THE CLONING AND CHARACTERIZATION OF A GENE ENCODING THE 75 KILODALTON PROTEIN OF *CHLAMYDIA TRACHOMATIS*

A Thesis Presented to the Department of Medical Microbiology

Faculty of Medicine

University of Manitoba

In Partial Fulfillment of the Requirements for the

Degree of Master of Science

---

by

Sandra Lynn Danilition

1987

ISBN 0-315-44212-3

THE CLONING AND CHARACTERIZATION OF A GENE ENCODING THE 75 KILODALTON PROTEIN OF *CHLAMYDIA TRACHOMATIS*

BY

SANDRA LYNN DANILITION

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Acknowledgements

I would like to thank Dr. R.C. Brunham for providing me with an interesting, challenging and exciting project, and for all the support he has shown me. I would like to acknowledge the people of Allelix Inc., especially Bruce Collier and Dr. S. Winston, for all of their help and enthusiasm. For helping in the preparation of this thesis, I thank Gerry Bilan, Carol Sigurdson, and Paul Hazelton. For their critical assessment of this thesis, I thank Dr. P. McNicol and Dr. G.W. Hammond. Lastly, but certainly not least, I thank my mother, Pat Wilson, and my sister, Heather Fox, for their constant love, and support of all of my endeavors.

Table of Contents

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

Abstract

The gene encoding the 75kd antigen of Chlamydia trachomatis was cloned, sequenced, and the transcription and translation regulatory sequences were characterized.

A portion of the gene was first identified in a library of EcoRI digested genomic DNA in Lambda gt11. Recombinants containing the 75 kilodalton protein gene fragment were identified by positive reaction on immunoblots with a monoclonal antibody to a species-specific epitope of the 75 kilodalton protein. These recombinants expressed a fusion protein of B-galactosidase and a 38 kilodalton peptide of the 75 kilodalton protein. The lambda recombinants (EDR's) contained an insert of 2.2 kilobases which consisted of approximately half of the coding region of the 75 kilodalton protein gene.

The complete gene was subsequently isolated from a library constructed with Sau3AI partially digested genomic DNA cloned into the plasmid vector pUC18. The isolate, pERU-S1, expressed a protein which migrated through sodium dodecyl sulfate polyacrylamide gels at the same rate as the 75 kilodalton protein of C. trachomatis. The novel protein reacted with the monoclonal antibody to the chlamydial 75 kilodalton protein in western blots.

The gene was subcloned on a 2.9 kilobase DNA fragment in pERU-S2. The subclone retained complete expression and immunoreactivity of the 75 kilodalton protein.

The DNA sequence of the gene was obtained through a

variety of subclones and deletions of pERU-S2 and EDR4-2. The coding region of the gene was determined to be 1956 nucleotides in length, coding for 652 amino acids. The gene sequence consisted of one stop codon, followed by a typical rho-independent transcription terminator 25 bases downstream. Upstream of the coding sequence, a Shine-Dalgarno sequence of 5 bases in length, beginning at -7, was identified. Further upstream, centered around -133 and -117, putative promoter elements were identified.

Introduction

Chlamydia trachomatis is an obligate intracellular bacterium that utilizes an intricate life cycle of growth. This life cycle consists of an extracellular elementary body (EB) which is non-metabolic, but infectious, and post infection, the EB undergoes a metabolic and morphologic change into the metabolically active form, the reticulate body (RB).

The membrane of chlamydiae possesses a variety proteins. The function(s) of these proteins has not been definitively determined. Studies have shown that some of the proteins, Major Outer Membrane Protein (MOMP), 60 kilodalton (kd) protein, 57kd protein, and 12kd protein, may be involved in providing the structural integrity of the membrane (Newhall and Jones, 1983; Hatch et al., 1984). The 32kd and 18kd membrane proteins have been implicated in the attachment step of infection by binding to host cell membranes (Hackstadt, 1986; Wenman and Meuser, 1986).

Studies have also shown that certain proteins may be important in eliciting antibody that provides protective immunity against ascending infection in human hosts. These antigens include the 100kd, 75kd, 60kd, 57kd, 32kd, and 29kd (Brunham et al., 1987).

Expression studies have suggested a developmentally regulated mechanism of gene expression for the 60kd, 57kd, 12kd, MOMP, and 74kd proteins (Hatch et al., 1984;Hatch et

al., 1986; Newhall, 1987; Palmer and Falkow, 1986ab). Synthesis of MOMP has been detected as early as 12hrs post infection, while the 60kd, 57kd and 12kd antigens are not produced in significant quantities until later in the reticulate body (RB) stage as RB's begin to reorganize into EB's (Newhall, 1987). Expression and synthesis of MOMP appears to occur significantly during the first 36hrs post infection, suggesting constitutive expression during the early RB stage and inhibition of expression prior to reorganization into EB's. Expression of 60kd, 57kd, and 12kd, on the other hand, appears to be repressed until later in the RB cycle when expression is induced.

Expression of the 74kd antigen appears to be controlled by a temporal mechanism, where antigen is produced early in the RB stage only (Palmer and Falkow, 1986a). Early expression of 74kd antigen is analogous to MOMP, and may be induced by a similar mechanism; however, it's expression is repressed sometime before that of the MOMP gene.

The intent of this project was to isolate a gene of Chlamydia trachomatis, which encodes one of the antigens that appear to elicit an immune response involved in protective immunity to invasive infection. The purpose of cloning the gene was to sequence the DNA and analyze the sequence in an attempt to identify genetic control mechanisms governing expression and provide some insight into the developmental regulation of the reticulate body.

Literature Review

A) Chlamydiae

(i) Health Problem

It has been well documented that C. trachomatis is a major health problem in the world today; it is the leading cause of preventable blindness in third world countries (Schachter, 1978), and one of the most common sexually transmitted pathogens in North America (Holmes, 1981). In fact, the incidence of chlamydial genital infections in North America has surpassed that of gonococcal infections (Schachter, 1985). It has been estimated that greater than 25% of women with acute salpingitis have chlamydial infection (Sweet et al., 1983) and that approximately 20% of infected women become infertile as a result (Westrom, 1980). Studies by Brunham and co-workers suggest that C. trachomatis is the cause of two-thirds of tubal infertility cases and one-third of ectopic pregnancies (Brunham et al., 1985; Brunham et al., 1986).

The other concern for women with chlamydial infection is transmission to sex partners and to a newborn child. The birth of a child to an infected mother can result in transmission of the organism to that child, and subsequent development of conjunctivitis and/or pneumonia. It is obvious then that chlamydial infections pose a serious health problem.

(ii) Chlamydial Membranes: Structure and Function

The genus Chlamydiales, consists of two species, C.

trachomatis and C. psittaci, which have been grouped together based on their similar growth characteristics and life cycle (Storz, 1971). C. trachomatis consists of three biovars; mouse (pneumonitis), lymphogranuloma venerum (LGV) (agent of LGV), and trachoma (agent of human oculogenital diseases other than LGV). Chlamydiae resemble gram-negative bacteria in that they posses an outer membrane that is disrupted in the presence of EDTA (Nariter et al. 1976) or polymixin B (Matsumoto and Manire, 1973). They also possess an inner membrane; however, contrary to other gram-negative bacteria, they possess neither a layer of peptidoglycan (Tamura and Manire, 1968) nor muramic acid (Garret et al., 1974; Manire and Tamura, 1967; Barbour et al., 1982).

Extracellularly, chlamydiae exist as elementary bodies (EB's) of 200-400 nm in diameter, and possess a rigid cell wall. It is during this stage that chlamydiae are infective, ie: they have the ability to attach to host cells and induce phagocytosis. EB's are non-metabolic, however upon internalization, the EB undergoes reorganization into the metabolically active reticulate body (RB) within 6-8 hours. The RB is a larger structure of 600-1,000nm in diameter which multiplies by binary fission inside a phagosome within the host cell. RB's are non-infectious, unstable outside the host and possess a fragile cell membrane.

Tamura and Manire (1967) first proposed that the EB cell wall rigidity was provided by extensive disulfide

linkages within the outer membrane since they found a lack of cysteine within the fragile membrane of RB's. In recent years, this theory has gained substantial support.

Caldwell et al. (1981) identified a 39.5Kd protein which contributes greater than 60% of the chlamydial outer membrane complex (COMC) protein. They suggested that this protein resembles the matrix proteins of other gram-negative bacteria and called it the chlamydial major outer membrane protein (MOMP). Extraction of MOMP from COMC's resulted in complete loss of structural integrity, suggesting that MOMP may function as a structural protein, conferring the rigidity found in elementary bodies (EB's).

Newhall and Jones (1983) found that MOMP exists as a large complex within the outer membrane as a result of extensive intra- and inter-chain disulfide cross linking between MOMP and other outer membrane proteins. They identified two other outer membrane proteins which appear to be integral components of the complex, the 15Kd and 60Kd proteins. A combination of their observations and other's, led them to propound a theory for chlamydial reorganization. They proposed that cleavage and formation of disulfide bonding within the outer membrane may be involved in the process of transformation of chlamydia from one form (EB) into the other (RB).

In an attempt to identify a difference in specific cysteine-rich polypeptides between EB's and RB's, Hatch et al. (1984) looked at the protein profiles of outer

membranes. They found that MOMP was present in both EB's and RB's in almost equal amounts. However, MOMP was dissociated from the RB membrane much more easily than from EB membranes, suggesting that cross linking is not as extensive in the RB. Three cysteine-rich proteins, 60Kd, 59Kd, and 12Kd, were identified in the EB membrane but not in the RB membrane. Commencement of synthesis of these cysteine-rich proteins does not begin until 18-24hours post-infection (Hatch et al., 1984; Hatch et al., 1986; Newhall, 1987) when RB's begin reorganization into EB's.

Bavoil et al.(1984) in a similar study, proposed a model of chlamydial reorganization to account for the observations. Proteins of the EB outer membrane are cross-linked by disulfide bonds rendering the organism rigid and impermeable to hydrophilic solutes. Once internalized within the phagosome, the EB is subjected to reducing conditions, membrane pores open, ATP and nutrients are taken in; thereby inaugurating reorganization into the RB. The RB multiplies until ATP and reducing power run down, resulting in a simultaneous decrease in metabolic activity and oxidation of free sulfhydryls, closing the pores and re-rigidifying the membrane.

Chlamydiae are obligate intracellular parasites, infecting cells of the columnar epithelial type. Successful infection of a host cell involves three basic steps: attachment, ingestion and inhibition of phagolysosomal fusion. Attachment is believed to be the rate limiting step

of chlamydial pathogenesis and was found to be enhanced by centrifugation of inocula onto cell cultures (Weiss and Dressler, 1960; Kuo and Grayston, 1976), and the presence of the polycation diethylaminoethyl (DEAE)-dextran (Harrison, 1970; Kuo et al., 1972; Kuo et al., 1973; Kuo and Grayston, 1976; Rata and Nichols, 1971).

Kuo and Grayston (1976) found that attachment of chlamydia could occur on cell lines such as fetal tonsil that are relatively insusceptible to infection. This observation implied that attachment is not a critical factor in determining susceptibility to infection, and this first step in chlamydial infection may involve a non-specific interaction with a ubiquitous component on host cells.

Treatment of L cells (Byrne, 1976; Bryne and Moulder, 1978) or HeLa 229 cells (Byrne and Moulder, 1978) with trypsin resulted in inhibition of infection by chlamydial organisms. Inhibition of infection led to the suggestion that a host cell membrane component of protein or glycoprotein may be acting as a receptor for chlamydiae. In an attempt to decipher the nature of this host "receptor", Levy (1979), showed that attachment of Chlamydia species to various cell lines could be blocked by the lectin wheat germ agglutinin (WGA). The blockage intimated that attachment involves interaction of chlamydia and an N-acetyl-D-glucosamine containing substance on the surface of host cells.

Soderlund and Kihlstrom (1983) suggested that

chlamydiae utilize the mechanism of receptor-mediated endocytosis via clathrin coated pits of the host for uptake. The hypothesis was disproved by Ward and Murray (1984) based on a variety of observations showing that chlamydial uptake did not meet the criteria necessary for receptor-mediated endocytosis.

Byrne and Moulder (1978) found that both L cells and also HeLa cells could phagocytose chlamydial species at a greater rate and to a greater extent than they could phagocytose either Escherichia coli cells or polystyrene latex spheres. Attachment and ingestion of chlamydiae occurred with ultraviolet light-killed EB's. This finding suggested that some component of chlamydia was able to induce the uptake of the organism and led to use of the term "parasite-specified phagocytosis." Inhibition of infection by interference with the attachment step has subsequently been shown to occur upon mild heating of inocula (Byrne, 1976; Kuo and Grayston, 1976; Byrne and Moulder, 1978), as well as treatment with antibody (neutralization), (Byrne and Moulder, 1978).

