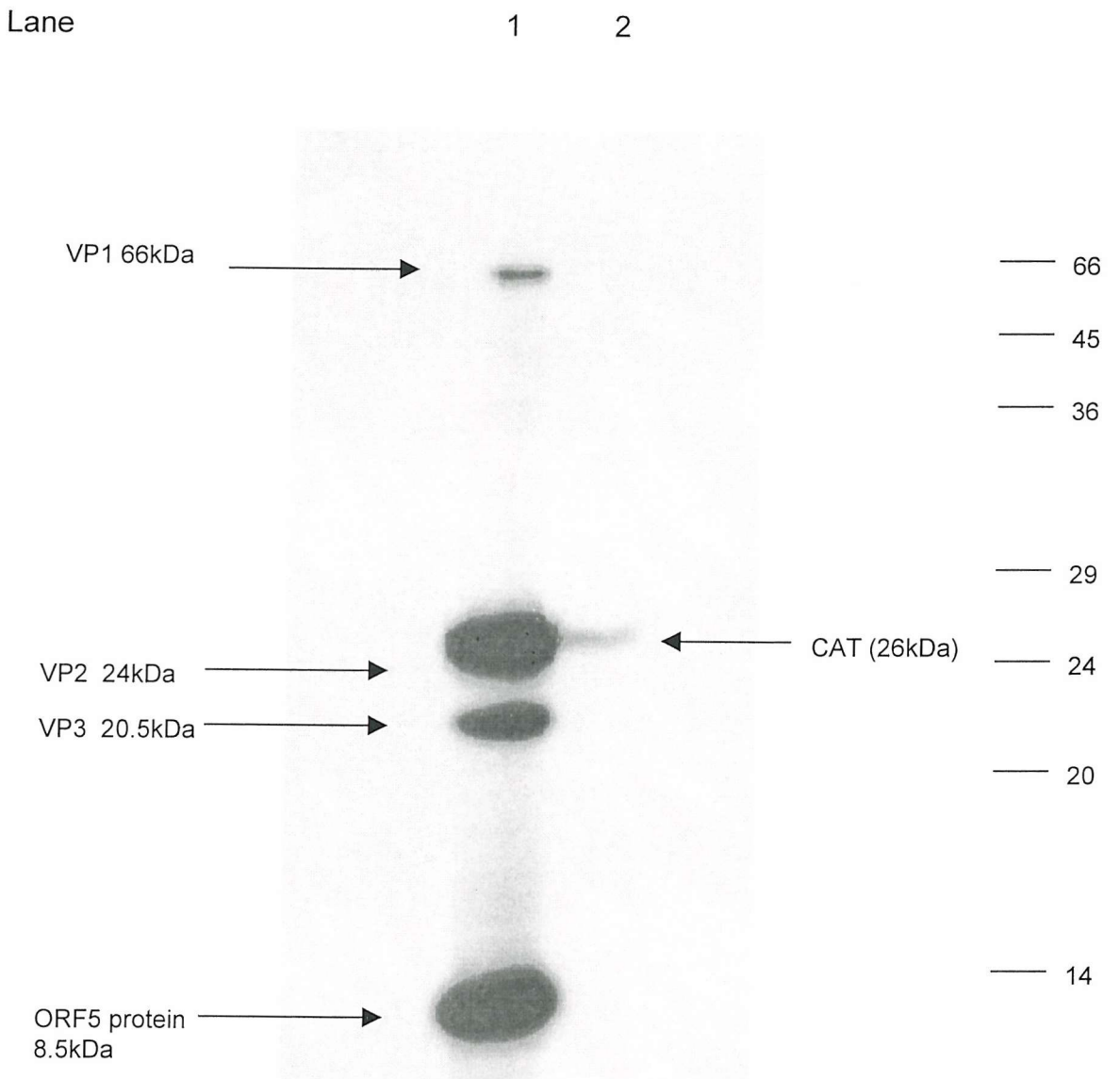



**Figure 5.13** An autoradiograph of a 12.5% SDS PAGE gel showing the products of S30 *in vitro* transcription/translation. The molecular weight markers are indicated. **Lane 1**. P15CATChp2; **Lane 2.** P15CAT.

### 5.4.2 Radioimmune precipitation assay (RIPA)

To identify the extra translation products observed in section 5.4.1, the radioactive translation products were analysed by using a Radio Immune Precipitation Assay using mAb 55 to detect VP1 and polyclonal antisera specific to, VP3, ORF4 protein and ORF5 protein, as described in section (2.15). The immunoprecipitation products were then analysed by SDS PAGE (2.12.1), and autoradiography (2.16). The antisera specific to ORF5 protein reacted with the 8.5kDa translation product, and the antisera specific to VP3 reacted to the 20kDa product. No antisera reacted to the 66kDa product or the 20.2kDa product. As the amount of the 66kDa product was low it is possible that there is not enough present for detection by this technique (figure 5.14). It is not surprising that the VP2 antisera did not react with the 20.2kDa product as this antisera did not show reactivity to VP2 in previous investigations (section 4.4).

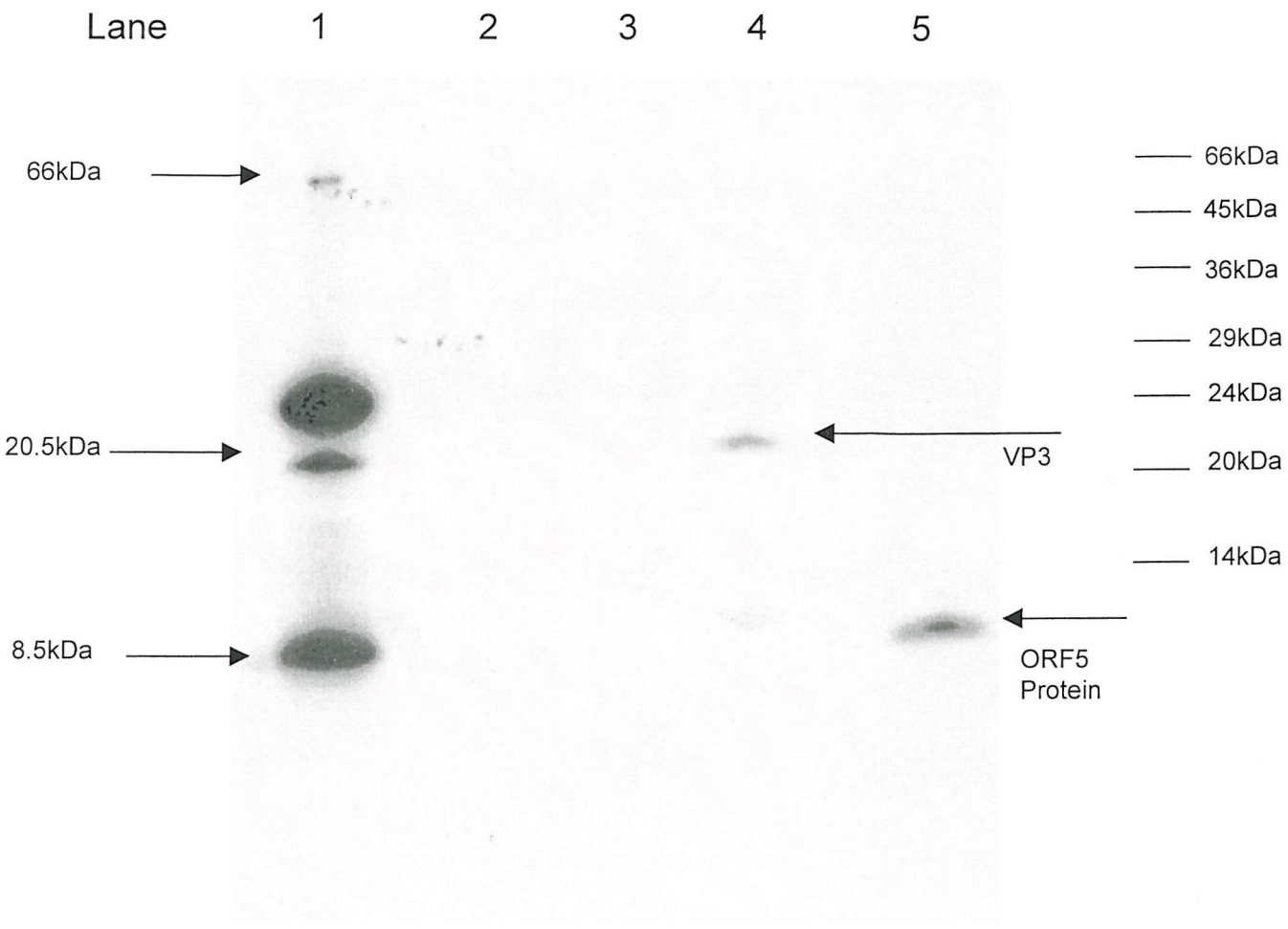


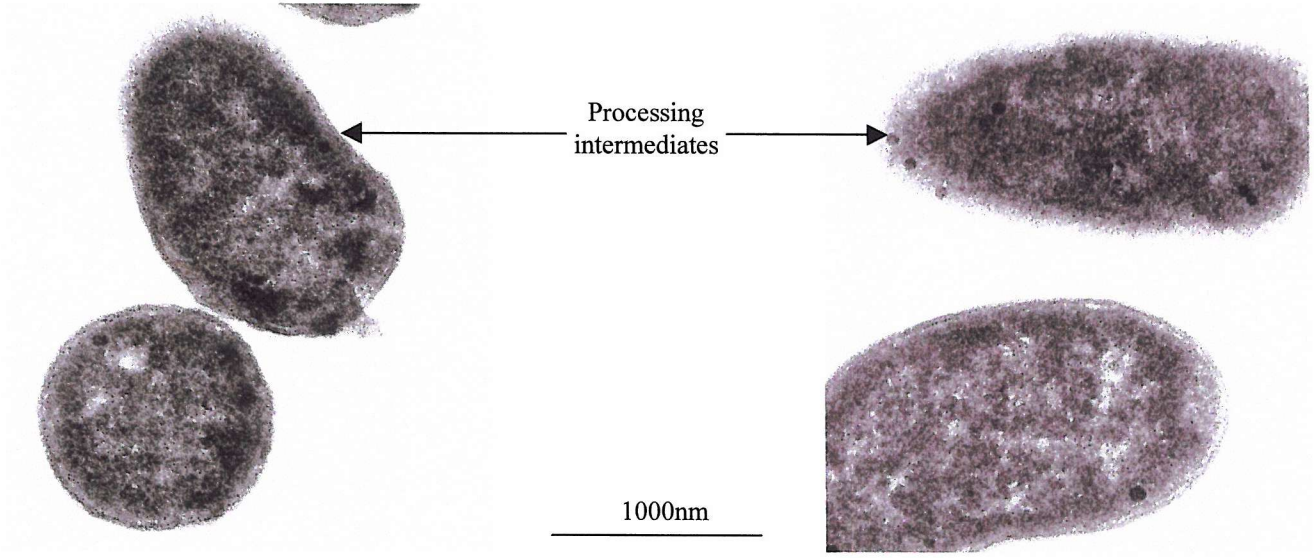

**Figure 5.14** An autoradiograph of a 12.5% SDS PAGE gel showing the results of a RIPA on P15CATChp2 *in vitro* transcription/translation products using antisera to VP1, VP2, VP3 and ORF5 protein. The molecular weight markers are indicated. **Lane 1**. *In vitro* transcription/translation products; **Lane 2**. VP1 monoclonal antibody; **Lane 3**. VP2 polyclonal antisera; **Lane 4**. VP3 protein polyclonal antisera; **Lane 5**. ORF5 protein polyclonal antisera.

## 5.5 Detection of Chp2 proteins *in vivo* in *E. coli*

The aim of this investigation was to determine if the $P_5$ promoter in the plasmid P15CATChp2 will function *in vivo* in *E. coli* cells and allow expression of the Chp2 proteins and subsequently the production of bacteriophage particles.

The construct containing the full length Chp2 genome ligated in the vector P15CAT via ORF4, plus a vector only control were each transformed into *E. coli* JM101 cells. The cells were then grown overnight at 37°C and a 1ml sample was taken out of each culture and analysed by negative stain microscopy to identify any bacteriophage particles (figure 5.15). No bacteriophage particles could be identified in the culture containing P15CATChp2.