Treatment of EB's with proteases in order to elucidate an involvement of surface proteins has been investigated by a number of researchers (Byrne and Moulder, 1978; Levy and Moulder, 1982; Hackstadt and Caldwell, 1985). In all of these studies, protease treatment did not significantly alter the ability of chlamydia to infect. Hackstadt and Caldwell (1985) suggested that the absence of inhibition may

be a result of most of the protein components remaining within the membrane post-proteolysis.

Recently, two proteins of the chlamydial outer membrane have been identified which specifically bind to HeLa cell membranes (Hackstadt, 1986; Wenman and Meuser, 1986). These outer membrane proteins, the 18Kd and 31/32Kd antigens, may be the chlamydial component through which attachment is attained to the host cell and triggering of the phagocytic process occurs.

It has been shown that purified cell envelopes of C. psittaci can be ingested by L cells and prevent phagolysosomal fusion (Eissenberg et al., 1983; Levy and Moulder, 1982). These experiments and the others mentioned previously, imply that components of the EB surface are involved in the three steps of infection: attachment, ingestion, and inhibition of phagolysosomal fusion; however, the exact mechanism of chlamydial infection is yet to be determined.

(iii) Chlamydial Antigens

The surface antigens of chlamydial species have been characterized by serological studies. Chlamydiae possess antigens distinguishing genus, species, subspecies, and serotype.

The genus-specific antibodies show cross reaction between the two species of chlamydia. The best characterized of these cross-reacting antigens is the lipopolysaccharide component which is heat stable and

complement-fixing (Allan, 1986). Other genus-specific antigens have been detected but not as well characterized.

The antigens inducing species-specific antibodies are those which share epitopes within one species of chlamydia. A total of 19 distinct species-specific antigens of C. trachomatis have been detected (Allan, 1986). This group of antigens includes the 60kd and 62kd cysteine rich polypeptides and certain epitopes of MOMP (Newhall et al., 1982;Caldwell and Schachter, 1982). A total of 15 distinct species specific antigens of C. psitacci have been detected (Allan, 1986).

The serovar antigens of chlamydia are those which distinguish the 15 different serovars of C. trachomatis. Certain serovars contain a MOMP which shows serovar-specific epitopes (Caldwell and Schacter, 1982).

In studying the humoral immune response to chlamydial infection, Newhall et al. (1982) found that sera of individuals with proven C. trachomatis genital infection almost always reacted with the 60kd and 62kd species-specific antigens, and rarely did the sera react to MOMP antigens.

Brunham et al. (1987) studied the humoral immune response in a group of women with C. trachomatis infections who underwent therapeutic abortions. They found that the serum from women who did not develop post abortal salpingitis contained antibodies to the 75kd, 60kd, and 57kd antigens more often than serum from women who developed post

abortal salpingitis. Furthermore, antibodies to the 100kd, 32kd, and 29kd antigens were present only in serum from women who did not develop salpingitis. As a result of their findings, Brunham et al. suggested that antibodies to these antigens may be important in inhibition of invasive chlamydial infection.

(iv) Molecular Biology

Another approach to understanding the pathogenesis and development of chlamydial organisms is through the use of recombinant DNA technology. Isolating and cloning specific genes provides a means by which both the genetics and the proteins of this organism may be examined. Determination of the DNA sequence of the genes can be used to look at homology both within and between different bacterial species to determine evolutionary relationships.

The chlamydial genome is approximately 600-850kb in size with a coding potential of 600 polypeptides (Allan, 1986). Cloning and expressing structural gene products of chlamydia could provide an abundance of protein that is easily accessible for purification. Purified protein can be used for studying the nature of the protein itself as well as using it in a variety of other experiments. The protein can be utilized for epitope mapping by producing polyclonal and/or monoclonal antibodies to it. Purified protein from clones has potential for studying the usefulness of that protein in vaccine production, as proposed by Kaul et al. (1987). This group cloned the 18Kd antigen of C.

trachomatis which has been identified in the binding of HeLa cell membranes (Hackstadt, 1986; Wenman and Meuser, 1986). The gene was isolated from a gene bank consisting of PstI partially digested chlamydial DNA of 4kb to 6kb inserted into the vector pUC8. The gene bank was screened by identifying recombinants expressing a novel protein which bound HeLa cell membranes. The expressed protein possessed the same membrane binding activity as the native chlamydial antigen (Kaul et al., 1987), and was located within the outer membrane of E. coli. The 18kd antigen was given the name clanectin to describe its cell binding activity.

The genes of several other chlamydial antigens have been cloned with the intention of gaining information on the biological nature of these proteins, hopefully leading to knowledge of their function and involvement in chlamydial development and pathogenesis.

The gene encoding MOMP and portions of this gene have been isolated by several investigators (Allan, et al., 1984; Stephens et al., 1985; Stephens et al., 1986; Nano et al., 1985; Pickett et al., 1987). Allan et al. (1984), isolated the gene from a library of Sau3AI partially digested chlamydial genomic DNA in the vector Lambda 1059. By screening the library with high-titer human serum arising from a serovar L1 infection, they identified a clone which expressed immunoreactive polypeptides of 40kd and 41kd in size. The authors believe these two products are different conformations of the same protein.

Stephens et al. (1985) isolated a clone which expressed a 15kd carboxyl-terminal peptide of MOMP as a fusion with B-galactosidase. The recombinant was identified in a library of DNase I digested C. trachomatis L2 DNA inserted into Lambda gt11 and screened with rabbit anti-C. trachomatis L2 serum. The recombinant protein reacted with monoclonal antibodies to the species-, subspecies-, and type-specific epitopes of MOMP. Therefore this clone provided information on important epitopes of the MOMP, locating them to the carboxy-terminus of the molecule. Cloned genes can be used in hybridization studies to identify the degree of homology between the same gene of different strains of an organism, thereby depicting the degree of relatedness between the strains and evolutionary trends. In order to confirm the antibody findings for the fusion protein, the DNA insert of the recombinant molecule was hybridized to DNA of the 15 other C. trachomatis serovars and to C. psittaci. The insert hybridized to all of the C. trachomatis serovars and weakly to C. psittaci, suggesting that this portion of the MOMP gene is highly conserved and may contain functional domains.

Often, when cloning genes, expression of the foreign protein results in death of the host cell or foreign protein instability (Young and Davis, 1985). For instance, membrane proteins which function as porins may be lethal to the host. Because of this problem, screening libraries by methods other than expression of foreign proteins must sometimes be

used to identify a cloned gene. These other methods of screening could involve the use of oligonucleotide probes, however this requires some initial knowledge of the sequence of the gene involved.

Cloning of the MOMP gene by screening with nucleotide probes was used by three groups using three different approaches. Stephens et al. (1986) used the DNA fragment containing the carboxy-terminus of MOMP to screen a library of BamHI digested chlamydial DNA in Lambda 1059. The MOMP gene was found in a recombinant clone containing a 9.2kb insert and the gene was located within approximately 3kb of the 5' end. The gene was subsequently analyzed by DNA sequencing.

Nano et al. (1985) generated a series of oligonucleotide probes whose sequences were derived from the known amino acid sequence of the MOMP amino terminal. The probes were used in screening a BamHI-PstI digest of chlamydial DNA for the complete MOMP gene by Southern blot. DNA in the size range of the hybridization signal was recovered and inserted into the plasmid vectors pEMBL8 and pEMBL9. The plasmid recombinants were subsequently screened with a mixture of monoclonal antibodies directed to MOMP, therefore this procedure still required the expression of MOMP for detection. One recombinant was isolated which expressed a 15kd peptide reactive to a type-specific monoclonal antibody to MOMP. Nano's group also took the direct approach of screening a gene bank in pUC8 of 4kb to

6kb Sau3AI digested chlamydial DNA, with a mixture of monoclonal antibodies to MOMP (Nano et al., 1985). This procedure yielded a recombinant molecule expressing a 51kd peptide containing a subspecies-specific epitope of MOMP.

The complete MOMP L1 gene was isolated by screening a lambda phage library of Sau3AI partially digested DNA with a 50 base oligonucleotide probe whose sequence was derived from the sequence of MOMP L2 (Pickett et al., 1987). The gene was localized to a 2.0kb fragment of the clone and further characterized by DNA sequence analysis.

Cloning of chlamydial proteins is another approach to the study of the developmental cycle of this organism. Isolation of genes encoding developmentally regulated proteins of chlamydia, such as the cysteine-rich 60kd, 59kd, and 12kd proteins, (Hatch et al., 1984; Hatch et al., 1986; Newhall, 1987), and the 18kd protein, clanectin (Wenman and Meuser, 1986; Kaul et al., 1987), could lead to important advances in understanding both their transcriptional/translational control and their biological function. Cloned genes can be used to look at the developmental regulation of RNA synthesis by probing RNA isolated at various time points post infection. Analysis of the DNA sequence bordering a gene can lead to information regarding the regulatory elements involved in the gene's expression. The amino acid sequence derived from DNA sequence could provide information as to the biological function of the protein by analysis of the protein's primary

and secondary structure. Computer homology searches of the DNA and amino acid sequences can generate evolutionary and functional information by comparison of these sequences to the sequence of proteins with known function.

The DNA sequence of the MOMP L1 and L2 serovar genes had been determined (Stephens et al., 1986; Pickett et al., 1987). Sequence comparison of these two genes reveals significant homology. Identical Shine-Dalgarno sequences (Shine and Dalgarno, 1974), or ribosome binding sites were identified for both gene sequences and putative promoter elements were suggested for the L1 serovar. The regulatory sequences were identified by comparison with analogous sequences of E. coli.

As previously discussed, the gene encoding the 18kd antigen has been cloned and expressed (Kaul et al., 1987); however, the gene is under the transcriptional control of the lac z promoter of the vector. The DNA sequence of this gene was determined, and no typical prokaryotic promoter-like sequences were identified. A Shine-Dalgarno sequence of 5'AGGA3' was found 10 base pairs (bps) upstream of the translational start site, suggesting that the gene is under typical prokaryotic translational control. Analysis of the amino acid sequence derived from the DNA sequence revealed two hydrophilic regions: amino acids 20 to 30 and 125 to 162. Therefore these two regions are likely surface-exposed and may elicit antigenic response. Two short amino acid sequences were identified which are similar to the cell

attachment domain of fibronectin; however, sequence comparison to E. coli genes failed to reveal any homology.

The genes encoding two of the cysteine-rich outer membrane proteins, 60kd and 12kd, have been isolated within a single clone (Bavoil et al., 1986). The genes were isolated from a partial gene bank of CpfI partially digested chlamydial DNA in Lambda-EMBL4. Detection of expression of these genes within E. coli by western blot with serum known to contain anti-60kd antibodies did not occur. When inserted into an in vitro transcription-translation system with ^{35}S-cysteine or ^{35}S-methionine, gene products of 60kd and 12kd were expressed. The presence of these two genes within the same recombinant molecule insinuated that they may be linked within the chlamydial genome and under the same regulatory control as in an operon (Bavoil et al., 1986). DNA sequence analysis of this recombinant could help to define the control which these genes are under.

Other antigens of chlamydia which have been cloned include the 19kd protein (Wenman and Lovett, 1982;Palmer and Falkow, 1986b), the 74kd protein (Kaul and Wenman, 1985; Palmer and Falkow, 1986b; Bavoil et al., 1986), and a genus-specific epitope of the lipopolysaccharide (LPS) antigen (Nano and Caldwell, 1985).

The gene encoding the 19kd antigen was identified in a library of Lambda-1059 containing Sau3AI digested genomic DNA. The library was screened with patients' sera known to contain antibodies to chlamydial antigens (Wenman and

Lovett, 1982). This gene has also been isolated by Palmer and Falkow (1986b).

The 74kd antigen of C. trachomatis has been cloned and expressed in a few vectors by several groups (Kaul and Wenman, 1985; Palmer and Falkow, 1986b; Bavoil et al., 1986). Kaul and Wenman (1985) isolated the gene from a library of PstI partially digested chlamydial DNA inserted into the vector pBR322. The gene banks were screened with rabbit anti-chlamydial antibodies. One of the purposes of this study was to look at the degree of homology of this gene between different strains of chlamydia. Using a 1.8kb PstI fragment of the 74kd gene, they found extensive homology between serovars L1, L2, J, and K, but only weak homology to the C. psittaci serovar Mn.

The gene encoding the 74kd antigen was also cloned by Palmer and Falkow, (1986b) and Bavoil et al. (1986). These genes were expressed in E. coli and identified with a patient's serum.

The ease with which the 74kd antigen gene is cloned and expressed in a variety of systems, as compared to other genes of chlamydia, implies that it is not lethal to the cell, is not degraded within the cell, and it may contain regulatory elements compatible with E. coli ribosomes. Palmer and Falkow (1986a) have suggested a temporal expression of this gene within C. trachomatis due to the observation of corresponding RNA during early infection only. Therefore, the 74kd antigen gene appears to be under

a different set of control mechanisms from other developmentally regulated genes of chlamydia which appear late in infection.

The ribosomal RNA genes of C. trachomatis have been cloned (Palmer and Falkow, 1986b; Engel and Ganem, 1987). Ribosomal genes are relatively conserved through evolution, therefore analysis of the chlamydial rRNA genes has potential for identification of the evolutionary trends of this organism.

Engel and Ganem (1987) isolated a series of clones from a library of EcoRI partially digested chlamydial DNA inserted into Lambda gtwes. The library was screened with cDNA probes to total chlamydial RNA. These probes were generated with the assumption that most of the RNA present would be ribosomal RNA. Two different cistrons were identified, both containing a single copy of the 16S and 23S rRNA genes. They concluded that the chlamydial genome contains at least two copies of these rRNA genes and that each copy is present as an operon containing one 16S rRNA gene and one 23S rRNA gene under the control of a common promoter.

Palmer and Falkow (1986b) isolated the 16S rRNA gene of C. trachomatis utilizing pure 16S rRNA as a probe. Restriction analysis and Southern hybridizations confirmed that the gene was duplicated within the genome. Approximately 90% of the gene has been sequenced and it shares 70% sequence homology with the 16S rRNA gene sequence

of E. coli. Southern hybridization studies of the 16S rRNA gene to genomic DNA of other bacterial species showed weak homology with C. psittaci, Vibrio cholerae, and Pseudomonas aeruginosa. Using an oligonucleotide complementary to a unique sequence within the 16S rRNA gene for Southern blot hybridization studies, these researchers showed significant homology between trachoma and LGV biovar strains, and slight homology to mouse pneumonitis and C. psittaci strain GPIC (Palmer et al., 1986).

The cryptic plasmids of chlamydial species, first described by Lovett et al. (1980), have been isolated and cloned, (Joseph,et al.,1986; Horn, etal.,1986; Palmer and Falkow, 1986ab; Clarke, 1986). The plasmid of C. trachomatis was determined to be 7.0-7.4kb in size (Joseph et al., 1986; Horn et al., 1986; Palmer and Falkow, 1986ab) while the plasmid of C. psittaci is only 6.2kb in size (Joseph, et al., 1986). Expression studies did not reveal any of the known antigenic protein genes of C. trachomatis or C. psittaci within the plasmid genomes (Joseph et al., 1986; Palmer and Falkow, 1986ab).

B) Promoter and Terminator Elements of Prokaryotes

(i) Transcription Promoters

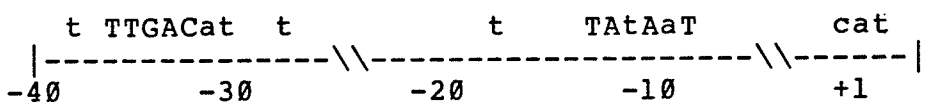
A promoter is a specific region of DNA, upstream from the coding sequence of a gene, which is involved in initiation of transcription by recognition and interaction with RNA polymerase holoenzyme. The specific recognition

and interaction is the first step in initiation of transcription, followed by formation of the open RNA-DNA complex upon release of sigma factor from the holoenzyme (Reznikoff and McClure, 1986).

Promoters of prokaryotes consist of a bipartite structure with regions of sequence conservation centered at -35 and -10 from the transcription start site. The general sequence homology within the -10 region (TATA or Pribnow box) of promoters, first shown by Pribnow (1975ab) and Schaller et al., (1975), consisted of the sequence TATAATG. The more distal upstream region, centered at -35, was also shown to consist of sequence homology (Takanami et al 1976; Seeburg et al 1977). Through the compilation of 112 E. coli promoters, Hawley and McClure (1983) deduced two concencus sequences for the -35 and -10 regions of the promoter. Figure 1 is a schematic representation of the E. coli promoter concencus sequences. Apart from the sequence homology, Hawley and McClure found that the -10 and -35 regions were usually separated by 17 $\pm$ 1 bases. Another region of sequence homology detected, involved three bases centered around the transcription start site, however the degree of homology in this region is low (see Figure 1).

It has been suggested that the -35 region functions in initial recognition by RNA polymerase (Schaller et al., 1975) and the -10 region is involved in melting-in of the polymerase near to the transcription start site. A minor variation to the bipartite function theory suggested by

Figure 1. Concensus Sequences for *E. coli* Promoters



Lower case letters represent those bases which occur in greater than 39% and less than 54% of the promoters examined. Upper case letters represent those bases which occur in greater than 54% of the promoters examined.

Adapted from Hawley and McClure, 1983.

Stefano and Gralla (1982) is that the RNA polymerase complex recognizes, simultaneously, the -35, and -10 sequences, and the space between them. Reznikoff and McClure (1986) have suggested that probably both theories are correct and that a given promoter utilizes one mechanism while another promoter may use the other mechanism for initiation. The route of choice of transcription initiation for a given promoter is based on the efficiency with which that gene must be transcribed within the cell. For example, genes which code for rRNA should have promoters which are strong and efficient in order to produce the large number of transcripts necessary for proper cell functioning; whereas, proteins which are needed only once per generation do not require a strong promoter. The most efficient route for transcription initiation is not known at this time.

Differences in the strength of a promoter may in part come about by variation in the concencus sequences at the minus 10 and -35 regions, however there are many other components involved in promoter function. Some other factors that can act to regulate the rate of transcription include, repressor or initiator proteins, extent of DNA supercoiling, and the presence of or absence of other promoters in close proximity (Reznikoff and McClure, 1986).

(ii) Transcription Termination

The formation of a complete mRNA transcript involves the initiation of transcription at the 5' end of an operon, and subsequent termination of transcription at the 3' end.

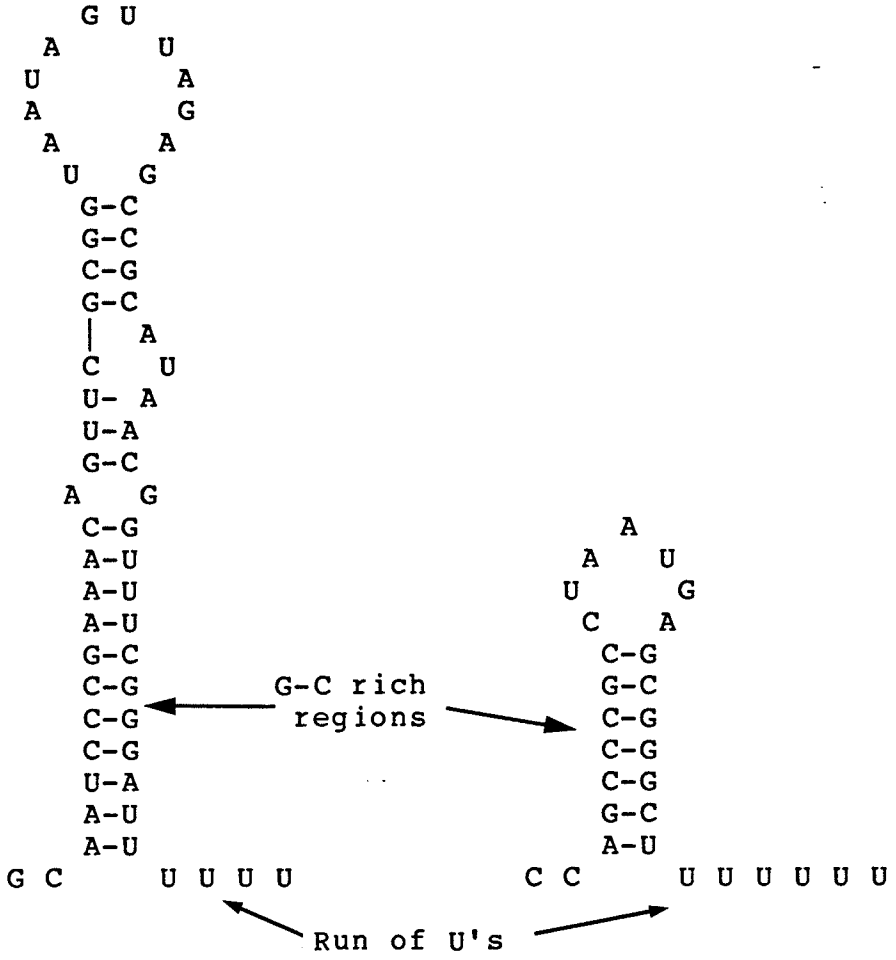
How does the RNA polymerase enzyme know when it has reached the end of the gene and to stop transcribing? The enzyme must recieve some signal which transduces the message: stop transcribing.

Analysis of the primary structure of bacterial mRNA transcripts revealed that many ended with GC-rich regions showing dyad symmetry followed by a series of uridine residues (Adhya and Gottesman, 1978), suggesting that this structure may be involved in transcription termination. This type of terminator, originally termed "simple terminators" by Platt and Bear (1983), are also known as rho-independent terminators, since they do not require rho-factor for termination in vitro. The general structure of rho-independent terminators is shown in Figure 2. It is believed that termination in this situation occurs as a result of formation of a stable intramolecular hairpin by the GC-rich dyad of the transcript, destabilizing the polymerase-template interaction. Dissociation of the transcript and template occurs as a result of the row of uridine residues which produces unstable pairing of rU-dA in RNA-DNA hybrids (Adhya and Gottesman, 1978; Rosenberg and Court, 1979; Platt and Bear, 1983).

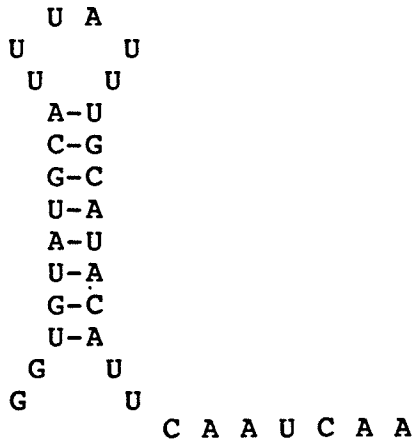
The other type of terminators, which Platt and Bear (1983) termed "complex terminators" , are all terminators which do not fit the criteria of simple terminators. In general, these terminators lack either dyad symmetry or a series of uridine residues and require the addition of an

Figure 2. Prokaryotic Transcription Terminators

Rho-independent terminators:



Rho-dependent terminator:



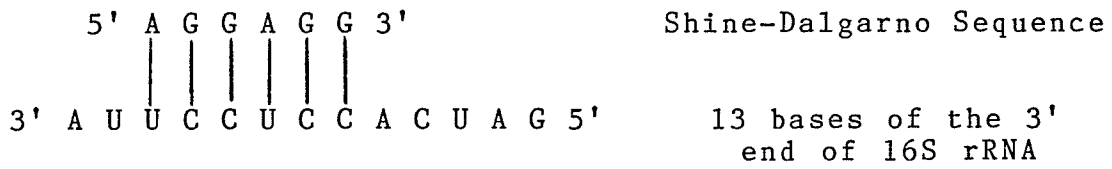
Adapted from Lewin, 1985.

accessory protein (rho-factor; nus-gene products) in order to carry out proper termination in vitro. In 1969, Roberts isolated a protein which he found catalyzed the release of mRNA transcripts and RNA polymerase from the DNA template in vitro. This accessory protein was termed rho-factor. General sequence homology for these terminators has not been found, however one of the typical structures is shown in Figure 2. The mechanism by which these terminators function is believed to involve the dissociation of the transcription complex as a result of the interaction of the accessory protein with RNA polymerase at the terminator structure.

(iv) Translation Initiation and Termination

The idea that ribosomes are intricately involved in the specificity of translation was first put forth by Gorini and Kataja in 1964 when they suggested that "the ribosomal structure could include the accuracy of the reading code during translation." Initial interaction of the ribosome and mRNA occurs with a sequence, upstream of the translation start codon, known as the Shine-Dalgarno sequence or ribosome binding site (RBS). It was first suggested by Shine and Dalgarno (1974) that a conserved sequence in the RBS of transcripts could interact with specific sequences in the 3' end of the 16S rRNA, based on their observation of sequence complementarity between these two regions. Figure 3 shows the thirteen 3' bases of 16S rRNA and their proposed interaction with an optimal RBS. A comparison of 124 E. coli translation initiation site sequences by Stormo and co-

Figure 3. Shine-Dalgarno Sequence and 16S rRNA Complementarity



workers (1982) revealed similarities from which they proposed some rules for typical RBS sequences. Almost all of the sequences had at least a three base complementarity to the 16S rRNA 3' end, within 15 bases of the translation start codon. Complementarity was an average length of 5 $\pm$ 1 bases and was spaced 7 $\pm$ 2 bases upstream of the translation start codon.