The two cultures were also analysed by SDS PAGE (2.12.1) to determine if any of the Chp2 proteins are expressed. No extra protein bands were observed for the lane that corresponded to the vector containing the Chp2 full-length genome (figure 5.16). However, this could be due to a low translation rate of these proteins and consequently obscurity by the *E. coli* proteins that were present in large amounts. To determine if the proteins were present in small quantities the proteins that had been separated by SDS PAGE were blotted on to a nitrocellulose membrane (2.12.4). An immunoblot was then performed on strips of the membrane with the antisera specific to VP3 and ORF5 protein and a polyclonal antisera to Chp2 particles. However, no extra protein bands were detected in the lane that corresponded to the construct containing the Chp2 full-length genome (figure 5.17/5.18).




*E. coli* JM101 containing the plasmid P15Cat Chp2.

*E. coli* JM101 containing the plasmid P15Cat.

**Figure 5.15** Negative stain electron micrograph of *E. coli* JM101 containing the plasmid P15CatChp2 and P15Cat.

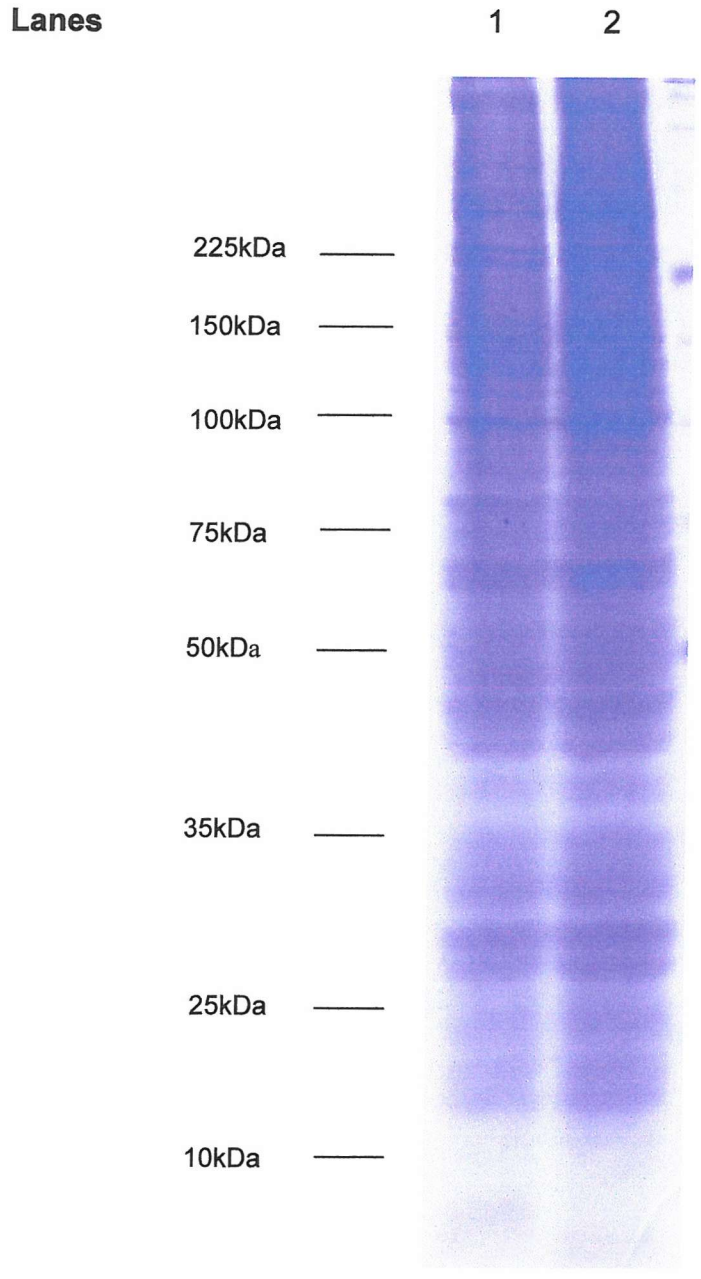



**Figure 5.16** A 12.5% SDS PAGE gel showing the proteins present in *E. coli* JM101 containing either the plasmid P15CatChp2 or P15Cat. **Lane 1**. *E. coli* JM101 containing the plasmid P15CatChp2; **Lane 2**. *E. coli* JM101 containing the plasmid P15Cat.

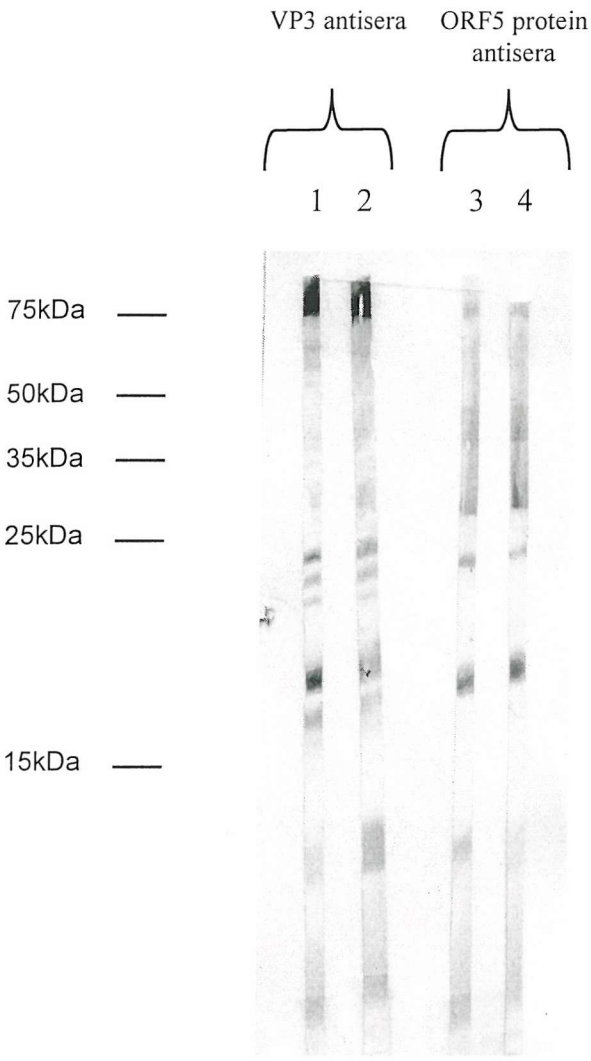



**Figure 5.17** Hyperimmune immunoblots with rat polyclonal antisera (1/1000) to VP3 and ORF5 protein. The antisera were blotted against the proteins present in *E.coli* JM101 containing the plasmid P15CatChp2 and P15Cat run on a 12.5% SDS-PAGE gel. **Lane 1**. *E. coli* JM101 containing P15CatChp2; **Lane 2**. *E. coli* JM101 containing P15Cat; **Lane 3.** *E. coli* JM101 containing P15CatChp2; **Lane 4**. *E. coli* JM101 containing P15Cat. VP3 is predicted to be 16.8kDa and ORF5 protein is predicted to be 9.5kDa.

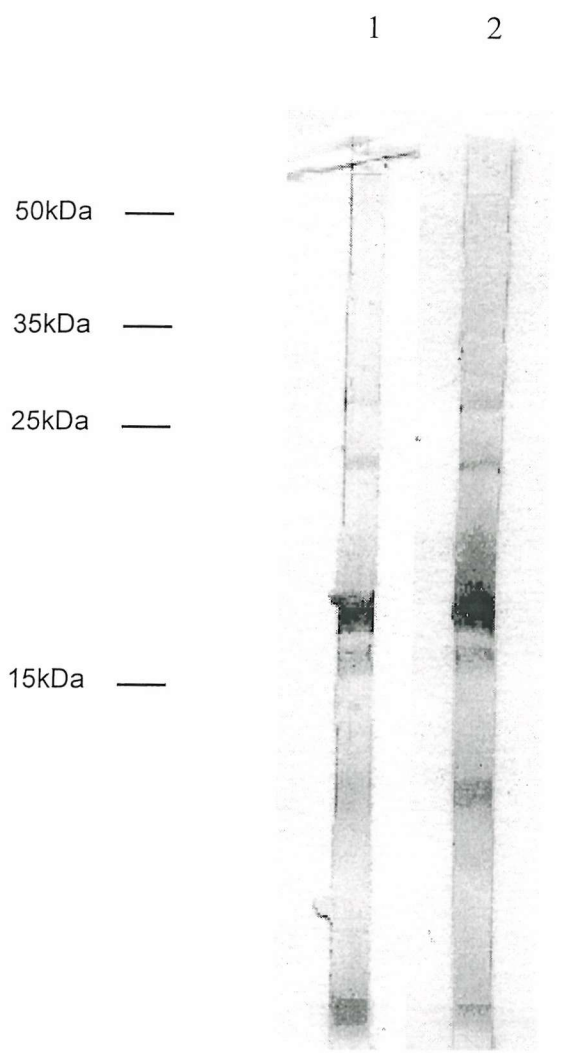



**Figure 5.18** Hyperimmune immunoblots with rabbit polyclonal antisera (1/1000) to Chp2 particles. The antisera was blotted against the proteins present in *E. coli* JM101 containing P15CatChp2 and P15Cat run on a 12.5% SDS PAGE gel **Lane 1**. P15Cat; **Lane 2**. P15Cat Chp2.

## 5.6 The effect of phage infection on Chlamydiae gene expression

The aim of this experiment was to determine if Chp2 infection causes inhibition of host chlamydial rRNA transcription during replication. In *E.coli*, ϕX174 infection does not cause host cell shut down, however, host transcription is slowed down during bacteriophage infection. This investigation involved infecting BGMK cell monolayers with *C. psittaci* B577 then extracting the total RNA at time points.

### 5.6.1 Infection of BGMK cells

BGMK cells were passaged into sixteen 25cm$^2$ tissue culture flasks, and grown until confluent. Eight of the flasks were then infected with *C. abortus* B577. The remaining eight flasks were infected with chlamydiae containing Chp2, by adding a standard inoculum of Chp2 (shown to be capable of infecting >99% of *C. abortus*) to *C. abortus* B577 prior to infection of BGMK cells.

### 5.6.2 RNA Extraction

The RNA was extracted from the cell monolayer infected with chlamydiae/chlamydiae and phage as described in section 2.4.3, at 12 hour time intervals over three days (0 hours, 12 hours, 24 hours etc.). The RNA was isolated using the SV RNA isolation system (2.4.3).

### 5.6.3 cDNA production

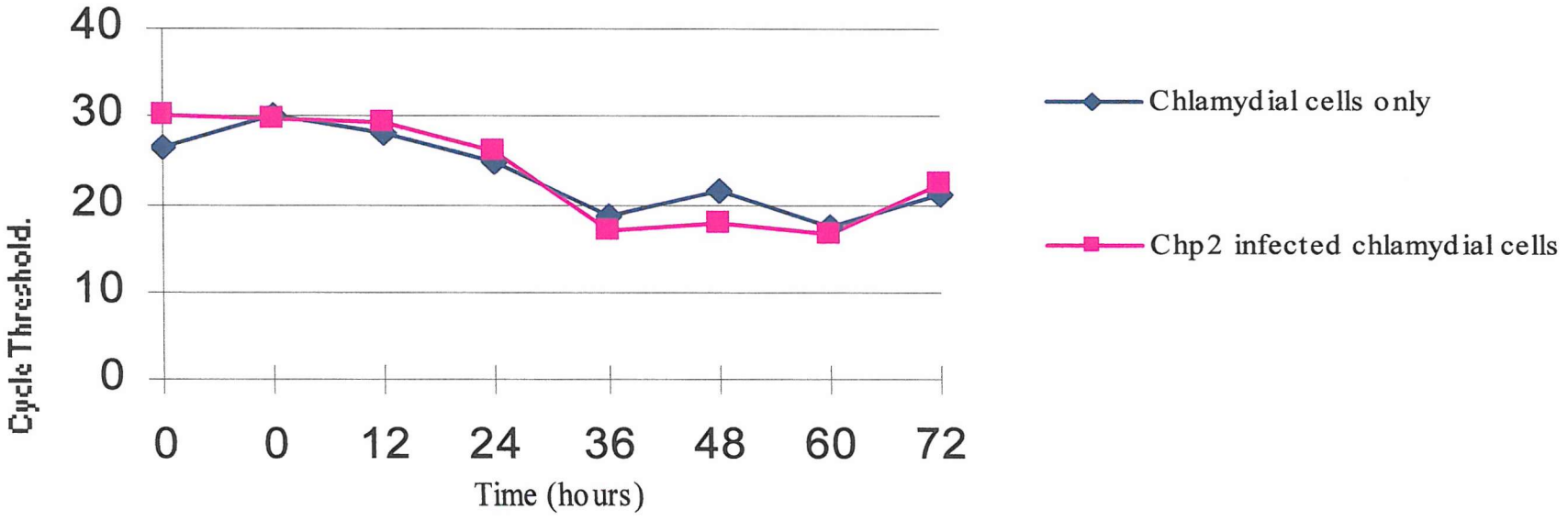
Immediately after the total RNA had been extracted from the cells, cDNA was prepared using a TaqMan$^R$ Gold RT-PCR Kit (Applied Biosystems) as described in section 2.7.4.1.