The most common intiation codon is AUG, found in 91% of E. coli sequences published as of 1984, (Stormo, 1986). The next most common are GUG at 8%, then UUG at 1%. The initiator tRNA codes for methionine and the anticodon is complementary to AUG; therefore, it is expected that AUG is the most common translation initiation codon.

Termination of translation is also dependent on the codon. There are three possible terminator codons, UAA, UAG, and UGA, which do not correspond to any tRNA and therefore do not code for an amino acid. Two proteins involved in the termination procedure, Release Factor (RF)1 and RF2, function in reading of the termination codons, but their exact mechanism is not yet known.

(iv) Chlamydial Promoters and Ribosome Binding Sites

A total of nine different chlamydial genes have been cloned, and three have had their DNA sequence determined. Of the chlamydial genes sequenced, only MOMP L1 and the rRNA operon have been shown to contain any sequence homology to the -10 and -35 transcription promoter elements of E. coli (Pickett et al., 1987; Engel and Ganem, 1987). The tandem

promoters found in the rRNA operon showed weak homology (45-50%) with typical E. coli promoters (Hawley and McClure, 1983). The chlamydial rRNA promoters were verified in vitro by primer extension and S1 nuclease digestion assays (Engel and Ganem, 1987). As a result of the weak homology to typical prokaryotic promoters, it was suggested that chlamydial promoter sequences are significantly different. In vivo expression studies of the rRNA genes showed that successful transcription initiating at the Proximal (P)2 promoter occurred ten times more often than at the P1 promoter. The differential expression implied that the P2 promoter is stronger or more efficient. This finding is interesting since the putative P2 promoter has the least homology to typical prokaryotic promoters.

The lack of typical promoter elements in chlamydial structural genes might be due, in part, to the presence of these genes within an operon, therefore under the control of a promoter far upstream. This genomic structure has been suggested for the 60kd and 12kd protein genes of chlamydia (Bavoil et al., 1986).

Identical rho-independent transcription terminators following the coding region of MOMP L1 and MOMP L2 genes have been detected (Pickett et al., 1987; Stephens et al., 1986). A terminator was not found following the 18kd antigen gene sequence (Kaul et al., 1987), but this may reflect the structure of an operon and transcription of the gene as a polycistronic mRNA transcript.

Putative Shine-Dalgarno sequences have been detected in upstream sequences of the 18kd protein, MOMP L2, and MOMP L1 genes. The putative Shine-Dalgarno sequence of the 18kd antigen gene consists of the four base sequence AGGA, and is located 10 bases upstream of the start codon (Kaul et al., 1987). The analogous sequence of both published MOMP gene sequences, L1 and L2, are identical, consisting of six bases, AGAGGT, beginning five bases upstream of the start codon (Stephens et al., 1986; Pickett et al., 1987). The observation of somewhat typical Shine-Dalgarno sequences in chlamydial genes is expected and not surprizing due to the extent of sequence and secondary structure conservation between C. trachomatis 16S rRNA and E. coli 16S rRNA (Palmer and Falkow, 1986b).

The apparent divergence of promoter element sequences supports the notion that chlamydial transcription may be initiated by a slightly different mechanism than E. coli transcription, ie: specific nucleotide interaction of chlamydial promoter sequences and chlamydial RNA polymerase. However, Palmer and Falkow (1986b) have reported good expression of chlamydial genes with in vitro transcription/ translation systems from E. coli, indicating that E. coli RNA polymerase and ribosomes are capable of interacting with chlamydial regulatory elements. Perhaps, the differences in promoter sequences reflect differences in accessory proteins involved in regulation of these mechanisms in chlamydial species.

The terminators of chlamydia transcription appear to be typical, but this is not surprizing since termination operates on a mechanical basis and is not wholly dependent upon sequence homology. Terminator structure shows that some chlamydial genes terminate by the rho-independent mechanism.

Materials and Methods

A) Growth and Maintenance of Cells

(i) HeLa 229 and Chlamydia trachomatis cells

Chlamydia trachomatis $UW_3E_{10}H33$ or serovar D and HeLa 229 cells were supplied by C.C.Kuo and S.P.Wang of the Department of Pathology, University of Washington, Seattle, Washington, U.S.A. The C. trachomatis strain was propagated in monolayer HeLa 229 cells as described by Jenkin, 1966, and Kuo et al., 1977.

HeLa 229 cells were grown in the presence of Eagle's Minimal Essential Medium (MEM) (Appendix A), supplemented with 10% fetal calf serum (Sigma), in 175cm^2 NUNC flasks at 37°C. Twenty-four hour confluent monolayers were treated for 20 mins., at room temperature with 3 mls of Hank's Balanced Salt Solution (HBSS) containing DEAE-dextran (30mg/ml) (Pharmacia). The HBSS was removed and 1ml of C. trachomatis stock containing $2x10^8$ Inclusion Forming Units (IFU's)/ml, (or $2x10^9$ particles/ml), was added. The flasks were left flat at room temperature for 2 hours to allow adhesion of the elementary bodies to the cells of the monolayer. The inoculum was then discarded and the monolayer was rinsed with a small volume of HBSS. Approximately 50 to 70 mls of MEM containing cycloheximide (1mg/ml), to inhibit cellular growth, was added to each flask and the flasks were incubated at 35°C for 72 hours.

If the cells in the flask had reached 80-100%

infectivity, the chlamydial elementary bodies were harvested for stocking or for purification. If the infection was less than 80%, the EB's were harvested and passed to more flasks in order to obtain efficient infection of the monolayer. Harvesting was carried out as follows. The growth medium was discarded, 10 mls of cold HBSS and approximately 30 glass beads (2mm) were added to each flask. The infected cells were removed by gently rolling the beads over the monolayer. The cell suspension was removed and placed into a centrifuge tube on ice. The flasks and beads were rinsed with a second 10 mls volume of HBSS. The cell suspension was sonicated with a Branson Sonifier cell disruptor 185 at a probe intensity of 35, for 20 seconds. The suspension was centrifuged at 500xg for 10 min at 4°C.

For the preparation of frozen stock cultures, the supernatant was removed and centrifuged at 30,000xg for 30 min at 4°C. The pellet was resuspended in 8 mls of sucrose, phosphate, glutamate (SPG) solution (Appendix A), using a blunt-ended spinal tap needle. The chlamydial suspensions were distributed into 1 ml aliquots and stored at -70°C.

For the preparation of pure elementary bodies, the supernatant recovered from centrifugation at 500xg was layered over 8 mls of 35% RENOgrafin (Squibb), and centrifuged at 43,000xg for 60 min at 4C in a SW27 rotor (Beckman). The pellet was resuspended in 10 mls of HBSS and layered over a discontinuous RENOgrafin gradient consisting of 5 ml of 52%, 8 ml of 44%, and 13 ml of 40% RENOgrafin in

0.01M HEPES (Gibco), 0.15M NaCl. The gradient was centrifuged at 50,000xg for 90 min at 4°C in an SW27 rotor. The elementary bodies were recovered from the 44%-52% interface. This suspension was diluted with 3 volumes of SPG and centrifuged at 30,000xg for 30 min at 4°C. The pellet was resuspended in SPG and centrifuged at 30,000xg for 30 min. at 4°C. The pellet was then resuspended in 1ml SPG for each original infected flask and stored at -70°C.

(ii) Escherichia coli Strains

The various E. coli strains used are listed in Table I. They were grown at 37°C in liquid media or on agar plates. The strains JM101 or JM109, used for propagating M13 phage, were initially grown on minimal media agar plates for isolation of single colonies. A single colony was picked and used to inocculate 2 mls of either Luria Bertani (LB) or 2x Yeast Tryptone (YT) liquid media (Appendix A) and incubated with vigorous shaking. Other strains were initially grown on either LB or 2YT agar plates before inocculation of liquid media.

The strains Y1089(r-)/Y1090(r-) were used for propagating the Lambda gt11 vector and recombinants. When lysogenized strains of E. coli Y1089(r-) harbouring a recombinant lambda genome were grown, the cells were incubated at 32°C in order to retain the lysogenic state.

The DH5-alpha strain of E. coli was used for propagating plasmids and plasmid recombinants.

TABLE I: E. COLI STRAINS

STRAIN	GENOTYPE	REFERENCE
Y1089(r-)	lacU169 proA$^+$ lon araD139 strA hflA [chr::Tn10] (pMC9)	Young and Davis (1983b)
Y1090(r-)	lacU169 proA$^+$ lon araD139 strA supF [trpC22::Tn10](pMC9)	Young and Davis (1983b)
DH5-alpha	DH1 (argF-laczya)U169 80d lacz M15	Bethesda Research Labs
JM101	supE thi (lac-proAB) [F'traD36 proAB lacI Z M15]	Yanisch-Perron et al., (1985)
JM109	recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 - (lac-proAB) [F'traD36 proAB lacI Z M15]	Yanisch-Perron et al., (1985)

B) DNA Extractions

(i) Chlamydial DNA

The procedure used for chlamydial DNA extraction was described by McClenaghan et al. (1984). Purified EB's from 12 flasks were pelleted by centrifugation and resuspended in 10 mls of 10mM Tris-HCl pH 7.5, 1mM EDTA (TE) buffer (all solutions and buffers used for DNA extractions are described in Appendix B). Five mls of chlamydial lysis solution was added and the mixture was incubated at 55°C for 15 min, then 37°C for 45 min. The lysate was extracted with phenol-saturated 10mM Tris-HCl, pH 7.5, 1mM EDTA, 10mM NaCl (TES). The aqueous phase was retained and re-extracted once with phenol:TES:chloroform, (1:1:1, v/v) and once with chloroform:isoamyl alcohol, (24:1, v/v). Nucleic acid was precipitated from solution by adding 1/10 the final volume of 3M ammonium acetate, twice the volume of 95% ethanol (EtOH), and holding at -20°C overnight or at -70°C for half an hour. Following centrifugation of 30,000xg at 4°C for 30 min, the DNA pellet was resuspended in an aqueous solution of 6% cesium chloride with 0.4ml of 10mg/ml ethidium bromide (EtBr). Centrifugation was carried out in a Beckman Ti70 rotor at 60,000 rpm for 18 hours at 25°C in a Beckman L8-70M ultracentrifuge. This procedure is known as dye-buoyant-density gradient centrifugation. The chromosomal band was visualized under ultraviolet (UV) light and removed. The EtBr was removed from the DNAby adding four times the solution volume of twice glass distilled H_2O (ddH_2O), and

extracting four times with TE/CsCl-saturated isobutanol. The nucleic acid was precipitated in the same manner as described above. After a 45min centrifugation, at 7700xg, the DNA pellet was resuspended in TE, precipitated, pelleted, and suspended in ddH_2O.

(ii) Lambda DNA

Lambda DNA was prepared by precipitating phage particles from a bacterial lysate and subsequently extracting with phenol. The lysates were prepared following the plate-lysate method of Maniatis et al. (1982). The lysate preparation is described in a later section. Precipitation of lambda phage was accomplished by centrifugation of the lysate at 20,000 rpm for 2 hrs at 4°C in a Ti60 rotor (Beckman). The supernatant was discarded and 0.2 ml of SM buffer was added to resuspend the pellet overnight at 4°C. Both DNase I (Boerhinger Manneheim) (5ul of 1mg/ml aqueous solution) and RNase A (Boerhinger Manneheim) (10ul of 10mg/ml aqueous solution) were added to the suspension and incubated at 37°C for 1 hr. Following incubation, Proteinase K (Sigma) to a final concentration of 1mg/ml and 10% Sodium Dodecyl Sulfate (SDS) to a final concentration of 0.1% were added. This mixture was incubated at 65°C for 20 min, followed with the addition of 0.25M EDTA to a final concentration of 5mM and incubated for an additional hour at 65°C. The solutions were extracted twice with phenol and once with phenol-chloroform. The DNA was pelleted as described for chlamydial DNA.

(iii) Plasmid Nucleic Acid

a) Large Scale Preparation

A 500 ml overnight culture of E. coli containing the plasmid of interest was centrifuged at 1500xg for 10 min. The pellet was resuspended in 2 mls of lysis solution and placed on ice for 20 min. A 2.4 ml volume of Triton X-100 (Sigma) solution (Appendix B) was added and the cell suspension was again left on ice for 15 min, or until air bubbles were seen indicating that lysis was complete. The mixture was centrifuged at 17,000xg for 15 min to remove the bacterial debris. The supernatant was placed into a clean 30 ml Corex tube and 0.9g of CsCl was added to each gram of supernatant. The solution was heated at 65°C for 15 min or until the CsCl dissolved. The solution was cooled to room temperature and then placed on ice for approximately 30 min. The solution was spun at 17,000xg for 15 min and the supernatant added to 0.1 ml of 10mg/ml EtBr solution in polyallomer tubes. The plasmid DNA was separated by dye-bouyant-density gradient centrifugation as described above. The plasmid DNA band was removed from the tube and purified as described for chlamydial DNA isolation.

b)Rapid DNA Mini-Preparation

This procedure was used for the preparation of small quantities of plasmid DNA or the replicative form (double stranded DNA) of M13 bacteriophage. This procedure has been adapted from Maniatis et al., (1982). A single colony or phage particles from a single plaque was used to

inoculate 1.5 ml of LB or 2YT broth containing ampicillin (50ug/ml, final concentration). The culture was grown for 12-16 hours, shaking at 37°C and then centrifuged for 5 mins at 14,000xg. If single stranded phage DNA was desired, the supernatant was stored at 4°C with 0.1 ml of chloroform. The pellet was resuspended in 0.1 mls of 50mM glucose, 10mM EDTA, and 25mM Tris-HCl, pH 7.5, with freshly prepared lysozyme (Sigma) at a final concentration of 2mg/ml. After holding on ice for 5 min, 0.2ml of a freshly prepared solution of 0.2M NaOH, 1.0% SDS was added. The solution was mixed and placed on ice for another 5 minutes. Then 150 ul of a 5M potassium acetate solution was added to the mixture. The solution was vortexed gently, and held on ice for an additional 10 minutes. After centrifugation for 10 minutes at 14,000xg, the supernatant was removed and 350 ul of ice cold isopropanol was added to the supernatant to precipitate the DNA. After 15 minutes on ice, then subsequent centrifugation at 14,000xg, the pellet was washed with cold 70% EtOH. The pellet was dried and then dissolved in 50-100 ul of TE. If desired, the solution was treated with RNase A (50ug/ml, final concentration) at room temperature for 1 hour. The RNase A was removed by phenol/chloroform extraction followed by EtOH precipitation of the DNA.

The plasmid "quick screen" is a crude method of rapid plasmid preparation that was used for screening large numbers of deletion mutants. The procedure utilizes alkali, at an elevated temperature for cell lysis. A large

proportion of a colony was picked and suspended into a 100 ul volume of alkali solution (50mM NaOH, 0.5% SDS, 5mM EDTA) in a microtitre well. The microtitre plate was incubated at 65°C for 1 hour, after which time an equal volume of loading dye (Appendix D) for agarose gel electrophoresis was added. Each sample was loaded onto an agarose gel and subjected to electrophoretic separation, described in detail in a later section. When separation was complete, the plasmid bands were visualized in the presence of EtBr by shortwave UV light. Plasmids in the desired size range can be crudely identified in this manner by comparison to the parent plasmid.

(iv) Single Stranded M13 DNA

M13 is a filamentous male specific coliphage. The virion is composed of a single stranded circular DNA genome in a protein coat. The single stranded genome has been modified to facilitate cloning and sequencing. The M13mp18 and M13mp19 series described by Norrander et al. (1983) were used.

To prepare single stranded DNA, infected cells were grown according to the method of Messing (1983) with some modifications. A 10 ul volume of the supernatant, saved from the double stranded mini-preps, and 10 ul of JM109 (or JM101) cells in exponential phase, were inoculated into 100 mls of 2YT broth. The culture was grown overnight at 37°C. The cells were pelleted by centrifugation at 17,000xg for 10 minutes and the supernatant was removed carefully with a

pipette. For every 100 mls of supernatant, 20mls of a solution of 20% polyethylene glycol (PEG), 2.5M NaCl, was added. The mixture was inverted gently and left at room temperature for 15 minutes to allow the λphage to precipitate. Phage was pelleted by centrifugation at 17,000xg for 10 minutes, and the supernatant was poured off. In order to remove as much of the PEG as possible, the tubes were inverted for 10 minutes and the insides were wiped out with a sterile cotton swab, gently so that the pellet was not disturbed. The pellet was resuspended in 2.4 mls TES for each 100 mls volume of supernatant, and extracted twice with phenol/chloroform and once with chloroform. The single stranded DNA was EtOH precipitated, pelleted by centrifugation, dried, and resuspended in 100 ul TE for each 25 ml volume of original supernatant. This DNA was used for preparing deletions and for sequencing.

C) Determination of DNA Concentrations

The concentration of nucleic acid in aqueous solution was determined spectrophotometrically, as described by Maniatis et al.(1982). Nucleic acid absorbs light at a wavelength of 260nm and protein absorbs light at 280nm. Dilutions of the nucleic acid solution were made in either ddH_2O or TE buffer. The absorption of light at 260nm was measured and the nucleic acid concentration could be calculated from a conversion equation. Nucleic acid in the form of double stranded DNA in solution at a concentration

of 50ug/ml will yield an optical density (OD) of 1. Measuring the OD at 280nm provides an estimate of the solution's purity. A ratio of OD 260nm/OD 280nm of 1.8 represents a pure preparation of DNA.

D) Restriction Endonuclease Digestion

(i) Digestion

Restriction endonucleases are DNA modifying enzymes which cut specific palindromic sequences in double stranded DNA. They were used to generate fragments of DNA for cloning or subcloning, and to identify locations of restriction sites on a given DNA fragment ie: create a restriction map. The method of producing a restriction map is outlined by Maniatis et al. (1982). Breifly, a given DNA fragment is digested with a series of restriction enzymes, individually and in combinations. The resultant fragments are sized by electrophoresis through agarose gels and comparison to a known standard. Restriction sites are assigned based on the size of DNA fragments.

Restriction enzyme digestion of DNA with a variety of enzymes was done following the manufacturer's specifications. Table II lists the restriction enzymes used, their recognition sequence, the buffer in which reactions were carried out and the optimal reaction temperature. All buffer solutions are described in Appendix C.

TABLE II: RESTRICTION ENZYMES

ENZYME	RECOGNITION SEQUENCE	BUFFER(1)/TEMP
AsuII	T T\C G A A	LOW/37oC
AvaI	C\Py C G Pu G	MEDIUM/37oC
AvaII	G\G T(A) C C	MEDIUM/37oC
BamHI	G\G A T C C	MEDIUM/37oC
BglII	A\G A T C T	LOW/37oC
BstEII	G\G T N A C C	MEDIUM/60oC
ClaI	A T\C G A T	ClaI/37oC
EcoRI	G\A A T T C	HIGH/37oC
HindIII	A\A G C T T	MEDIUM/37oC
KpnI	G G T A C/C	MEDIUM/37oC
PstI	C T G C A/G	MEDIUM/37oC
PvuII	C A G|C T G	MEDIUM/37oC
SacI	G A G C T/C	LOW/37oC
SalI	G\T C G A C	HIGH/37oC
Sau3AI	\G A T C	MEDIUM/37oC
SmaI	C C C|G G G	SmaI/37oC
SphI	G C A T G/C	HIGH/37oC
XbaI	T\C T A G A	HIGH/37oC

(1) High, med, and low refers to salt concentration; all buffer components are listed in Appendix C

(ii)Agarose Gel Electrophoresis

Based on size, and conformation, DNA can be separated through a gel matrix of agarose by an electrical current. Electrophoresis was done as described by Maniatis et al. (1982), with some minor modifications. The procedure of DNA electrophoresis was carried out in agarose gels which ranged from 0.5% to 1.0% agarose in ELFO buffer (Appendix D) at room temperature. All electrophoresis buffers and solutions are described in Appendix D. Ethidium bromide was added to the buffer and gel mixture at a final concentration of 0.5 ug/ml, to allow visualization of the DNA bands in the gel when exposed to a UV light source. Pictures of the gels were taken with Polaroid Type 57 film.

Sometimes a specific fragment of DNA from a digestion is desired for subcloning or for use as a probe. This fragment can be isolated by electrophoresing through low melting temperature (LMT) agarose and excising the fragment from the gel. Gels of 1.0% agarose (Seaplaque) in ELFO buffer were used and electrophoresis was done at a low voltage (ca. 25v.) overnight and often at 4°C. If the DNA fragment was desired for use as a probe, it had to be purified out of the agarose. This was achieved with the Schleicher and Schuell ELUTIP-d columns. Purification was performed as described by the manufacturer, with the exception that filters were not used.

(iii) Subcloning and Ligation

If a particular DNA fragment was desired for subcloning

into a vector, the fragment was left in the agarose and the procedure for ligations in LMT agarose described by Struhl (1985) was used. Briefly, electrophoresis and isolation of the DNA was done as described above. The gel slice was trimmed so that only agarose containing DNA was left. The remaining slice containing a known amount of DNA is weighed and the assumption is made that 1.0mg=1.0ul of gel slice. The concentration of the DNA in the slice can then be calculated with the following equations:

$$\text{I:}\quad \text{pmols DNA} = \text{ug DNA} \times \frac{1 \times 10^{12}}{\#\text{base pairs} \times 660 \times 10^{6}}$$

$$\text{II:}\quad \text{concentration} = \frac{\#\text{pmols DNA}}{\#\text{ul of gel slice}}$$

Gel slices containing the DNA fragments to be ligated were melted at 70°C for 5 minutes and combined in optimal proportions for ligation (Tegerski and Robberson, 1985). At 37°C, 1ul 10x ligation buffer, and 1ul ligase enzyme are added and the final volume brought to 10ul with ddH_2O. The buffers and solutions for use of all DNA modifying enzymes are listed in Appendix E. The solution is mixed and incubated at 14°C overnight. This procedure was also used for ligation of DNA not in LMT agarose.

If forced orientation cloning is not being used the ends of the vector must be dephosphorylated to increase the efficiency of ligation of recombinant molecules, and decrease the background of non-recombinant molecules.

Dephosphorylation was carried out as follows. The vector, linearized with the appropriate enzyme, was incubated with Bacterial Alakaline Phosphatase (BAP) (International Biotechnology Inc., IBI) at 65°C for 1hour as described by the manufacturer. The enzyme was inactivated by the addition of Proteinase K and incubation at 37°C for 30 mins. The solution was phenol extracted and the DNA was EtOH precipitated. The vector is then ready for ligation. The various vectors used are presented in Table III.

a) Construction of Lambda Library

a.1) Lambda gt11 Cloning System

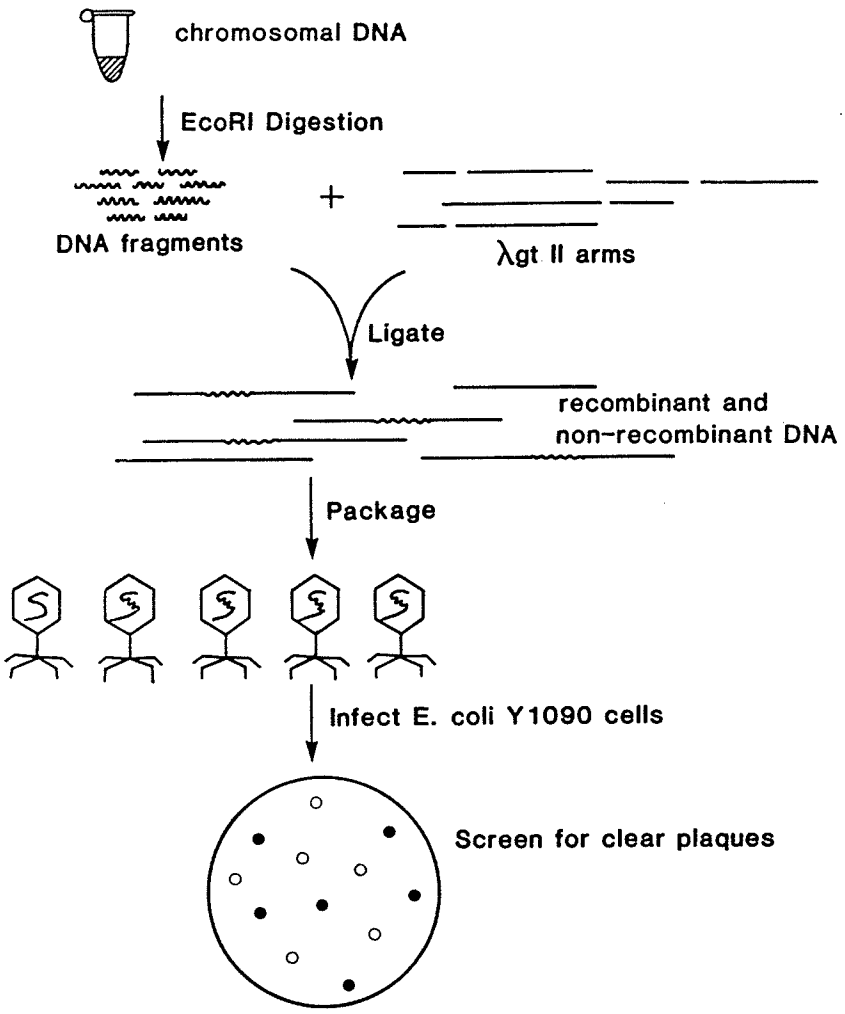
Bacteriophage lambda is a double stranded DNA virus which contains a large genome of 50kb in size. Approximately one third of the genome is nonessential for lytic growth of the phage; therefore, it can be replaced with large quantities of foreign DNA (Maniatis et al., 1982). The lambda vector, gt11, was described by Young and Davis (1983a). The general scheme of constructing a Lambda gt11 library is shown in Figure 4.