### 5.6.4 Assay of Chlamydial 16S rRNA

Real time PCR (2.7.4.2) was used to monitor the amount of chlamydiae 16S rRNA at each of the time points. There was no reference used to standardise the amount of rRNA extracted because if the phage causes host cell shut down a huge drop in RNA will occur.

The result showed that there was no significant difference between infected and uninfected chlamydiae cells, this indicates that the phage does not cause host cell 16S rRNA shutdown (figure 5.19).

**Figure 5.19** A graph to show the effect of Chp2 infection on chlamydial 16S rRNA production.



## 5.7 Discussion and Future Work

The attempts at identification of Chp2 transcripts by Northern blotting were not successful. This was due to the very low abundance of these molecules and the masking effect of Chp2 ssDNA or RF DNA that was not removed completely during the RNA extraction procedure. The SV RNA extraction kit includes a DNase I incubation step to remove any DNA from the sample, however, the vast amounts of ssDNA produced by the bacteriophage during stage II replication was not removed completely.

Two promoter regions were identified in the Chp2 genome that showed activity in *E. coli* cells. $P_5$ that is located upstream of ORF5 translation-initiator codon and $P_4$ which is situated upstream of the ORF4 translation-initiator codon. The relative overall effectiveness (i.e. strength) of each promoter region was determined by measuring the enzyme activities of the *E. coli* JM109 cells containing the $P_4$ and $P_5$ promoter clones. The $P_5$ promoter was investigated further by reducing its size to 67bp and 134bp. The 67bp $P_5$ promoter clone showed the highest enzyme activity, followed by the 200bp $P_5$ promoter clone then the 134 bp $P_5$ promoter clone. The $P_4$ promoter clone showed the lowest enzyme activity that was only slightly greater (12.2 U/ml) than the enzyme activity threshold (5 U/ml) (Miller, 1972). This data shows that the upstream flanking region of the $P_5$ promoter sequence has an effect on the overall strength of the promoter in *E. coli*. The sequence of DNA that the RNA polymerase interacts with (promoter region) largely determines the strength of the promoter. However, other effector molecules are also known to interact with the DNA template and influence (positively or negatively) the interaction between the promoter region and the RNA polymerase. The fact that the 67bp $P_5$ promoter clone showed a higher enzyme activity than the 200bp and 134bp $P_5$ promoter clones may indicate that a transcriptional inhibitor binds to the additional upstream region present in these clones. The difference in activities between the 200bp and 134bp promoter clones could be explained by the presence of a positive transcriptional regulator that binds to the DNA template in the additional 66bp present in the 200bp promoter clone. However, the variation in activity can only be speculated at this time due to differences that exist between the *E. coli* and chlamydial RNA polymerase and

promoter sequences, and that the positive or negative regulators active in *E. coli* may not be present in chlamydiae.

ORF4 proteins shows homology with the ϕX174 gene A nicking/closing protein. Protein A is one of the first proteins required in the ϕX174 replication cycle. It is needed in stage II replication for the production of copies of the bacteriophage genome. It is also required later in the replication cycle for stage III replication. It has been shown that gene A transcripts decay very quickly, this is thought to be a method of gene regulation to ensure that this protein is not over-expressed. The lower efficiency of the $P_4$ promoter could also be a mechanism of regulating the expression of this gene in chlamydiae.

Using an *in vitro* transcription/translation system and a radio immune precipitation assay (RIPA), it was shown that the $P_5$ promoter transcribes the subsequent ORFs, ORF1, ORF2 and ORF3 *in vitro*. This indicates that the $P_5$ promoter is the main Chp2 promoter, producing a transcript from ORF5 up to ORF3 or ORF8 (possibly not detected due to its small size). This is consistant with the location of a hairpin loop situated after the ORF8, which is probably the mRNA terminator.

ORF5 protein is thought to be to chlamydiaphage equivalent of the ϕX174 protein C based on a similar size, charge and location. This protein is a ssDNA binding protein that is required for initiation of stage III DNA synthesis. Gene C is situated in a similar genomic location in ϕX174 to ORF5 in Chp2 and is transcribed via $P_B$ situated upstream of gene B, which encodes the ϕX174 internal scaffolding protein.

In ϕX174 terminators are situated after genes J, F, G and H. The promoter $P_B$ produces four different length transcripts to each of the terminators. The amount of each transcript produced is dependent on the efficiency of the terminator events, which acts as a mechanism for gene regulation. Chp2 may utilise a similar type of gene regulation. Based on the location of the promoter $P_5$, transcripts containing ORF5 must be produced in the highest amount followed by ORF1, ORF2, ORF3 and ORF8. Similarly ϕX174, gene C mRNA is one of the most abundantly produced transcripts, followed by the gene F and H. Surprisingly, ORF3, which is thought to

encode a Chp2 scaffolding protein, is situated at the distal end of an operon. In contrast gene B which encodes the ϕX174 scaffolding protein is situated directly after $P_B$ and is one of the most abundant ϕX174 transcripts. Similarly ORF8, which is thought to encode the ϕX174 protein J analogue, is transcribed at the distal end of the transcript. In ϕX174 gene J is transcribed in the most abundant mRNA. This observation indicates that either due to the simpler structure of the Chp2 virion and genome, scaffolding proteins and DNA binding proteins are not required in great quantities, or that there is another Chp2 promoter located somewhere between $P_5$ and the translation-initiation codon of ORF3. This could be investigated further using the promoter-cloning vector used in section 5.2.6.

Chp2 proteins could not be detected *in vivo* in *E. coli* cells transformed with the plasmid P15CATChp2. This is probably due to the presence of proteases in *E. coli* cells. It is known that Chp2 does not cause shut off of host cell rRNA production from section 5.6. However, ϕX174 is known to inhibit host chromosomal DNA replication and division during replication. This mechanism may also be utilised by Chp2 for the reduction in the transcription and consequently translation of host cell proteases.

It has not been possible to locate the transcriptional initiation points for either of the Chp2 promoters. Investigations into the promoter regions P0 and P1 of SpV4, involved cloning potential promoter regions into an *E. coli* promoter selection vector (pKK 232-8) and then the mapping the transcriptional initiation start point in *E. coli* by primer extension. To determine the transcriptional start point of the promoters situated upstream of ORF5 and ORF4, primer extension studies could be carried out on the promoter fragment inserted into the pNM480/1/2 promoter-cloning vector. The region where the RNA polymerase binds to the promoter regions could be identified in *E. coli* using the $P_4$ and $P_5$ promoter clones. Foot-printing could then be used, using the *E. coli* polymerase as a DNA binding protein to identify the promoter regions.

The translation of Chp2 proteins could not be detected *in vivo* in *E. coli* cells, which, as mentioned previously, could possibly be due to *E. coli* nucleases or proteases. The Chp2 genes may still be transcribed in *E. coli* cells from the $P_5$ promoter region. This

could be determined by carrying out Northern blot hybridisations on *E. coli* cells grown with the plasmid using radiolabelled probed to the Chp2 major ORFs. If any transcripts are detected both the 5' initiation start point could be mapped using the method mentioned above, and the termination site. The terminator site is predicted to be located after ORF8.

# Chapter 6

## Final Discussion

Chlamydiae are obligate intracellular pathogens that cause a wide range of diseases in humans and animals. Due to their unique obligate intracellular developmental cycle progress into the study of chlamydial molecular biology has been slow compared to free-living bacteria. This is mainly due to the inaccessibility of the bacteria within the chlamydial phagosome and its dependence on host cells (Stephens, 1993). There is no current method for gene transfer in chlamydiae and it is difficult to isolate conditional lethal mutants. One solution to this problem would be the development of a cloning vector. Two of the most frequently used types of cloning vectors are plasmids and bacteriophages. Some chlamydial strains are known to possess plasmids, although studies leading to the development of a plasmid cloning vector have so far been unsuccessful. The discovery of bacteriophages that infect chlamydiae may provide the basis of a genetic transfer system. Therefore it is crucial to obtain a detailed knowledge about the molecular biology of these bacteriophages.

Five chlamydiaphages have been identified so far, infecting various hosts: Chp1 (*C. psittaci*), Chp2 (*C. abortus*), ϕCPAR39 (*C. pneumoniae*), ϕCPG1 (*C. caviae*) and Chp3, which was identified during these studies from a *C. pecorum* strain of chlamydiae (Richmond *et al.*, 1982a; Hsia *et al.*, 1996; Liu *et al.*, 2000; Read *et al.*, 2000a). Chp3 is closely related to the Chp2-like group of chlamydiaphages. The discovery of this bacteriophage confirms the hypothesis that infection of chlamydiae with these bacteriophages is widespread. The genomes of Chp2, Chp3, ϕCPAR39 and ϕCPG1 are extremely similar, whilst Chp1 shows high divergence throughout its genome.

The host range of Chp1 is thought to be restricted to *C. psittaci* strains (Richmond *et al.*, 1982b). Chp2 and Chp3 have identical host ranges (*C. abortus*, *C. caviae*, *C. felis* and *C. pecorum*) (Everson *et al.*, 2003; Everson *et al.*, 2002). ϕCPAR39 is also able to infect *C. abortus*, *C. caviae*, and *C. pecorum*, however, it is unable to infect *C. felis* and can additionally infect *C. pneumoniae* (Everson *et al.*, 2003). This suggests that Chp2/Chp3 and ϕCPAR39 have different host receptors and that both receptors are only present on the

hosts they share (*C. abortus*, *C. caviae* and *C. pecorum*). The host range of ϕCPG1 is not yet known. It is predicted that two discrete regions situated in VP1 (IN5 and Ins) encode the chlamydiaphage receptor-recognition domain in the Chp2-like bacteriophages (Chipman *et al.*, 1998; Read *et al.*, 2000b). These two regions are similar between Chp2/Chp3 and ϕCPG1/ϕCPAR39. This suggests that ϕCPG1 should have an identical host range to ϕCPAR39.

Chlamydiaphages belong to the *Microviridae* family of bacteriophages, which consists of twelve members that have been subjected to complete genome sequence analysis (including Chp3). These species fall into two distinct subfamilies, one subfamily which includes bacteriophages that infect enterobacteriaceae e.g. ϕX174 and G4. The second subfamily contains the chlamydiaphages, SpV4 (a bacteriophage that infects *Spiroplasma*) and MH2k (a bacteriophage that infects a *Bdellovibrio sp*) (Renaudin *et al.*, 1984; Brentlinger *et al.*, 2002). The main differences between these two subfamilies, is the absence of scaffolding proteins in the chlamydiaphage subfamily.

The ϕX174 subfamily attaches to receptive *E. coli* via an LPS receptor on the bacterial surface. Chlamydiae are also known to contain LPS in their outer membrane however, as the host range of Chp2 is restricted to four of the nine species of chlamydiae and periodate treatment of *C. abortus* does not reduce Chp2 binding, it is thought that LPS is not the chlamydiaphage receptor (Everson *et al.*, 2002). The receptor for chlamydiaphages has been shown to be protease sensitive and therefore it is highly likely to be an outer membrane protein (Everson *et al.*, 2002).

The chlamydiaphage genome has been predicted via computer analysis to be organised into eight major ORFs (Storey *et al.*, 1989), a ninth ORF was identified during the sequence analysis of Chp3. Each of these ORFs contain a ribosome-binding site, suggesting that they encode protein. ORF1, ORF2 and ORF3 are thought to encode the main structural proteins VP1, VP2 and VP3 respectively (Storey *et al.*, 1989). It is thought that ORF4 encodes the chlamydiaphage equivalent of protein A due to two regions of homology between these two proteins (Storey *et al.*, 1989). As the protein encoded by the chlamydiaphage ORF5 would be a similar protein to protein C it is hypothesized that ORF5 protein has a similar role to protein C during the Chp2 replication cycle. The protein

encoded by ORF8 protein would be a small basic protein similar to the ϕX174 protein J. The functions of the proteins encoded by ORF6, ORF7 and ORF9 are not known.

During these studies immuno-specific reagents to the proteins encoded by ORF2-ORF8 were produced, and a monoclonal antibody specific to VP1 was identified. These reagents were then used to screen Chp2 infected *C. abortus* inclusions. The expression of VP1, VP3, ORF4 protein ORF5 protein and ORF7 protein was demonstrated. This shows that ORF1, ORF3, ORF5, ORF5 and ORF7 encode proteins and are expressed in Chp2 infected chlamydiae. The expression of the proteins encoded by ORF2, ORF6 and ORF8 could not be demonstrated.