The Lambda gt11 cloning system was purchased from Promega Biotec, and construction of the library was carried out according to the manufacturers suggestions. C. trachomatis DNA was digested to completion with the restriction enzyme EcoRI, and ligated to the vector arms. Packaging of the recombinant molecules into phage heads and tails was done using the Packagene kit, purchased from Promega Biotec. The procedure followed was described by the

TABLE III: VECTORS

VECTOR	REFERENCE
Lambda gtll	Young and Davis (1983a)
M13mpl8/M13mpl9	Norrander et al. (1983)
pUCl8/pUCl9	Norrander et al. (1983)
pWR59Ø-1	Guo et al. (1984)

Figure 4. Strategy of Cloning in Lambda gt11



manufacturer.

a.2) Transfections

Transfection of the packaged recombinants was also done as described by the manufacturer. The recombinant phage were pre-adsorbed to E. coli Y1090(r-) (the lytic host) for 20 mins at room temperature. Melted top-agar containing $MgCl_2$, isopropyl B-D-thiogalactopyranoside (IPTG), and 5-bromo-4-chloro-3-indolyl-B-D-galactopyranoside (X-gal), was added to the transfections, poured onto LB plates and incubated overnight at 42°C. Recombinants were identified as clear plaques within the lawn of bacteria. Recombinant plaques were picked and replated until pure.

The E. coli host strain Y1089(r-) is a high frequency lysogenization strain used for the generation of lysogens where the lambda genome becomes integrated into the host chromosome. This strain was used for generating lysogens of the lamda recombinants so that large quantities of the recombinant protein could be produced for SDS-PAGE analysis. Transfection of the Y1089(r-) host was done in the same manner as described for the Y1090(r-) host, with the exception that the plated infection was incubated at 32°C to retain the lysogenized state, a shift to 42°C will induce the lytic cycle.

a.3) Lysate Preparation

Recombinant phage were produced in large quantity by the plate lysate method of Maniatis et al., (1982). Breifly, phage were pre-adsorbed to a compatible bacterial

strain and poured onto plates in the same manner as described for transfections. The plates were incubated at 42°C until lysis was confluent. To each plate, 5 ml of SM buffer was added and incubated at 4°C for a few hours. The SM buffer, containing phage particles, was removed and the plates were rinsed with an additional 5 ml aliquot of SM buffer which was added to the first harvest. In order to remove bacterial contamination, 0.1 ml of chloroform was added to the phage suspension, the suspension was vortexed, and centrifuged briefly to pellet cellular debris. The recovered supernatant was stored at 4°C, with 0.3% chloroform.

b) Plasmid Constructions

b.1) Transformations

A transformation is the method by which a DNA molecule is inserted into a compatible host bacterium. It was used for plasmid and M13 dsDNA molecules. The host bacterium must first be rendered competent by treatment with ice-cold $CaCl_2$, as described by Maniatis et al., (1982). A single plate colony of JM101 or JM109 (for M13 dsDNA) or DH5-alpha (for plasmid DNA) was inoculated into 50 mls of 2YT broth and incubated at 37°C, shaking vigorously, until the culture reached an OD_{600nm} of 0.15-0.2. The culture was placed on ice for 10 mins and then pelleted at 3,000xg for 10 min at 4°C. The pellet was resuspended in ca. 25 mls of ice-cold 0.1M $CaCl_2$ solution and re-incubated on ice for 20 mins. Following this incubation period, the cells were pelleted,

resuspended in 2 to 5 mls of 0.1M $CaCl_2$, and stored on ice until used. Competent cells were either made fresh for each transformation or bought from Pharmacia.

For the transformation, 0.2ml to 0.5ml of the competent cell suspension was used. DNA of interest was added to the cells and the mixture was chilled on ice for 20 to 60 minutes. The reaction was heated at 42°C for 2 mins, to induce heat shock in the cells. Heat shocking allows for a more efficient transformation. Following heat shock, the reaction was replaced on ice. For M13 vectors and recombinants, 3 mls of soft 2YT agar medium containing X-gal (50ul of a 2.0% solution in N,N-dimethyl formamide) and IPTG (20ul of a 2.0% aqueous solution) was added. The reaction was mixed and poured over 2YT plates. For plasmids, 1 ml of 2YT broth was added after heat shock and incubated for 1 hour at 37°C. The suspensions, divided into three different volumes (700ul, 200ul, and 100ul) were poured onto 2YT plates containing ampicillin (50ug/ml), X-gal and IPTG. The plates were incubated for 12-16 hours at 37°C.

E) Plaque and Colony Immunoblot Screening

(i) Plaque Immunoblotting

Recombinant molecules were screened for genes of interest by the immunoblot screening method described by Stephens et al. (1985). Nitrocellulose discs, saturated with IPTG, were overlaid on the plates and marked for orientation. For lambda plaque screening, the plates with

the discs were incubated for an additional 2 hrs at 37°C. The discs were removed, washed with phosphate buffered saline (PBS), and blocked with PBS containing 5% bovine serum albumin (BSA), at 37°C for 1 hour. The discs were then treated with a pool of monoclonal antibodies with specificities for outer membrane antigens of C. trachomatis. Table IV lists the monoclonal antibodies used and thier specificities. The antibodies were diluted 1:2000 in Tris-buffered saline (TBS), 0.05% Tween 20, and incubated with the discs for 1 hour at 37°C. The discs were washed three times with TBS-Tween 20, shaking at 37°C, for 10 mins each wash. The horseradish peroxidase-conjugated anti-mouse IgG antibody (Cedarlane Laboratories, Hornby, Ontario) was diluted 1:2000 in TBS-Tween and reacted with the discs in the same manner as described above. Again, the discs were washed three times and subsequently developed in a solution containing 0.05ug 3,3'-diaminobenzidine/ml and 0.003% H_2O_2 in PBS. Positive plaques were identified as those which developed a small purplish-brown spot on the nitrocellulose discs.

The positive recombinants were picked, purified and subsequently screened individually with each monoclonal antibody, to identify which antigen they were expressing.

(ii) Colony Immunoblotting

In screening lysogenized strains or plasmid libraries, a similar method of immunoblotting was used, (Young and Davis, 1983a). Whole cells were transferred onto

TABLE IV: MONOCLONAL ANTIBODIES

ANTIBODY DESIGNATION	ANTIGEN SPECIFICITY
2.F2.D3	LPS (1)
2.H9.G5	29KD OMP (2)
4.H10.B6	32KD OMP
5.D5.F6	MOMP
3F2	60KD OMP
14-E6	75KD OMP
7A8-G12	70KD OMP
7-C9-C6	57KD OMP

(1) LPS = lipopolysaccharide

(2) OMP = outer membrane protein

nitrocellulose by overlaying the plates with discs and marking them for orientation. The discs were replated and incubated for 1 hr at 37°C (lambda lysogens were incubated at 42°C). The cells were lysed in a $CHCl_3$ atmosphere for 15 mins. The nitrocellulose filters were placed in 3 mls of Buffer A (0.17M NaCl, 0.01M Tris-HCl, pH 7.5, 0.1mM phenylmethylsulfonyl fluoride) with 0.01% SDS for 1 hour at room temperature. The filters were incubated for an additional 10 mins in Buffer A with 2ug/ml DNase I, and rinsed with Buffer A. Again, to reduce non-specific antibody binding, the filters were incubated for 1 hour in Buffer A containing 3% BSA. Monoclonal antibodies were added, at a dilution of 1:2000, in Buffer B (Buffer A with 0.1% SDS, 0.1% Triton X-100, 1mM EDTA) for 3hrs, gently shaking. The filters were washed three times, 10 mins each, with Buffer B and reacted with the anti-mouse IgG conjugate at a dilution of 1:2000 in Buffer B for 3 hrs, gently shaking. Again, the filters were washed three times with Buffer B and subsequently developed as described for the plaque immunoblots. Alternatively, following cell lysis, the nitrocellulose filters were treated in the same manner as western blots (described later).

Positive colonies were isolated, purified, and the DNA was isolated for subsequent analysis.

F) Colony DNA Hybridizations

Screening of plasmid libraries was also done by DNA-DNA

hybridization of transformed colonies with a DNA probe nick translated with alpha-^{32}P-dATP (New England Nuclear, NEN).

(i) Preparation of Transformant DNA

The procedure used for colony DNA hybridization was essentially that of Grunstein and Hogness (1975) with some modifications. Nitrocellulose discs were blotted onto plates containing transformants, marked for orientation and removed. The discs were replated on media containing chloramphenicol, at 37°C, for 2 hrs, to increase the plasmid copy number. The colonies were lysed and the DNA was immobilized on the nitrocellulose. The nitrocellulose discs were air dried and baked at 80°C for 2 hrs, or not baked at all when using Gene Screen Plus nitrocellulose (BIO-RAD). The entire procedure is outlined in Maniatis et al. (1982) for in situ hybridization of bacterial colonies.

(ii) Nick Translation of Probes

Radioactive labelling of double stranded DNA probes was accomplished using the method described by Maniatis et al. (1982) with one modification. The process involves creating multiple nicks within the DNA molecule with DNase I and subsequently repairing these nicks with one labelled and three unlabelled nucleotides and DNA polymerase I (Kornberg fragment).

Usually, about 200 ng of DNA was labelled per translation. The DNA along with the nucleotides were incubated in nick translation buffer (Appendix E) with ca. 0.004 units of DNase I at 15°C for 10 mins. Following the

nicking reaction, DNA polymerase I (Kornberg) (Pharmacia) was added and the mixture was incubated at 15°C for at least 1 hr at which time the reaction was stopped with the addition of TES containing bromophenol blue dextran. Labelled DNA was separated from unincorporated nucleotides by passage through a Sephadex G-50 column and isolation of the faint blue band.

G) Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

(i) Electrophoresis

Separation of proteins based on their molecular weight was accomplished using the method of SDS polyacrylamide gel electrophoresis (SDS-PAGE) described by Laemlli, (1970). All buffers and solutions used are described in Appendix D. Electrophoresis was carried out in gels of 10% and 12.5% acrylamide. Protein samples were solubilized by adding an equal volume of solubilizing buffer and boiling for 5min. The solubilized samples were added to the wells within the stacking gel and an electrical current of 20m amperes per gel was applied to the system. Electrophoresis was generally carried out for 4 to 6 hours, until the bromophenol blue dye front neared the bottom of the gel.

(ii) Silver Staining

Following electrophoretic separation, gels destined for silver staining were carefully removed from the glass plates and the upper right hand corner was snipped for orientation purposes. The method of silver staining used was the

modified method of Morrisey, (1981). Briefly, gels were prefixed in methanol:acetic acid:ddH_2O, (50%:10%:40%, v/v), by agitating for 30 min. This is followed by a second prefix reaction in methanol:acetic acid:ddH_2O, (5%:7%:88%, v/v) for another 30min. The gel was subsequently soaked in 10% glutaraldehyde for 30 min, rinsed in ddH_2O and left soaking overnight in a large volume of ddH_2O. The next day, the gel was soaked for 30 min. in fresh ddH_2O. The gel was shaken for 30 min. in an aqueous solution containing 0.5mg dithiothreitol (DTT). The DTT solution was poured off and silver nitrate solution (0.1% in ddH_2O) was added, and the gel was again shaken for 30min. The gel was rinsed three times, quickly, with small volumes of water each time. The developer was added (50ul of 37% formaldehyde in 100ml of 3% sodium carbonate solution) until bands reach the desired intensity, at which point the developer was poured off and the gel was rinsed many times with water.

(iii) Western Blotting

Western blotting involves the electrophoretic transfer of proteins in a gel matrix onto nitrocellulose membranes. The procedure followed was an adaption of the method described by Towbin et al., (1979). The gel, recovered following electrophoresis, was soaked in Blot buffer for 1hr to remove SDS present in the gel. The gel was carefully laid over a nitrocellulose membrane cut to fit the size of the gel and sandwiched between a layer of 3MM Whatman paper and two sponges. The electrophoretic apparatus used was the

Trans-blot system purchased from BIO-RAD. Electrophoresis was carried out in Blot buffer, overnight, at 0.2 amps, followed by an increase to 0.3 amps for 1hr, the next day.

The nitrocellulose membranes were probed with antibody in essentially the same manner as described for plaque immunoblots.

H) DNA Sequencing

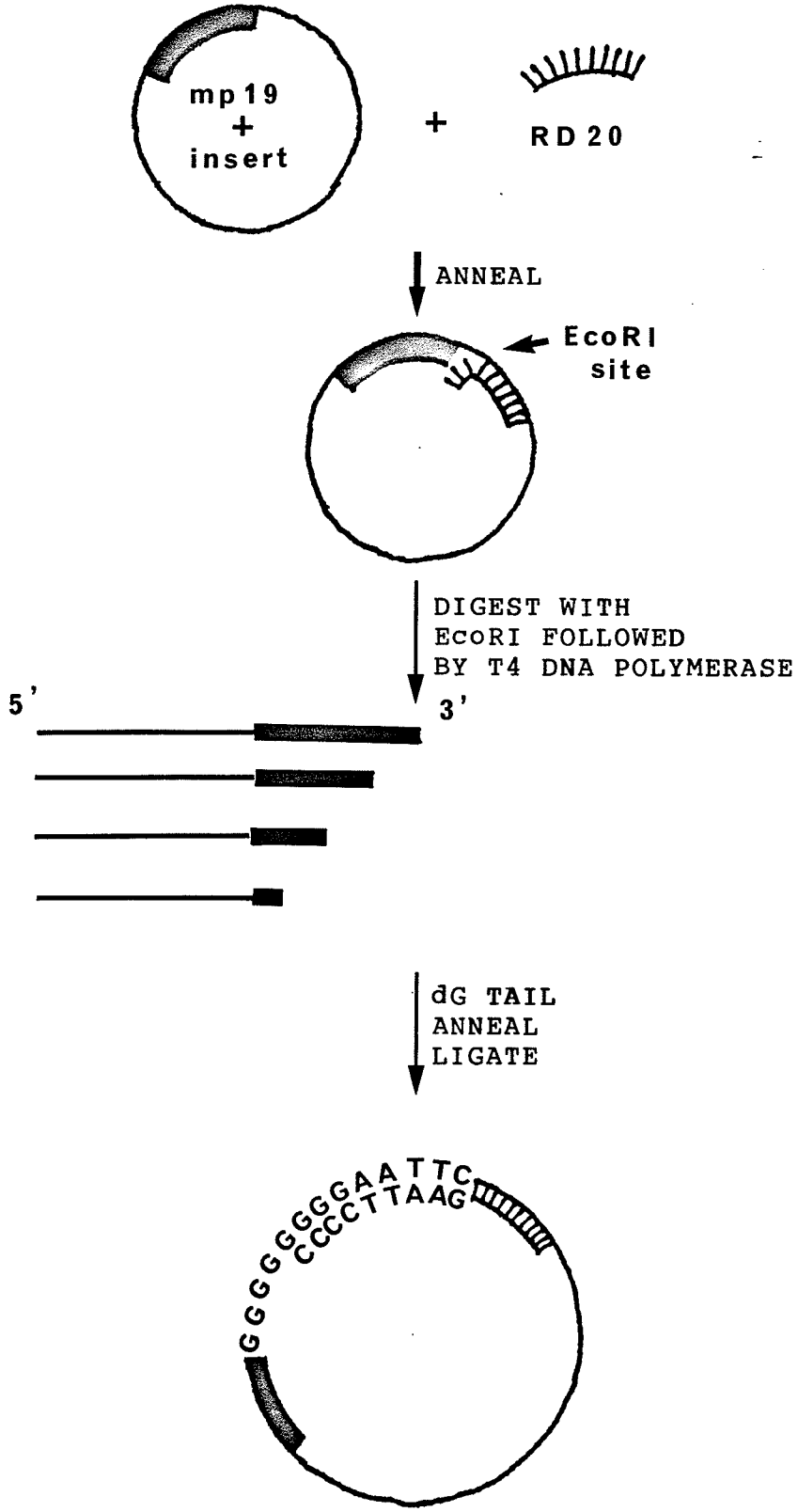
(i) T4 DNA Polymerase Deletion

Deletions of inserts within M13 vectors can be created using the enzyme T4 DNA Polymerase, as described by Dale et al., (1985). The basic strategy of this procedure using an M13mp18 recombinant is shown in Figure 5. A 20 base oligonucleotide, RD 20 (IBI), is annealed to the single stranded recombinant. This is followed by linearization with EcoRI. The linearized molecule is then subjected to digestion with T4 DNA Polymerase (Pharmacia) which has specific 3' to 5' exonuclease activity on single stranded DNA. Digestions were stopped at various time-points to create a variety of deletions. A homopolymeric tail of guanine residues is added to the 3' end with the enzyme terminal transferase (BRL) and dGTP. The linear molecules are ligated and transformed into competent JM101 or JM109 cells. Double stranded DNA from recombinant plaques is then screened for desired deletions.

(ii) Exonuclease III Deletion

Another method for creating deletions in recombinant

Figure 5. Creating Deletions with T4 DNA Polymerase



Adapted from Dale et al., 1985.

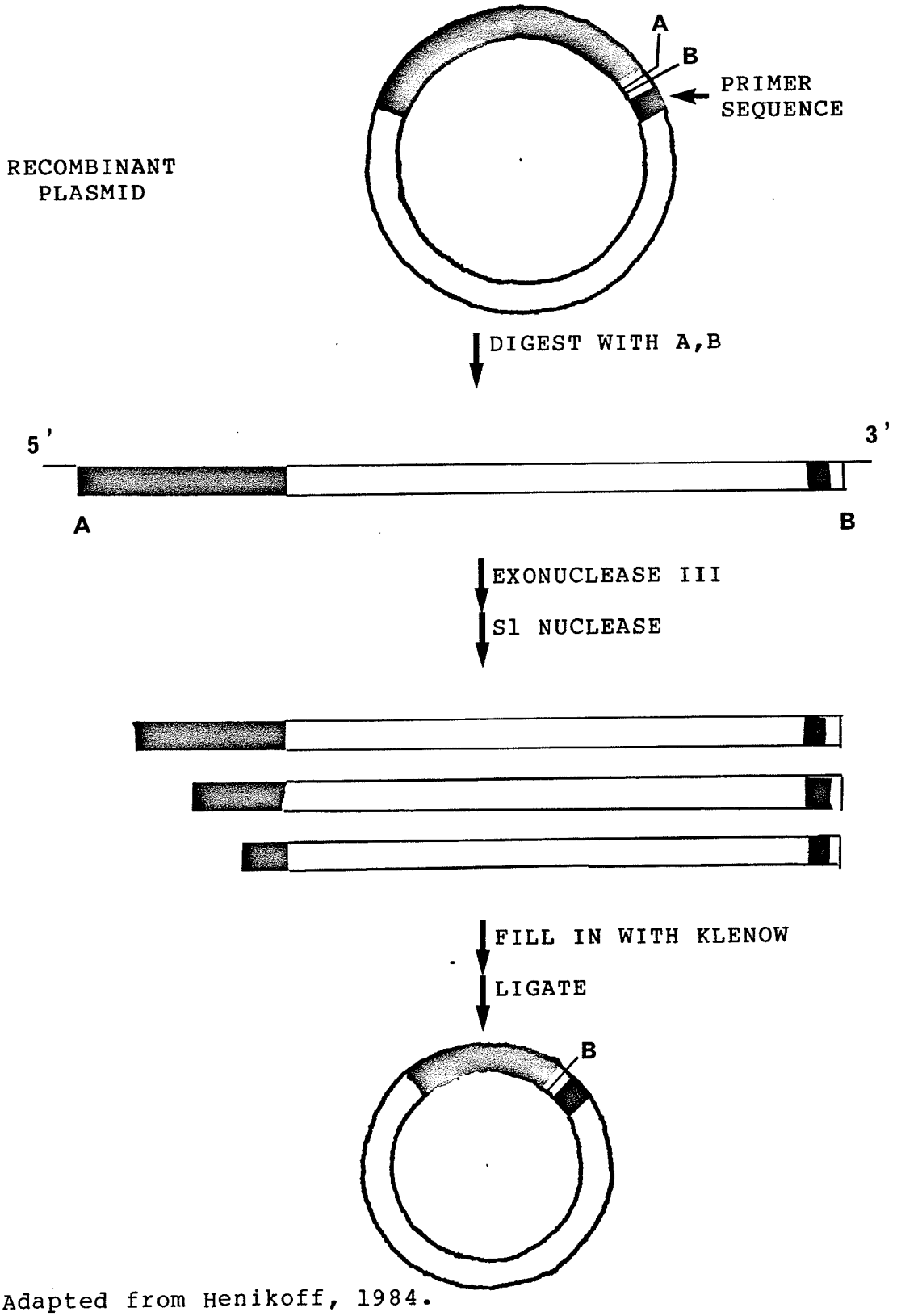
molecules utilizes the enzymes Exonuclease III and S1 nuclease (Guo et al., 1983; Henikoff, 1984). This procedure creates deletions in double stranded plasmids and is outlined in Figure 6. Briefly, the plasmid must be double digested with one endonuclease which creates a 5' overhang (A) and one which creates a 3' overhang (B). The enzyme Exonuclease III has specific single stranded 3' to 5' exonuclease activity on double stranded molecules with a 5' overhang. The linearized plasmid is incubated with Exonuclease III (BRL) and removed at various time points to create a variety of deletion lengths. The deletions are treated with S1 nuclease (BRL) to remove the resultant 5' overhangs, and then with Klenow polymerase I (in the same manner as nick translations) to fill-in any nicks that might have been created by S1 nuclease, and the 3' overhang. The molecules are blunt-end ligated and transformed into competent cells. The plasmids of the resultant deletions were screened by the quick screen method described earlier.

(iii) Single Stranded DNA Sequencing

The method of single stranded dideoxy-sequencing of DNA by Sanger et al. (1977), is the most widely used procedure for sequencing DNA. Various manufacturers have produced kits with all the reagents necessary for carrying out the dideoxy-sequencing method.

The M13/Klenow sequencing kit from BIO-RAD, and the M13/Sequenase kit from IBI were used. The BIO-RAD kit follows the standard dideoxy-sequencing procedures with the

Figure 6. Creating Deletions with Exonuclease III



Adapted from Henikoff, 1984.

use of the large fragment of DNA polymerase I, the Klenow fragment. The IBI kit consists of an novel enzyme, Sequenase, which replaces the function of the Klenow fragment, ie: synthesis of DNA. Sequenase is a recombinant molecule derived from bacteriophage T7 DNA polymerase with modifications to improve sequencing ability (Tabor el al., 1987). The benefits of this enzyme over the Klenow fragment, as claimed by the manufacturer, include, high speed of reaction, low background, and the ability to obtain twice as much sequence than with reactions using the Klenow fragment. The Sequenase procedure is based on the concept of dideoxy-sequencing, but uses slight modifications in reagent concentrations from the Sanger et al. (1977) method. In general, it requires much lower concentrations of the nucleotide reagents.

The procedures used for sequencing of single stranded DNA recombinants in M13 were those described by the manufacturers of the two kits discussed above.

(iv) Double Stranded DNA Sequencing

Sequencing of double stranded DNA is sometimes desirable since it eliminates the need for subcloning into a single stranded phage. Some sequences, such as bacterial promoters, highly repeated sequences, large inverted repeats, and inserts over 1000 nucleotides, cannot be stably cloned into M13 vectors (Chen and Seeburg, 1985). If the DNA to be sequenced is located within a plasmid vector containing sequences homologous to the universal

oligonucleotide primers, double stranded sequencing can be done. If homologous sequences to the universal primers are not present within the vector, a primer can be generated; however, it would be much simpler in this case to subclone into a vector with homologous sequences.