It is likely that the chlamydiaphages regulate gene expression in a similar way to ϕX174. Initiation of translation in ϕX174 begins immediately after stage I replication and occurs concurrently with transcription. Translation is regulated by a simple attenuation mechanism, which is based on the arrangement of promoters and terminators in the bacteriophage genome. Promoters are simply positioned in front of the genes that encode proteins required in the highest amounts. These genes are therefore most frequently transcribed and consequently translated in larger amounts. Alternatively, genes that encode proteins which are required in smaller amounts, are situated at the distal ends of operons and are consequently less frequently transcribed and therefore translated in smaller quantities (Hayashi *et al.*, 1988).

A significant discovery during this investigation was the discovery of two chlamydiaphage promoters situated upstream of the ORF4 and ORF5 translation-initiation codons that were active in an *E. coli in vitro* transcription/translation system. It was shown that the promoter situated upstream of the Chp2 ORF5 translation-initiation codon possibly produces a polycistronic mRNA consisting of ORF5, ORF1, ORF2, ORF3 and possibly ORF8, as the expression of ORF5 protein, VP1, VP2 and VP3 was demonstrated *in vitro*. The chlamydiaphage terminator is predicted to be located after ORF8. However, the expression of ORF8 protein was not observed possibly due to the small size of this protein or a low translation rate due to the position of ORF8 at the distal end of an operon. These two promoter regions are situated in a similar genomic location to promoters in SpV4 and ϕX174 and are positioned in front of genes encoding proteins that are predicted to be

required in large amounts (ORF4 protein, ORF5 protein) based on the amounts of protein A and protein C that are produced during ϕX174 replication.

The promoter situated upstream of the ORF5 translation-initiation codon was identified as the 'strongest' promoter based on the results obtained in a beta-galactosidase assay. The promoter situated upstream of ORF4 was identified as a 'weaker' *E. coli* promoter. The relative strength of the promoter regions may act as a method to regulate gene expression and control the amount of each protein that is expressed.

# Future Work

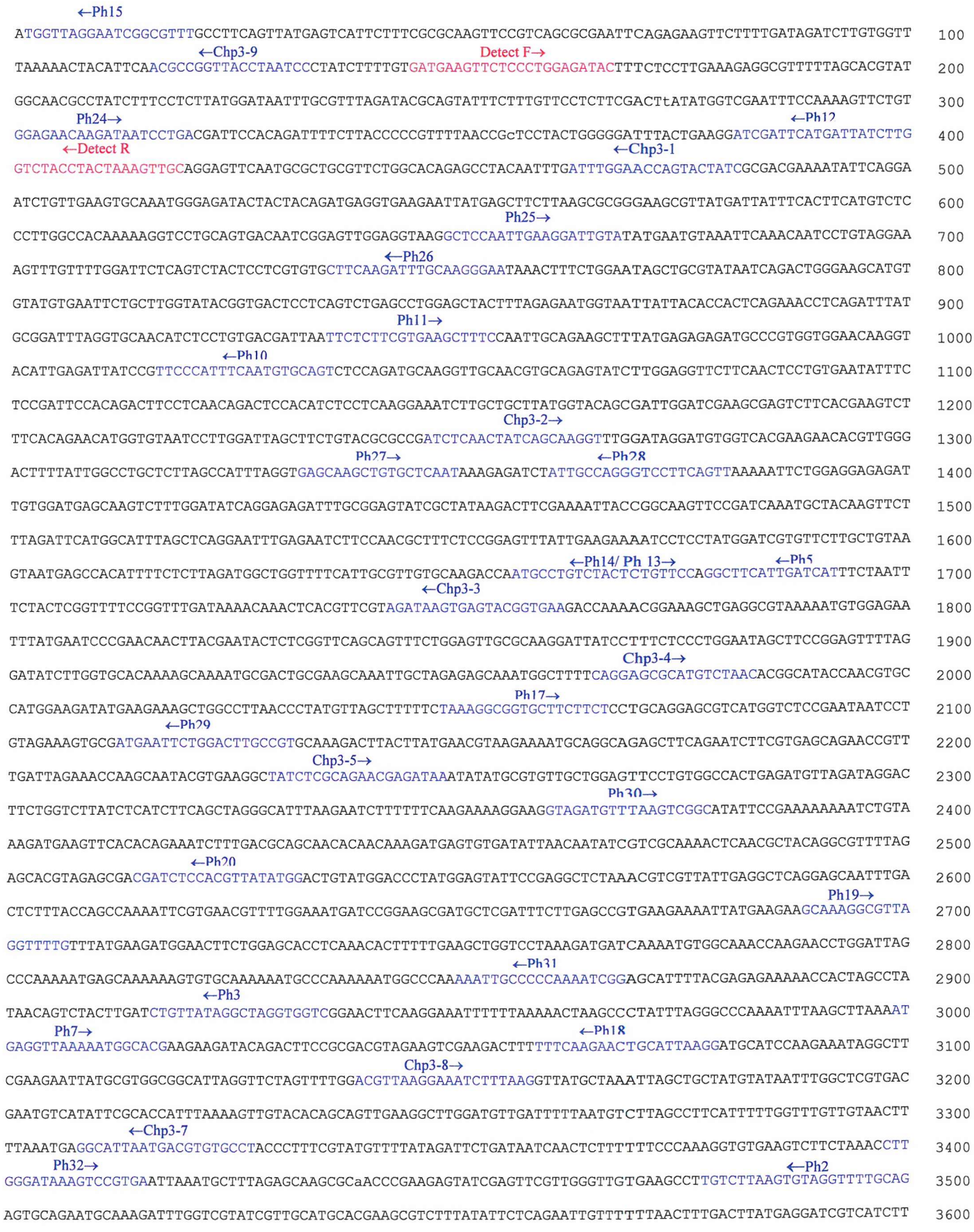
The functions of the proteins encoded by the Chp2 ORFs can only be speculated at this time. The development of immuno-specific reagents to detect each protein is an important step in providing tools for determining the function of each protein. Based on the experiments used to identify the functions of the ϕX174 proteins, similar investigations could be carried out to determine the functions of the Chp2 proteins, using the immuno-specific reagents to detect each protein.

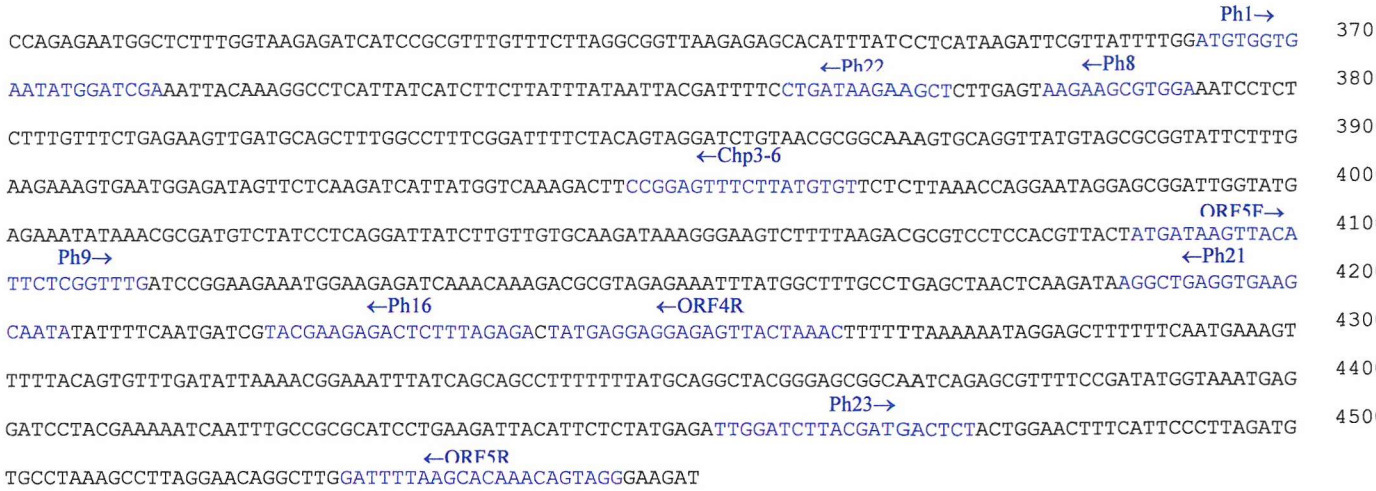
The precise transcriptional initiation and terminator sites have not yet been determined for the chlamydiaphage transcripts. The identification of two Chp2 promoter regions that are functional in *E. coli* may enable these points to be identified in *E. coli*, using plasmid vectors that contain the Chp2 genome.

The identification of the host receptors for Chp2/Chp3 and ϕCPAR39/ϕCPG1 would be an important step in the study of these bacteriophages and may lead to the production of a surrogate host for the genetic manipulation of these bacteriophages. This could lead to the development of a packaging system for these bacteriophages to generate mutants.

# APPENDIX 1

## Chp3 Oligonucleotide primers




Ph1→
3700
←Ph22
←Ph8
3800
3900
←Chp3-6
4000
ORF5F→
4100
Ph9→
←Ph21
4200
←Ph16
←ORF4R
4300
4400
Ph23→
4500
←ORF5R

# APPENDIX 2

## Chp3 sequence

VP1→
ATGGTTAGGAATCGGCGTTTGCCTTCAGTTATGAGTCATTCTTTCGCGCAAGTTCCGTCAGCGCGAATTCAGAGAAGTTCTTTTGATAGATCTTGTGGTT 100

TAAAAACTACATTCAACGCCGGTTACCTAATCCCTATCTTTTGTGATGAAGTTCTCCCTGGAGATACTTTCTCCTTGAAAGAGGCGTTTTTAGCACGTAT 200

GGCAACGCCTATCTTTCCTCTTATGGATAATTTGCGTTTAGATACGCAGTATTTCTTTGTTCCTCTTCGACTtATATGGTCGAATTTCCAAAAGTTCTGT 300

GGAGAACAAGATAATCCTGACGATTCCACAGATTTTCTTACCCCCGTTTTAACCGcTCCTACTGGGGGATTTACTGAAGGATCGATTCATGATTATCTTG 400

GTCTACCTACTAAAGTTGCAGGAGTTCAATGCGCTGCGTTCTGGCACAGAGCCTACAATTTGATTTGGAACCAGTACTATCGCGACGAAAATATTCAGGA 500

ATCTGTTGAAGTGCAAATGGGAGATACTACTACAGATGAGGTGAAGAATTATGAGCTTCTTAAGCGCGGGAAGCGTTATGATTATTTCACTTCATGTCTC 600

CCTTGGCCACAAAAAGGTCCTGCAGTGACAATCGGAGTTGGAGGTAAGGCTCCAATTGAAGGATTGTATATGAATGTAAATTCAAACAATCCTGTAGGAA 700

AGTTTGTTTTGGATTCTCAGTCTACTCCTCGTGTGCTTCAAGATTTGCAAGGGAATAAACTTTCTGGAATAGCTGCGTATAATCAGACTGGGAAGCATGT 800

GTATGTGAATTCTGCTTGGTATACGGTGACTCCTCAGTCTGAGCCTGGAGCTACTTTAGAGAATGGTAATTATTACACCACTCAGAAACCTCAGATTTAT 900

ORF7→
GCGGATTTAGGTGCAACATCTCCTGTGACGATTAATTCTCTTCGTGAAGCTTTCCAATTGCAGAAGCTTTATGAGAGAGATGCCCGTGGTGGAACAAGGT 1000
ORF7*

ACATTGAGATTATCCGTTCCCATTTCAATGTGCAGTCTCCAGATGCAAGGTTGCAACGTGCAGAGTATCTTGGAGGTTCTTCAACTCCTG**TGA**ATATTTC 1100

TCCGATTCCACAGACTTCCTCAACAGACTCCACATCTCCTCAAGGAAATCTTGCTGCTTATGGTACAGCGATTGGATCGAAGCGAGTCTTCACGAAGTCT 1200

TTCACAGAACATGGTGTAATCCTTGGATTAGCTTCTGTACGCGCCGATCTCAACTATCAGCAAGGTTTGGATAGGATGTGGTCACGAAGAACACGTTGGG 1300

ACTTTTATTGGCCTGCTCTTAGCCATTTAGGTGAGCAAGCTGTGCTCAATAAAGAGATCTATTGCCAGGGTCCTTCAGTTAAAAATTCTGGAGGAGAGAT 1400

TGTGGATGAGCAAGTCTTTGGATATCAGGAGAGATTTGCGGAGTATCGCTATAAGACTTCGAAAATTACCGGCAAGTTCCGATCAAATGCTACAAGTTCT 1500

TTAGATTCATGGCATTTAGCTCAGGAATTTGAGAATCTTCCAACGCTTTCTCCGGAGTTTATTGAAGAAAATCCTCCTATGGATCGTGTTCTTGCTGTAA 1600

VP1*
GTAATGAGCCACATTTTCTCTTAGATGGCTGGTTTTCATTGCGTTGTGCAAGACCAATGCCTGTCTACTCTGTTCCAGGCTTCATTGATCATTTC**TAA**TT 1700

ORF6→
TCTACTCGGTTTTCCGGTTTGATAAAACAAACTCACGTTCGTAGATAAGTGAGTACGGTGAAGACCAAAACGGAAAGCTGAGGCGTAAAAATGTGGAGAA 1800

VP2→
TTTATGAATCCCGAACAACTTACGAATACTCTCGGTTCAGCAGTTTCTGGAGTTGCGCAAGGATTATCCTTTCTCCCTGGAATAGCTTCCGGAGTTTTAG 1900

ORF6*
GATATCTTGGTGCACAAAAGCAAAATGCGACTGCGAAGCAAATTGC**TAG**AGAGCAAATGGCTTTTCAGGAGCGCATGTCTAACACGGCATACCAACGTGC 2000

CATGGAAGATATGAAGAAAGCTGGCCTTAACCCTATGTTAGCTTTTTCTAAAGGCGGTGCTTCTTCTCCTGCAGGAGCGTCATGGTCTCCGAATAATCCT 2100

GTAGAAAGTGCGATGAATTCTGGACTTGCCGTGCAAAGACTTACTTATGAACGTAAGAAAATGCAGGCAGAGCTTCAGAATCTTCGTGAGCAGAACCGTT 2200

TGATTAGAAACCAAGCAATACGTGAAGGCTATCTCGCAGAACGAGATAAATATATGCGTGTTGCTGGAGTTCCTGTGGCCACTGAGATGTTAGATAGGAC 2300

VP2* VP3→
TTCTGGTCTTATCTCATCTTCAGCTAGGGCATTTAAGAATCTTTTTTCAAGAAAAGGAAGG**TAG**ATGTTTAAGTCGGCATATTCCGAAAAAAAATCTGTA 2400

AAGATGAAGTTCACACAGAAATCTTTGACGCAGCAACACAACAAAGATGAGTGTGATATTAACAATATCGTCGCAAAACTCAACGCTACAGGCGTTTTAG 2500

AGCACGTAGAGCGACGATCTCCACGTTATATGGACTGTATGGACCCTATGGAGTATTCCGAGGCTCTAAACGTCGTTATTGAGGCTCAGGAGCAATTTGA 2600

CTCTTTACCAGCCAAAATTCGTGAACGTTTTGGAAATGATCCGGAAGCGATGCTCGATTTCTTGAGCCGTGAAGAAAATTATGAAGAAGCAAAGGCGTTA 2700

GGTTTTGTTTATGAAGATGGAACTTCTGGAGCACCTCAAACACTTTTTGAAGCTGGTCCTAAAGATGATCAAAATGTGGCAAACCAAGAACCTGGATTAG 2800

VP3*
CCCAAAAA**TGA**GCAAAAAAGTGTGCAAAAAATGCCCAAAAAATGGCCCAAAAATTGCCCCCAAAATCGGAGCATTTTACGAGAGAAAAACCACTAGCCTA 2900

TAACAGTCTACTTGATCTGTTATAGGCTAGGTGGTCGGAACTTCAAGGAAATTTTTTAAAAACTAAGCCCTATTTAGGGCCCAAAATTTAAGCTTAAAAT 3000

ORF8→
GAGGTTAAAAATGGCACGAAGAAGATACAGACTTCCGCGACGTAGAAGTCGAAGACTTTTTTCAAGAACTGCATTAAGGATGCATCCAAGAAATAGGCTT 3100

ORF8*
CGAAGAATTATGCGTGGCGGCATTAGGTTC**TAG**TTTTGGACGTTAAGGAAATCTTTAAGGTTATGCTAAATTAGCTGCTATGTATAATTTGGCTCGTGAC 3200

GAATGTCATATTCGCACCATTTAAAAGTTGTACACAGCAGTTGAAGGCTTGGATGTTGATTTTTAATGTCTTAGCCTTCATTTTTGGTTTGTTGTAACTT 3300

ORF4→
TTAAATGAGGCATTAATGACGTGTGCCTACCCTTTCGTATGTTTTATAGATTCTGATAATCAACTCTTTTTTTCCCAAAGGTGTGAAGTCTTCTAAACCTT 3400

GGGATAAAGTCCGTGAATTAAATGCTTTAGAGCAAGCGCaACCCGAAGAGTATCGAGTTCGTTGGGTTGTGAAGCCTTGTCTTAAGTGTAGGTTTTGCAG 3500

AGTGCAGAATGCAAAGATTTGGTCGTATCGTTGCATGCACGAAGCGTCTTTATATTCTCAGAATTGTTTTTTTAACTTTGACTTATGAGGATCGTCATCTT 3600

CCAGAGAATGGCTCTTTGGTAAGAGATCATCCGCGTTTGTTTCTTAGGCGGTTAAGAGAGCACATTTATCCTCATAAGATTCGTTATTTTGGATGTGGTG 3700

AATATGGATCGAAATTACAAAGGCCTCATTATCATCTTCTTATTTATAATTACGATTTTCCTGATAAGAAGCTCTTGAGTAAGAAGCGTGGAAATCCTCT 3800

CTTTGTTTCTGAGAAGTTGATGCAGCTTTGGCCTTTCGGATTTTCTACAGTAGGATCTGTAACGCGGCAAAGTGCAGGTTATGTAGCGCGGTATTCTTTG 3900

AAGAAAGTGAATGGAGATAGTTCTCAAGATCATTATGGTCAAAGACTTCCGGAGTTTCTTATGTGTTCTCTTAAACCAGGAATAGGAGCGGATTGGTATG 4000

AGAAATATAAACGCGATGTCTATCCTCAGGATTATCTTGTTGTGCAAGATAAAGGGAAGTCTTTTAAGACGCGTCCTCCACGTTACTATGATAAGTTACA 4100

TTCTCGGTTTGATCCGGAAGAAATGGAAGAGATCAAACAAAGACGCGTAGAGAAATTTATGGCTTTGCCTGAGCTAACTCAAGATAAGGCTGAGGTGAAG 4200

ORF4* ORF5→

CAATATATTTTCAATGATCGTACGAAGAGACTCTTTAGAGACTATGAGGAGGAGAGTTAC**TAA**ACTTTTTTAAAAAAATAGGAGCTTTTTTCAATGAAAGT 4300

TTTTACAGTGTTTGATATTAAAACGGAAATTTATCAGCAGCCTTTTTTTATGCAGGCTACGGGAGCGGCAATCAGAGCGTTTTCCGATATGGTAAATGAG 4400

GATCCTACGAAAAATCAATTTGCCGCGCATCCTGAAGATTACATTCTCTATGAGATTGGATCTTACGATGACTCTACTGGAACTTTCATTCCCTTAGATG 4500

ORF5*

TGCCTAAAGCCTTAGGAACAGGCTTGGATTTTAAGCACAAACAG**TAG**GGAAGAT

Ribosome Binding sites are indicated in Red.
Transcriptional initiation codons are indicated by →
Transcriptional terminators are indicated by *.

# Chapter 7

## Reference List


1. Alexander,E.R., Wang,S.P., Grayston,J.T. (1967). Further classification of TRIC agent from ocular trachoma and other sources by the mouse toxicity test. American Journal of Opthalmology *63*, 1469-1478.

2. Aoyama,A., Hamatake,R.K., Hayashi,M. (1983a). *In vitro* synthesis of bacteriophage ϕX174 by purified components. Proc.Natl.Acad.Sci.USA *80* , 4195-4199.

3. Aoyama,A., Hamatake,R.K., Mukai,R., Hayashi,M. (1983b). Purification of ϕX174 gene C protein. J.Biol.Chem *258*, 5798-5803.

4. Aoyama,A., Hayashi,M. (1986). Synthesis of bacteriophage ϕX174 *in vitro:* Mechanism of switch from DNA replication to DNA packaging. Cell *47*, 99-106.

5. Baker,R., Tessman,T. (1969). The circular genetic map of phage S13. Proc.Natl.Acad.Sci.USA *58*, 1438.

6. Bayer,M.E., Starkey,T.W. (1972). The adsorption of bacteriophage ϕX174 and its interaction with *Escherichia coli*; a kinetic and morphological study. Virology *49*, 236.

7. Belland,R.J., Scidmore,M.A., Crane,D.D., Hogan,D.M., Whitmire,W., McClarty,G., Caldwell,H.D. (2001). *Chlamydia trachomatis* cytotoxicity associated with complete and partial cytotoxin genes. Proc.Natl.Acad.Sci.USA *98*, 13984-13989.

8. Benbow,R.M., Zuccarelli,A.Z., Davis,A.Z., Sinsheimer,R.L. (1974). Genetic recombination in bacteriophage ϕX174. Journal of Virology *13*, 898.

9. Berdal,B.P., Scheel,O. (1993). *Chlamydia pneumoniae*-pathogenesis and perspectives. Tidsskr Nor Laegeforen *110*, 859-861.

10. Bernhardt,T.G., Roof,W.D., Young,R. (2000). Genetic evidence that the bacteriophage ϕX174 lysis protein inhibits cell wall synthesis. Proceedings Of The National Academy Of Sciences Of The United States Of America *97*, 4297-4302.

11. Bernhardt,T.G., Struck,D.K., Young,R. (2001). The lysis protein E of ϕX174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. Journal of Biological Chemistry. *276*, 6093-6097.

12. Bird, M., Dawson, C. R., Miao, Y., Schachter, J., and Lietman, T. Does the Clinical Exam Adequately Identify Ocular Chlamydial Infection in Trachoma. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 519-522. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

13. Birkelund,S., Johnsen,H., Christiansen,G. (1994). *Chlamydia trachomatis* serovar L2 induces protein tyrosine phosphorylation during uptake by HeLa cells. Infection and Immunity *64*, 4908.

14. Birkelund,S., Lundenmose,A.G., Christiansen,G. (1988). Chemical Cross-Linking of *Chlamydia trachomatis*. Infection and Immunity *56*, 654-659.

15. Birkelund,S., Lundenmose,A.G., Christiansen,G. (1989). Characterization of Native and Recombinant 75 Kilodalton Immunogens from *Chamydia trachomatis* Serovar-L2. Infection and Immunity *57*, 2683-2690.

16. Brentlinger,K.L., Hafenstein,S., Novak,C.R., Fane,B.A., Borgon,R., McKenna,R., Agbandje-McKenna,M. (2002). *Microviridae*, a family divided: Isolation, characterization and genome sequence of a ϕMH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus*. Journal of Bacteriology *184*, 1089-1094.

17. Brown,D.R., Schmidt-Glenewinkel,T., Reinberg,D., Hurwitz,J. (1983). DNA Sequences Which Support Activities of the Bacteriophage ϕX174 Gene A protein. Journal of Biological Chemistry. 8402-8412.

18. Buckley,K.J., Hayaski,M. (1987). The role of premature translational termination in the regulation of expression of the ϕX174 genes. Journal of Molecular Biology *198*, 599.

19. Burgess,A.B., Denhardt,D.T. (1969). Studies on ϕX174 Proteins. Journal of Molecular Biology *44*, 377-386.

20. Buxton,D., Anderson,I.E., Longbottom,D., Livingstone,M., Wattegedera,S., Entrican,G. (2002). Ovine chlamydial abortion: characterization of the inflammatory immune response in placental tissues. J Comp Pathol *127*, 133-141.

21. Cello,R.M. (1967). Ocular infections in animals with PLT (Bedsonia) group agents. American Journal of Opthalmology *63*, 1270-1273.

22. Chipman,P.R., Agbandje-McKenna,M., Renaudin,J., Baker,T.S., McKenna,R. (1998). Structural analysis of the spiroplasma virus, SpV4: implications for evolutionary variation to obtain host diversity among the *Microviridae*. Structure *6*, 135-145.

23. Clements,J.B., Sinsheimer,R.L. (1975). Process of infection with bacteriophage ϕX174. XXXVII. RNA metabolism in ϕX174 infected cells. Journal of Virology *15*, 151.

24. Collett,B.A., Newhall,W.J., Jersild,R.A.Jr., Jones,R.B. (1989). Detection of surface-exposed epitopes on *Chlamydia trachomatis* by immune electron microscopy. Journal of General Microbiology *135*, 85-94.

25. Cook,P.J., Davies,P., Tunnicliffe,W., Ayres,J.G., Honeybourne,D., Wise,R. (1998). *Chlamydia pneumoniae* and asthma. Thorax *53*, 254-259.

26. Darougar,S., Forsey,T., Brewerton,D.A., Rogers,K.L. (1980). Prevalence of antichlamydial antibody in London blood donors. Br J Verner Dis *56*, 404-407.

27. Dawson,C., Schachter,J. (1999). Can blinding trachoma be eliminated worldwide? Arch Ophthalmol *117*, 974.

28. Denhardt,D.T. (1977). The Isometric Single-Stranded DNA Phages. In: Comprehensive Virology, ed. H.Fraenkel-Conrat, R.R.WaynerNew York: Plenum, 1-104.

29. Dokland,T., Bernal,R.A., Burch,A., Pletnev,S., Fane,B.A., Rossmann,M.G. (1999). The role of scaffolding proteins in the assembly of the small single-stranded DNA virus øX174. Journal of Molecular Biology *288*, 595-608.

30. Dokland,T., McKenna,R., Ilag,L.L., Bowen,B.R., Incardona,N.L., Fane,B.A., Rossmann,M.G. (1997). Structure of a viral assembly intermediate with molecular scaffolding. Nature *389*, 308-313.

31. Douglas,A.L., Hatch,T.P. (1996). Mutagenesis of the P2 promoter of the major outer membrane protein gene of *Chlamydia trachomatis*. Journal of Bacteriology *178*, 5573-5578.

32. Dowell,C.E., Jansz,H.S., Zandberg,J. (1981). Infection of *Escherichia coli* K12 by bacteriophage ϕX174. Virology *114*, 252.

33. Duensing,T.D., Wing,J.S., van Putten,J.P. (1999). Sulfated polysaccharide-directed recruitment of mammalian host proteins: a novel strategy in microbial pathogenesis. Infection and Immunity *67*, 4463-4468.

34. Dwyer,R.S., Treharne,J.D., Jones,B.R., Herring,J. (1972). Chlamydial infection. Results of micro-immunofluorescence tests for the detection of type-specific antibody in certain chlamydial infections. Br J Vener Dis *48*, 452-459.

35. Eisenberg,S., Ascarelli,R. (1981). The A* protein is an inhibitor of DNA replication. Nucleic Acids Research *9* , 1991-2001.

36. Eisenberg,S., Scott,J.F., Kornberg,A. (1976). An enzyme system for replication of duplex circular DNA: The replicative form of ϕX174. Proc.Natl.Acad.Sci.USA *73*, 1594.

37. Eissenberg,L.G., Wyrick,P.B. (1981). Inhibition of phagolysosome fusion is localized to *Chlamydia psittaci*-laden vacuoles. Infection and Immunity *32*, 889-896.

38. Emini E.A., Hughes J.V., Perlow D.S., Boger J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific peptide. Journal of Virology *55*, 836-839.

39. Euzéby,J.P. (1980). Approved lists of bacterial names. Med J Aust *2*, 3-4.

40. Everett,K.D., Andersen,A.A. (1997). The ribosomal intergenic spacer and domain I of the 23S rRNA gene are phylogenetic markers for *Chlamydia* spp. International Journal of Systematic Bacteriology *47*, 461-473.

41. Everett,K.D.E., Bush,R.M., Andersen,A.A. (1999a). Emended description of the order *Chlamydiales*, proposal of *Parachlamydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new species, and standards for the identification of organisms. International Journal of Systematic Bacteriology *49*, 415-440.

42. Everett,K.D.E., Hornung,L.J., Andersen,A.A. (1999b). Rapid detection of the *Chlamydiaceae* and other families in the order *Chlamydiales*: Three PCR tests. Journal of Clinical Microbiology *37*, 575-580.

43. Everson,J.S., Garner,S.A., Fane,B., Liu,B.-L., Lambden,P.R., Clarke,I.N. (2002). Biological Properties and Cell Tropism of Chp2, A Bacteriophage of the Obligate Intracellular Bacterium *Chlamydophila abortus*. Journal of Bacteriology *184*, 2748-2754.

44. Everson,J.S., Garner,S.A., Lambden,P.R., Fane,B.A, Clarke,I.N. (2003). Host Range of Chlamydiaphages ϕCPAR39 and Chp3. Journal of Bacteriology *185*, 21, 6490-6492.

45. Fawaz FS, van Ooij,C., Homola,E., Mutka,S.C., Engel,J.N. (1997). Infection with *Chlamydia trachomatis* alters the tyrosine phosphorylation and/or localization of several host cell proteins including cortactin. Infection and Immunity *65*, 5301-5308.

46. Francis,T.J., Magill,T.P. (1938). An unidentified virus producing acute meningitis and pneumonitis in experimental animals. Journal of Experimental Medicine *68*, 147-163.

47. Frangioni,J.V., Neel,B.G. (1993). Solubilization and purification or enzymatically active glutathione-S-transferase (pGEX) fusion proteins. Anal.Biochem *210*, 179-187.

48. Freymeyer,D.K.2., Shank,P.R., Edgell,M.H., Hutchison,C.A.3., Vanaman,T.C. (1977). Amino acid sequence of the small core protein from bacteriophage ϕX174. Biochemistry *16*, 4550-4556.

49. Fujisawa,H., Hayashi,M. (1976). Gene A product of ϕX174 is required for site specific endonucleolytic cleavage during single-stranded DNA synthesis *in vivo.* Journal of Virology *19*, 416.

50. Fujisawa,H., Hayashi,M. (1977). Functions of gene C and gene D products of bacteriophage ϕX174. Journal of Virology *21*, 506-515.

51. Fukushi,H., Hirai,K. (1992). Proposal of *Chlamydia pecorum* sp. nov. for *Chlamydia* strains derived from ruminants. Int J Syst Bacteriol *42*, 306-308.

52. Grant,G.G.N., Jesse,J., Bloom,F.R., Hanahan,D. (1990). Differential plasmid rescue from transgenic mouse DNAs transformed inth *Esherichia coli* methylation-resistant mutants. Proc.Natl.Acad.Sci.USA *87*, 4645-4649.

53. Grayston,J.T. (1989). *Chlamydia pneumoniae,* strain TWAR. Chest *95*, 664-669.

54. Grayston,J.T., Wang,S.-P. (1975). New knowledge of chlamydiae and the diseases they cause. Infectious Agents and Disease *132*, 87-105.

55. Griffiths,P.C., Philips,H.L., Dawson,M., Clarkson,M.J. (1992). Antigenic and morphological differentiation of placental and intestinal isolates of *Chlamydia psittaci* of ovine origin. Veterinary Microbiology *30*, 167-177.

56. Griffiths,P.C., Plater,J.M., Martin,T.C., Hughes,S.L., Hughes,K.J., Hewinson,R.G., Dawson,M. (1995). Epizootic Bovine Abortion in a Dairy-Herd - Characterization of a *Chlamydia psittaci* Isolate and Antibody-Response. British Veterinary Journal *151*, 683-693.

57. Hackstadt,T. (1991). Purification and N-terminal amino acid sequences of *Chlamydia trachomatis* histone analogs. Journal of Bacteriology *173*, 7046-7049.

58. Hackstadt,T., Rockey,D.D., Heinzen,R.A., Scidmore,M.A. (1996). *Chlamydia trachomatis* interrupts an exocytic pathway to acquire endogenously synthesized sphingomyelin in transit from the Golgi apparatus to the plasma membrane. EMBO Journal *15*, 964-977.

59. Hackstadt,T., Scidmore,M.A., Rockey,D.D. (1995). Lipid metabolism in *Chlamydia trachomatis*-infected cells: directed trafficking of Golgi-derived sphingolipids to the chlamydial inclusion. Proc.Natl.Acad.Sci.USA *92*, 4877-4881.

60. Halberstaedter.L, ,v.P.S. (1907). Arb Gesundhamt *26*, 44-47.

61. Hamatake,R.K., Aoyama,A., Hayashi,M. (1985). The gene J of ϕX174. I. *In vitro* analysis of J protein function. Journal of Virology *54*, 345-350.

62. Hanahan,D. (1983). Studies on transformation of *E. coli* with plasmids. Journal of Molecular Biology *16*, 557-580.

63. Hare,M.J., Taylor-Robinson,D., Cooper,P. (1982). Evidence for an association between *Chlamydia trachomatis* and cervical intraepithelial neoplasia. Br J Obstet Gynaecol *89*, 489-492.

64. Harshbarger.J.C. (1976). *Chlamydia* (with phages), Mycoplasmas, and Rickettsiae in Chesapeake Bay Bivalves. Science *196*, 666-668.

65. Hatch,T.P. (1975). Competition between *Chlamydia psittaci* and L cells for host isoleucine pools: a limiting factor in chlamydial multiplication. Infection and Immunity *12*, 211-220.

66. Hatch,T.P. (1996). Disulfide cross-linked envelope proteins: the functional equivalent of peptidoglycan in chlamydiae? Journal of Bacteriology *178*, 1-5.

67. Hatch,T.P., Al-Hossainy,E., Silverman,J.A. (1982). Adenine nucleotide and lysine transport in *Chlamydia psittaci*. Journal of Bacteriology *150*, 662-670.

68. Hayashi,M. (1978). Morphogenesis of the Isometric Phages. In: The Single-Stranded DNA Phages, ed. D.T.Denhardt, D.Dressler, D.S.RayCold Springs Harbor, NY: Cold Springs Harbor Laboratory, 531-547.

69. Hayashi,M., Aoyama,A., Richardson,D.L., Hayashi,M.N. (1988). Biology of the Bacteriophage ϕX174. In The Bacteriophages. Plenum Publishing corporation, New York.

70. Hayashi,M., Fujimura,F.K., Hayashi,M.N. (1976). Mapping of *in vivo* messenger RNAs for bacteriophage ϕX174. Proc.Natl.Acad.Sci.USA *73*, 3519.

71. Hayashi,M., Hayashi,Y. (1970). Fractactionation of ϕX174 specific messenger RNA. Cold Spring Harbour Symp *35*, 17.

72. Hayes,L.J., Pickett,M.A., Conlan,J.W., Ferris,S., Everson,J.S., Ward,M.E., Clarke,I.N. (1990). The major outer-membrane proteins of *Chlamydia trachomatis* serovars A and B: intra-serovar amino acid changes do not alter specificities of serovar- and C subspecies-reactive antibody-binding domains. Journal of General Microbiology *136*, 1559-1566.

73. Hein,J.J. (1990). Unified approach to alignment and phylogenies. Methods in Enzymology. *183*, 626-645.

74. Henderson,I.R., Lam,A.C. (2001). Polymorphic proteins of *Chlamydia* spp. - autotransporters beyond the Proteobacteria. Trends in Microbiology *9*, 573-578.

75. Hendrix,R.W., Smith,M.C.M., Burns,R.N., Ford,M.E., Hatfull,G.F. (1999). Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proc.Natl.Acad.Sci.USA *96*, 2192-2197.

76. Holmes,D.S., Quigley,M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal.Biochem *114*, 193-197.

77. Horn,M., Wagner,M., Muller,K.D., Schmid,E.N., Fritsche,T.R., Schleifer,K.H., Michel,R. (2000). Neochlamydia hartmannellae gen. nov., sp nov (Parachlamydiaceae), an endoparasite of the amoeba Hartmannella vermiformis. Microbiology-UK *146*, 1231-1239.

78. Hsia, R-C., Ohayon, H., Gounton, P., and Bavoil, P. M. Phage infection of *Chlamydia psittaci* GPIC. 48. 1996. Bologna, Italy, Societa Editrice Esculapio. Proceedings of the Third European Society for Chlamydia Research. Stary, A.
Ref Type: Conference Proceeding

79. Hsia,R.-C., Pannekoek,Y., Ingerowski,E., Bavoil,P.M. (1997). Type III secretion genes identify a putative virulence locus of *chlamydia*. Molecular Microbiology *25*, 351-359.

80. Hsia,R.C., Ohayon,H., Gounon,P., Dautry-Varsat,A., Bavoil,P.M. (2000a). Phage infection of the obligate intracellular bacterium, *Chlamydia psittaci* strain guinea pig inclusion conjunctivitis. Microbes And Infection *2*, 761-772.

81. Hsia,R.C., Ting,L.M., Bavoil,P.M. (2000b). Microvirus of *Chlamydia psittaci* strain Guinea pig Inclusion Conjunctivitis: isolation and molecular characterization. Microbiology-Uk *146*, 1651-1660.

82. Iliffe-Lee,E.R., McClarty,G. (1999). Glucose metabolism in *Chlamydia trachomatis*: the 'energy parasite' hypothesis revisited. Molecular Microbiology *33*, 177-187.

83. Incardona,N.L., Selvidge,L. (1973). Mechanism of adsorption and eclipse of bacteriophage øX174. Journal of Virology *11*, 775-782.

84. Ishiwa,H., Tessman,I. (1968). Control of host DNA synthesis after infection with bacteriophages S13 and ϕX174. Journal of Molecular Biology *37*, 467-474.

85. Jameson B.A., Wolf H. (1988). The antigenic index: a novel algorithm for predicting antigenic determinants. Comput.Appl.Biosci. **4**, 181-186.

86. Jazwinski,S.M., Lindberg,A.A., Kornberg,A. (1975a). The gene H spike protein of bacteriophages ϕX174 and S13. II. Relation to synthesis of the parental replicative form. Virology *66*, 283.

87. Jazwinski,S.M., Marco,R., Kornberg,A. (1975b). The gene H spike protein of bacteriophages ϕX174 and S13. Virology *66*, 294-305.

88. Jennings,B., Fane,B.A. (1997). Genetic analysis of the øX174 DNA binding protein. Virology *227*, 370--377.

89. Johansson,M., Schon,K., Ward,M., Lycke,N. (1997). Genital tract infection with *Chlamydia trachomatis* fails to induce protective immunity in gamma interferon receptor- deficient mice despite a strong local immunoglobulin A response. Infection and Immunity *65*, 1032-1044.

90. Jones, R. B. and Van der Pol, B. *Chlamydia trachomatis* Isolates Resistant to Multiple Antibiotics. Chlamydial Infections. Proceedings of the Seventh International Symposium on Human Chlamydial Infections - Harrison Hot Springs, British Columbia, Canada. 24-29 June 1990., 519-522. 1990.
Ref Type: Abstract

91. Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Hyman,R.W., Olinger,L., Grimwood,J., Davis,R.W., Stephens,R.S. (1999). Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. Nature Genetics *21*, 385-389.

92. Kaltenboeck,B., Kousoulas,K.G., Storz,J. (1993). Structures of and allelic diversity and relationships among the major outer membrane protein (ompA) genes of the four chlamydial species. J Bacteriol *175*, 487-502.

93. Karunakaran,K.P., Blanchard,J.F., Raudonikiene,A., Shen,C., Murdin,A.D., Brunham,R.C. (2002). Molecular Detection and Seroepidemiology of the *Chlamydia Pneumoniae* Bacteriophage (ϕCpn1). Journal of Clinical Microbiology *40*, 4010-4014.

94. Kellerman,O.K., Ferenci,T. (1982). Maltose binding protein from *E coli.* Methods in Enzymology. *90*, 459-463.

95. Kornberg,A. (1980). DNA Replication., ed. FreemanSan Franscisco.

96. Krull,M., Klucken,A.C., Wuppermann,F.N., Fuhrmann,O., Seybold,J., Hippenstiel,S., Hegemann,J.H., Jantos,C.A., Suttorp,N. (1999). Signal transduction pathways activated in endothelial cells following infection with *Chlamydia pneumoniae*. Journal of Immunology *162*, 4834-4841.

97. Kuo,C.-C., Grayston,T. (1976). Interaction of *Chlamydia trachomatis* organisms and HeLa 229 cells. Infection and Immunity *13*, 1103-1109.

98. Langeveld,S.A., van Mansfeld,A.D., Baas,P.D., Jansz,H.S., Van Arkel,G.A., Weisbeek,P.J. (1978). Nucleotide sequence of the origin of replication in bacteriophage ϕX174 RF DNA. Nature *271*, 417-420.

99. Lehrach,H., Diamond,D., Wozney,J.M., Boedtker,H. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical re-examination. Biochemistry *16*, 4743-4751.

100. Lindqvist,B.H., Sinsheimer,R.L. (1967). The process of infection with bacteriophage ϕX174. XV. Bacteriophage DNA synthesis in abortive infections with a set of conditional lethal mutants. Journal of Molecular Biology *30*, 69-80.

101. Linney,E.A., Hayashi,M.N., Hayashi,M. (1972). Gene A of ϕX174. I. Isolation and Identification of its products. Virology *50*, 381-387.

102. Liu,B.L., Everson,J.S., Fane,B., Giannikopoulou,P., Vretou,E., Lambden,P.R., Clarke,I.N. (2000). Molecular characterization of a bacteriophage (Chp2) from *Chlamydia psittaci*. Journal of Virology *74*, 3464-3469.

103. Longbottom,D., Findlay,J., Vretou,E., Dunbar,S.M. (1998a). Immunoelectron microscopic localisation of the OMP90 family on the outer membrane surface of *Chlamydia psittaci*. FEMS Microbiology Letters *164*, 111-117.

104. Longbottom,D., Russell,M., Dunbar,S.M., Jones,G.E., Herring,A.J. (1998b). Molecular cloning and characterization of the genes coding for the highly immunogenic cluster of 90-kilodalton envelope proteins from the *Chlamydia psittaci* subtype that causes abortion in sheep. Infection and Immunity *66*, 1317-1324.

105. Majeed,M., Krause,K.H., Clark,R.A., Kihlstrom,E., Stendahl,O. (1999). Localization of intracellular Ca2+ stores in HeLa cells during infection with *Chlamydia trachomatis*. Journal of Cell Science *112*, 35-44.

106. Manire,G.P. (1966). Structure of purified cell walls of dense forms of meningopneumonitis organisms. Journal of Bacteriology *91*, 413.

107. Mathews,S.A., Sriprakash,K.S. (1994). The RNA polymerase of *Chlamydia trachomatis* has a flexible sequence requirement at the -10 and -35 boxes of its promoters. Journal of Bacteriology *176*, 3785-3789.

108. Mathews,S.A., Volp,K.M., Timms,P. (1999). Development of a quantitative gene expression assay for *Chlamydia trachomatis* identified temporal expression of σ factors. Federation of European Biochemical Societies Letters *458*, 354-358.

109. Matsumoto,A. (1982). Surface projections of *Chlamydia psittaci* elementary bodies as revealed by freeze-deep-etching. Journal of Bacteriology *151*, 1040-1042.

110. Mayol,R.F., Sinsheimer,R.L. (1970). Process of Infection with Bacteriophage ϕX174. Journal of Virology *6*, 310-319.

111. McClarty,G. (1994). Chlamydiae and the biochemistry of intracellular parasitism. Trends Microbiol *2*, 157-164.

112. McClenaghan,M., Herring,A.J., Aitken,I.D. (1984). Comparison of *Chlamydia psittaci* isolates by DNA restriction endonuclease analysis. Infection and Immunity *45*, 384-389.

113. Meijer, A. and Ossewaarde, J. M. Description of a Wider Diversity Within the Order *Chlamydiales* Than Currently Classified. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 13-16. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

114. Meyer,K.F. (1965). Ornithosis., ed. H.E.Biester, L.H.SchwarteAmes, Iowa: Iowa State University Press, 675.

115. Miller,J.H. (1972). Experiments in Molecular Genetics.New York: Cold Spring Harbor.

116. Minton.N.P (1984). Improved plasmid vectors for isolation of translational lac gene fusions. Gene *31*, 269-273.

117. Moulder, J. W. The Psittacosis Group as Bacteria. 95. 1964. New York: Wiley.
Ref Type: Pamphlet

118. Moulder,J.W. (1966). The relation of the psittacosis group (Chlamydiae) to bacteria and viruses. Ann.Rev.Microbiol *20*, 107-130.

119. Moulder,J.W. (1988). *Chlamydia* As Organisms. In: Microbiology *of Chlamydia*, ed. Barron.A.L. CRC Press, 4-15.

120. Mygind,P.H., Christiansen,G., Roepstorff,P., Birkelund,S. (2000). Membrane proteins PmpG and PmpH are major constituents of *Chlamydia trachomatis* L2 outer membrane complex. FEMS Microbiology Letters *186*, 163-169.

121. Newbold,J.E., Sinsheimer,R.L. (1970a). Kinetic of the attachment and eclipse steps of the infection. Journal of Virology *5*, 427-431.

122. Newbold,J.E., Sinsheimer,R.L. (1970b). Process of infection with bacteriophage ϕX174. Journal of Virology *5*, 427-431.

123. Nicolaidis,A.A., Drabble,W.T. (1979). Plasmid replication in *Salmonella typhimurium* LT2 and *Escherichia coli* K12: a differential effect of nalidixic acid. F.E.M.S.Letters *230*, 349-361.

124. Paavonen.J. *Chlamydia* and Cancer. Saikku, P. 1[1], 239-240. 2000. Helsinki, Finland. Proceedings, Fourth Meeting of the European Society for *Chlamydia* Research.
Ref Type: Conference Proceeding

125. Petersen.E.E., Bätz.O., and Bättcher, M. *Chlamydia trachomatis* serology: Prevalence of antibodies in women attending IVF centres. Saikku.P. 1[1], 133. 2000. Helsinki. Proceedings, Fourth Meeting of the European Society for *Chlamydia* Research.
Ref Type: Conference Proceeding

126. Pollock,T.J., Tessman,E.S., Tessman,I. (1978a). Identification of lysis protein E of bacteriophage ϕX174. Journal of Virology *28*, 408.

127. Pollock,T.J., Tessman,I., Tessman,E.S. (1978b). Radiological mapping of functional transcription units of bacteriophages phiX174 and S13. The function of proximal and distal promoters. Journal of Molecular Biology *124*, 147-160.

128. Potter, M. E., Kaufmann, A. K., and Brian D.Plikaytis, B. D. CDC. Psittacosis surveillance report, 1979. 32(SS01), 27-31. 1979. Atlanta. Annual Summary 1978.
Ref Type: Report

129. Pribnow,D. (1975a). Bacteriophage T7 early promoters: Nucleotide sequences of two RNA polymerase binding sites. Journal of Molecular Biology *99*, 419-443.

130. Pribnow,D. (1975b). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc.Natl.Acad.Sci.USA *72*, 784-789.

131. Pudjiatmoko, Fukushi,H., Ochiai,Y., Yamaguchi,T., Hirai,K. (1997). Phylogenetic analysis of the genus *Chlamydia* based on 16S rRNA gene sequences. Int J Syst Bacteriol *47*, 425-431.

132. Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G.,

Salzberg,S.L., Eisen,J., Fraser,C.M. (2000a). Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Research *28*, 1397-1406.

133. Read,T.D., Fraser,C.M., Hsia,R.C., Bavoil,P.M. (2000b). Comparative analysis of *Chlamydia* bacteriophages reveals variation localized to putative receptor binding domain. Microb Comp Genomics *5*, 223-231.

134. Renaudin,J., Bove,J.M. (1994). SpV1 and SpV4, Spiroplasm viruses with circular, single-stranded DNA genomes, and their contribution to the molecular biology of spiroplasmas. Advances in Virus Research *44*, 429-463.

135. Renaudin, J., Ricard, B., Garnier, M., and Bove, J. M. Purification and partial characterization of SpV4. Yale Journal of Biology and Medicine 57[6], 903. 1984.
Ref Type: Abstract

136. Ricard,B., Garnier,M., Bové,J.M. (1982). Characterization of SpV3 from spiroplasmas and discovery of a new spiroplasma virus (SpV4). Rev Infect Dis *4*, S275.

137. Richmond, S., Stirling, P., and Ashley, C. *Chlamydia* have phage too. Mardh, P-A., Holmes, K. K., Oriel, J. D., Piot, P., and Schachter, J. 41-44. 1982. Amsterdam, The Netherlands, Elsevier Biomedical Press. Fifth International Symposium on Human Chalmydial Infection. 15-6-0082.
Ref Type: Conference Proceeding

138. Rockey,D.D., Heinzen,R.A., Hackstadt,T. (1995). Cloning and characterization of a *chlamydia psittaci* gene coding for a protein localized in the inclusion membrane of infected cells. Molecular Microbiology *15*, 617-626.

139. Romero,H., Zavala,A., Musto,H. (2000). Codon usage in *Chlamydia trachomatis* is the result of strand-specific mutational biases and a complex pattern of selective forces. Nucleic Acids Research *28*, 2084-2090.

140. Saikku, P. *Chlamydia pneumoniae* - An update on Clinical Disease. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 443-454. 2002. Turkey.
Ref Type: Conference Proceeding

141. Sanger,F., Coulson,A.R., Friedmann,C.T., Air,G.M., Barrell,B.G., Brown,N.L., Fiddes,J.C., Hutchison,C.A., Slocombe,P.M., Smith,M. (1978). The nucleotide sequence of bacteriophage øX174. Journal of Molecular Biology *125*, 225-246.

142. Sanger,F., Nicklen,S., Coulson,A.R. (1977). DNA sequencing with chain-terminating inhibitors. Biochemistry *74* , 5463-5467.

143. Schachter,J., Caldwell,H.D. (1980). Chlamydiae. Ann.Rev.Microbiol *34*, 285-309.

144. Schachter,J., Meyer,K.F. (1969). Lymphogranuloma venereum II. Characterisation of some recently isolated strains. Journal of Bacteriology *99*, 636-638.

145. Schachter, J. and Moncada, J. Nucleic Acid Amplication Tests to Diagnose *Chlamydia Trachomatis* Genital Infection...The Glass is More Than Half Full. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 379-388. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

146. Schachter,J., Stephens,R.S., Timms,P., Kuo,C., Bavoil,P.M., Birkelund,S., Boman,J., Caldwell,H., Campbell,L.A., Chernesky,M., Christiansen,G., Clarke,I.N., Gaydos,C.,

Grayston,J.T., Hackstadt,T., Hsia,R., Kaltenboeck,B., Leinonnen,M., Ocjius,D., McClarty,G., Orfila,J., Peeling,R., Puolakkainen,M., Quinn,T.C., Rank,R.G., Raulston,J., Ridgeway,G.L., Saikku,P., Stamm,W.E., Taylor-Robinson,D.T., Wang,S.P., Wyrick,P.B. (2001). Radical changes to chlamydial taxonomy are not necessary just yet. Int J Syst Evol Microbiol *51*, 251-253.

147. Schaller,H., Gray,C., Herrmann,K. (1975). Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. Proc.Natl.Acad.Sci.USA *72*, 737-741.

148. Seed,B. (1982). Attachment of Nucleic Acids to Nitrocellulose and Diazonium-Subsituted Supports. In: Genetic Engineering: Principles and Methods, ed. J.K.Setlow, A.HollaenderNew York: Plenum publishing.

149. Shaw,A.C., Gevaert,K., Demol,H., Hoorelbeke,B., Vandekerckhove,J., Larsen,M.R., Roepstorff,P., Holm,A., Christiansen,G., Birkelund,S. (2002). Comparative proteome analysis of *Chlamydia trachomatis* serovar A, D and L2. Proteomics *2*, 164-186.

150. Shaw,D.C., Walker,J.E., Northrop,F.D., Barrell,B.G., Godson,G.N., Fiddes,J.C. (1978). Gene K, a new overlapping gene in bacteriophage G4. Nature *265*, 702-705.

151. Shen,L., Shi,Y., Douglas,A.L., Hatch,T.P., O'Connell,C.M., Chen,J.M., Zhang,Y.X. (2000). Identification and characterization of promoters regulating tuf expression in *Chlamydia trachomatis* serovar F. Arch Biochem Biophys *379*, 46-56.

152. Shlomai,J., Kornberg,A. (1980). An *Escherichia coli* replication protein that recognizes a unique sequence within a hairpin region in ϕX174 DNA. DNA, Pro.Natl.Acad.Sci.USA *77*, 799.

153. Shlomai,J., Polder,L., Arai,K., Kornberg,A. (1981). Replication of ϕX174 DNA with purified enzymes. I. Conversion of viral DNA to a supercoiled, biologically active duplex. J Biol Chem *256*, 5233-5238.

154. Siden,E.J., Hayashi,M. (1974). Role of gene B product in bacteriophage ϕ X174 development. Journal of Molecular Biology *89*, 1-16.

155. Sinsheimer,R.L. (1968). Bacteriophage ϕX174 and related viruses. Prog Nucleic Acid Res Mol Biol *8*, 115-169.

156. Smith,D.B. (1993). Purification of glutathione-S-transferase fusion proteins. Methods Mol.Cell Biol *4*, 220-229.

157. Spears,P., Storz,J. (1979). Biotyping of *Chlamydia psittaci* based on inclusion morphology and response to diethylaminoethyl-dextran and cycloheximide. Infection and Immunity *24*, 224-232.

158. Stamburski,C., Renaudin,J., Bove,J.M. (1990). Characterization of a promoter and a transcription terminator of *Spiroplasma melliferum* virus SpV4. Journal of Bacteriology *172*, 5586-5592.

159. Stamp,J.T., McEwen,A.D., Watt,J.A.A., Nisbet,D.I. (1950). Enzootic abortion in ewes I. Transmission of the disease. The Veterinary Record *62*, 251-256.

160. Stephens, R. S. Chlamydial Evolution: A Billion Years and Counting. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 3-12. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

161. Stephens,R.S., Kalman,S., Lammel,C., Fan,J., Marathe,R., Aravind,L., Mitchell,W., Olinger,L., Tatusov,R.L., Zhao,Q., Koonin,E.V., Davis R.W. (1998). Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. Science *282*, 754-759.

162. Stephens,R.S., Koshiyama,K., Lewis,E., Kubo,A. (2001). Heparin-binding outer membrane protein of Chlamydiae. Molecular Microbiology *40*, 691-699.

163. Stephens,R.S. (1993). Challenge of *Chlamydia* research. Infectious Agents and Disease *1*, 279-293.

164. Storey,C.C., Lusher,M., Richmond,S.J. (1989a). Analysis of the complete nucleotide sequence of Chp1, a phage which infects avian *Chlamydia psittaci*. Journal of General Virology *70*, 3381-3390.

165. Storey,C.C., Lusher,M., Richmond,S.J., Bacon,J. (1989b). Further characterization of a bacteriophage recovered from an avian strain of *Chlamydia psittaci*. Journal of General Virology *70*, 1321-1327.

166. Storz,J. (1963). Superinfection of pregnant ewes latently infected with a psittacosis-lymphogranuloma agent. Cornell Veterinarian *53*, 469-480.

167. Storz,J. (1988). Overview of Animal Diseases Induced by Chlamydial Infections. In: Microbiology of Chlamydia, ed. Barron.A.L. CRC Press., 4-15.

168. Studdert,M.J., Studdert,V.P., Wirth,H.J. (1981). Isolation of *Chlamydia psittaci* from cats with conjunctivitis. Aust Vet J *57*, 515-517.

169. Studier,F.W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. Journal of Molecular Biology *219*, 37-44.

170. Su,H., Watkins,N.G., Zhang,Y.X., Caldwell,H.D. (1990). *Chlamydia trachomatis*-host cell interactions: role of the chlamydial major outer membrane protein as an adhesin. Infection and Immunity *58*, 1017-1025.

171. Takahashi,T., ,M.M., Tsuruno,T., Mori,Y., Takashima,I., Hiramune,T., Kikuchi,N. (1997). Phylogenetic analyses of *Chlamydia psittaci* strains from birds based on 16S rRNA gene sequence. J Clin Microbiol *35*, 2908-2914.

172. Tan,M., Engel,J.N. (1996). Identification of sequences necessary for transcription in vitro from the *Chlamydia trachomatis* rRNA P1 promoter. Journal of Bacteriology *178*, 6975-6982.

173. Taraktchoglou,M., Pacey,A.A., Turnbull,J.E., Eley,A. (2001). Infectivity of *Chlamydia trachomatis* serovar LGV but not E is dependent on host cell heparan sulfate. Infection and Immunity *69*, 968-976.

174. Taylor-Robinson, D. and Thomas.B.J. *Chlamydia pneumoniae* in the vascular system. Saikku, P. 1[1], 373-374. 2000. Helsinki, Finland. Proceedings, Fourth Meeting of the European Society for Chlamydia Research.
Ref Type: Conference Proceeding

175. Tessman,E.S. (1965). Complementary groupd in phage S13. Virology *25*, 303.

176. Tessman,E.S. (1966). Mutants of bacteriophage S13 blocked in infectious DNA synthesis. Journal of Molecular Biology *17*, 218-236.

177. Tessman,E.S., Tessman,I. (1978). The Genes of Isometric Phages and Their Functions. In: The Single-Stranded DNA Phages, ed. D.T.Denhardt, D.Dressler, D.S.RayCold Spring Harbor, NY.: Cold Spring Harbor Laboratory, 9-29.

178. Tessman,E.S., Tessman,I., Pollock (1980). Gene K of bacteriophage ϕ X174 codes for a nonessential protein. Journal of Virology *33*, 557--560.

179. Thylefors,B., Negrel,A.D., Pararajasegaram,R., Dadzie,K.Y. (1995). Global data on blindness. Bulletin World Health Organisation *73*, 115-121.

180. Timms, P. and Mathews, S. Molecular Biology of *Chlamydia*. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 1[1], 585-594. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

181. Todd,W.J., Caldwell,H.D. (1985). The interaction of *Chlamydia trachomatis* with host cells: ultrastructural studies of the mechanism of release of a biovar II strain from HeLa 229 cells. Journal of Infectious Diseases *151*, 1037-1044.

182. Tuffrey,M., Folder,P., Gale,J., Taylor-Robinson,D. (1986). Salpingitis in mice induced by human strains of *Chlamydia trachomatis*. British Journal of Experimental Pathology *67*, 605-616.

183. Van Der Avoort,H.G.A.M., Teertstra,R., Versteeg,R., Weisbeek,P.J. (1983). Genes and Regulatory sequences of Bacteriophage ϕX174. Biochem.Biophysic.Acta. *741*, 94-102.

184. Van Der Avoort,H.G.A.M., Van Arkel,G.A., Weisbeek,P.J. (1982). Cloned Bacteriophage ϕX174 DNA Sequence Interferes with Synthesis of the Complementary Strand of Infecting Bacteriophage ϕX174. Journal of Virology *42*, 1-11.

185. Vandahl, B., Christiansen, G., and Birkelund, S. 2D-PAGE Analysis of the *Chlamydia puenmoniae* outer membrane complex. Schachter, J., Christiansen, G., Clarke, I. N., Hammerschlag, M. R., Kaltenboeck, B., Kuo, C-C., Rank, R. G., Ridgway, G. L., Saikku, P., Stamm, W. E., Stephens, R. S., Summersgill.J.T., Timms, P., and Wyrick, P. B. 547-550. 2002. Turkey. Proceedings of the Tenth International Symposium on Human Chlamydial Infections.
Ref Type: Conference Proceeding

186. Vandahl,B.B., Birkelund,S., Demol,H., Hoorelbeke,B., Christiansen,G., Vandekerckhove,J., Gevaert,K. (2001). Proteome analysis of the *Chlamydia pneumoniae* elementary body. Electrophoresis *22*, 1204-1223.

187. Vogel,H.J., Bonner,D.M. (1956). Acetylornithinease of *Escherichia Coli*: partial purification and some properties. Journal of Biological Chemistry. *218*, 97-106.

188. Wang,S.P., Grayston,J.T. (1963). Classification of trachoma virus strains by protection of mice from toxic death. Journal of Immunology *90*, 849-856.

189. Ward,M.E. (1983). Chlamydial classification, development and structure. British Medical Bulletin *39*, 109.

190. Wills,J.M., Watson,G., Lusher,M., Mair,T.S., Wood,D., Richmond,S.J. (1990). Characterization of *Chlamydia psittaci* isolated from a horse. Veterinary Microbiology *24*, 11-19.

191. Wuppermann,F.N., Hegemann,J.H., Jantos,C.A. (2001). Heparan sulfate-like glycosaminoglycan is a cellular receptor for *Chlamydia pneumoniae*. Journal of Infectious Diseases *184*, 181-187.

192. Wylie,J.L., Hatch,G.M., McClarty,G. (1997). Host cell phospholipids are trafficked to and then modified by *Chlamydia trachomatis*. Journal of Bacteriology *179*, 7233-7242.

193. Wyllie,S., Raulston,J.E. (2003). Identifying regulators of transcription in an obligate intracellular pathogen: a metal-dependent repressor in *Chlamydia trachomatis*. Molecular Microbiology *40*, 1027-1036.

194. Wyrick,P.B. (2000). Intracellular survival by *Chlamydia*. Cellular Microbiology *2*, 275-282.

195. Yanisch-Perron,C., Vieira,J., Messing,J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC18 vectors. Gene *33*, 103-119.

196. Zhang,J.P., Stephens,R.S. (1992). Mechanism of *C. trachomatis* attachment to eukaryotic host cells. Cell *69*, 861-869.

197. Zhong,G., Fan,P., Ji,H., Dong,F., Huang,Y. (2001). Identification of a chlamydial protease-like activity factor responsible for the degradation of host transcription factors. Journal of Experimental Medicine *193*, 935-942.