Sequencing can be done with either double stranded phage or plasmid DNA. The method used was essentially that of Chen and Seeburg (1985). Briefly, double stranded DNA, obtained as described earlier, was denatured by alkali, resulting in covalently closed, circular, single stranded molecules. The solution was neutralized and sequencing can be carried out as previously described (Sanger et al., 1977). Sequencing was done using the BIO-RAD kit and the procedure outlined by that manufacturer, with the exception of the annealing reaction and subsequent incubation temperatures. The annealing reaction must be reduced to 37°C, to prevent rehybridization of the double strands, and subsequent incubation steps were done at 37°C, instead of room temperature. The only other difference between the procedure used and that of Chen and Seeburg (1985) was that sequencing reactions were not dessicated and resuspended in a smaller volume prior to electrophoresis. The reactions were loaded directly into the gels in 3ul quantities.

(v) Electrophoresis and Autoradiography

Sequenced fragments are separated by electophoresis through denaturing polyacrylamide gels. The combination of urea and high voltage (generating high temperatures), act as

the DNA denaturants. Maintenance of the sequenced DNA in a denatured state is essential to retain single strands during electrophoresis. The procedures followed for preparation of reagents, apparatus, pouring of gels, and electrophoresis of sequenced DNA were as described by the sequencing gel apparatus manufacturers, BRL and BIO-RAD.

Electrophoresis was generally carried out in 0.4mm thick gels of 6% polyacrylamide, 8.3M urea, and a temperature of 50-60°C. The gels were loaded with 3ul of each reaction mix, between the teeth of a sharkstooth comb. When the gel apparatus of 50cm (BIO-RAD) in length was used, electrophoresis for 3hrs was needed for the bromophenol blue dye front to reach the bottom. If longer sequence data was desired, the gel could be run for an additional 3 hrs, with sequential loading of the reactions for shorter sequence data. All buffers and solutions are described in Appendix D.

At the end of electrophoresis, the apparatus must be taken apart very gently so that the gel does not become stretched or damaged. The gels can be transferred to Whatman 3MM paper for a protective support and then dried under a vacuum. If sequencing was done with a ^{32}P isotope, drying is not necessary, and old X-ray films can be used for a gel support.

Autoradiography was carried out in a metal cassette, with the gel and X-ray film held in close contact. The X-ray film used was either Kodak XAR-5 or X-Omat RP. Exposure

was usually for about 12hrs, at which time the film was developed through a developing machine (Health Sciences Center, X-ray Department, Winnipeg, Canada).

Autoradiograms are read from the bottom to the top. Sequence is read starting with the first band at the bottom and continuing up the ladder with bands in each lane representing the base corresponding to the reaction mixture of that lane. The sequence data is read in the 5' to 3' direction from the primer, and represents the complementary sequence of the template DNA.

I) Computer Analysis of DNA Sequences

The DNA sequences were read and analyzed with the aid of computer programs. Reading of autoradiograms was done utilizing the Pustell DNA Sequence Analysis (IBI) program and the University of Wisconsin BESTFIT program. These programs were used to find overlapping regions between the various sequences and homology with other chlamydial and bacterial sequences. The University of Wisconsin PEPLOT, and FOLD programs were used to analyse the amino acid sequence and mRNA secondary structure, respectively.

Results

A) Genomic Library of Chlamydia DNA in Lambda gt11

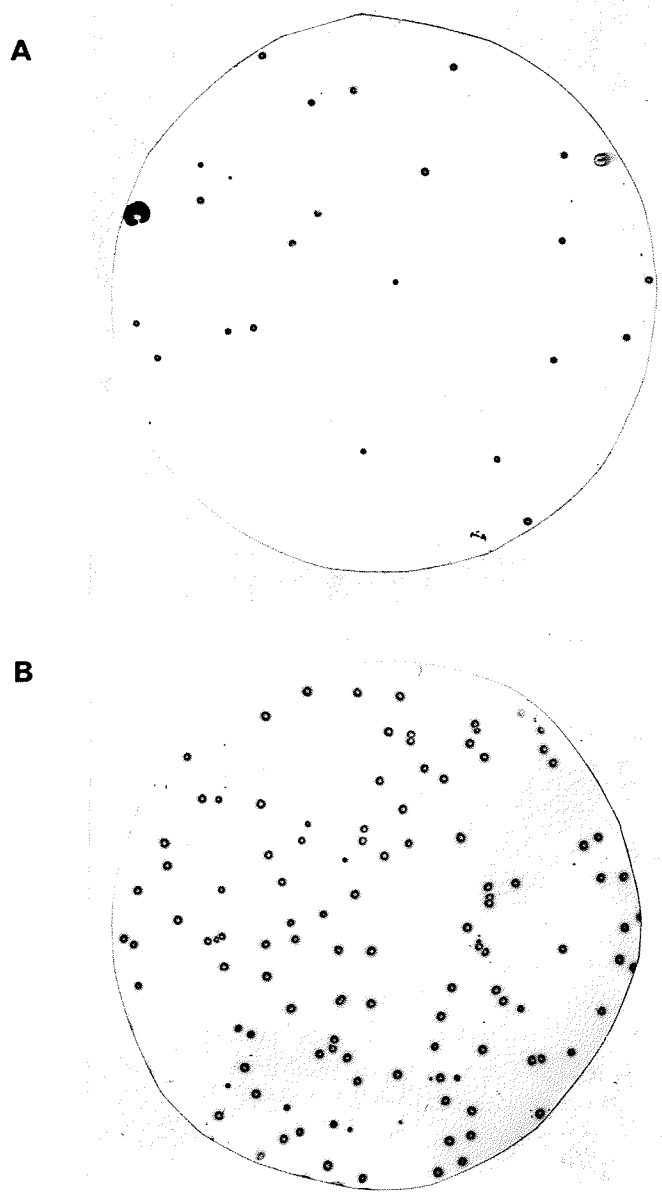
(i) Recombinants of the Lambda gt11 Library

Chlamydial chromosomal DNA, digested to completion with Eco R1, was inserted into the vector Lambda gt11. The recombinant molecules were packaged into phage heads and tails and adsorbed onto *E. coli* Y1090(r-). Recombinants were detected as clear plaques, in the presence of X-gal, indicating that the expression of B-galactosidase has been interrupted by insertion of foreign DNA within the gene.

Five clear plaques (EDR-1 to 5) were picked and amplified in order to screen the recombinants for those harbouring a gene encoding one of the immunologically relevant antigens of *C. trachomatis*. The immunoblot method was used for screening the recombinants with the series of monoclonal antibodies listed, along with their specificity, in Table III. Only one of the monoclonal antibodies reacted with the phage recombinants screened. The antibody, 14-E6, has specificity for an epitope of the 75Kd outer membrane protein of *C. trachomatis*. This result suggested that the EDR recombinants expressed a protein containing an epitope of the 75Kd outer membrane protein (omp). Figure 7 consists of immunoblots screened for recombinant phage.

Lambda phage can be induced to integrate its genome into the chromosome of a susceptible host, thereby entering the lysogenized state. This method was used to produce large amounts of the recombinant fusion protein for SDS-PAGE

Figure 7. Immunoblots of Lambda Library



A: Immunoblot of lambda library plaques; screened with pool of monoclonal antibodies.

B: Immunoblot of purified EDR recombinant; screened with monoclonal antibody 14-E6.

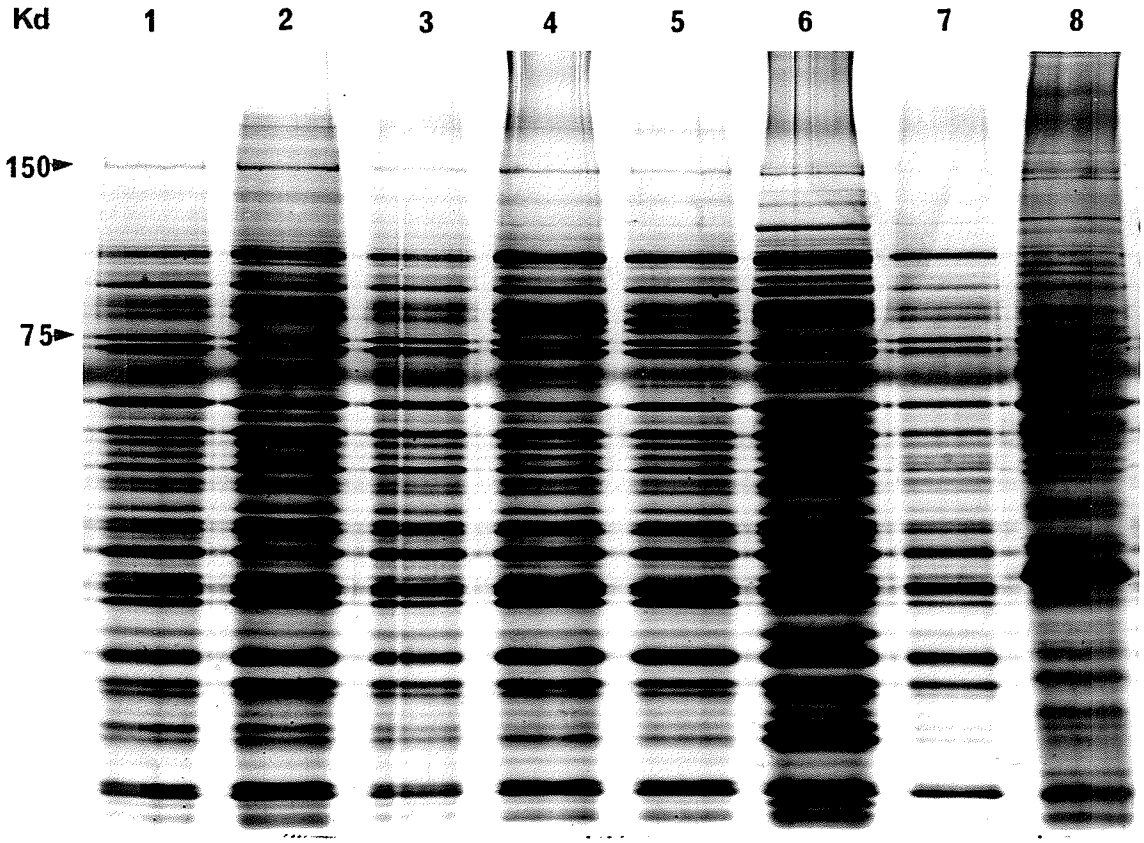
analysis. Small quantities of overnight cultures of the lysogenized recombinants were added to equal volumes of solubilizing buffer and boiled. These whole cell extracts were electrophoresed in 12.5% SDS polyacrylamide gel, in duplicate. The silver stained gel and the Western blot are shown in Figure 8. Production of a novel protein by the lysogenized strains was observed in the silver stain (Figure 8A, lanes 1-5). This protein was not present in cell extracts of either wild type E. coli Y1089(r-) (lane 7) or E. coli lysogenized with the non-recombinant phage (lane 6). The novel protein migrated through the gel to a size of 150 Kd. The novel protein reacted with the monoclonal antibody 14-E6 in the western blot of the same gel (Figure 8B, lanes 1-5). The monoclonal antibody 14-E6 did not react with proteins of either E. coli Y1089(r-) (lane 7) or E. coli lysogenized with the non-recombinant phage (lane 6). A large difference in size was observed between the recombinant protein and the native antigen of C. trachomatis (lane 8).

Restriction analysis of the recombinant EDR phage DNA revealed an insert of approximately 2300 nucleotides, as shown in Figure 9. The recombinants were digested with the endonuclease EcoRI.

(ii) Restriction Mapping and Orientation Determination

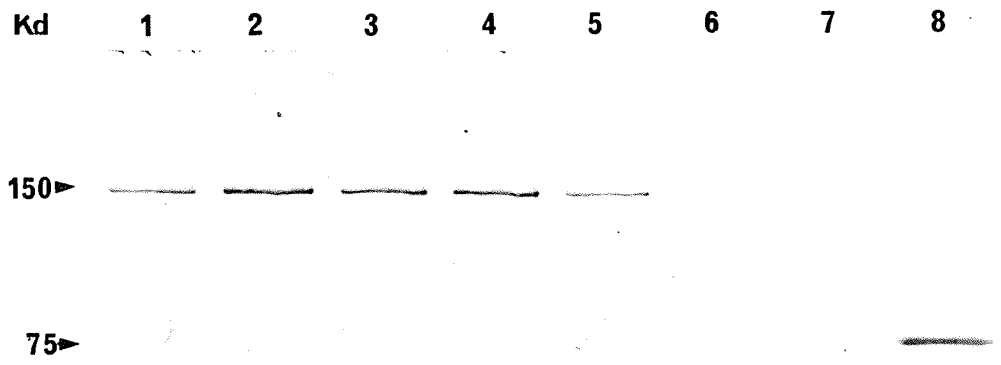
The 2.3kb fragment was subcloned into the EcoRI site of the vectors pWR590-1, and pUC19, in order to determine the direction of transcription and generate a restriction map,

Figure 8. 12.5% SDS-PAGE Analysis of EDR Recombinants



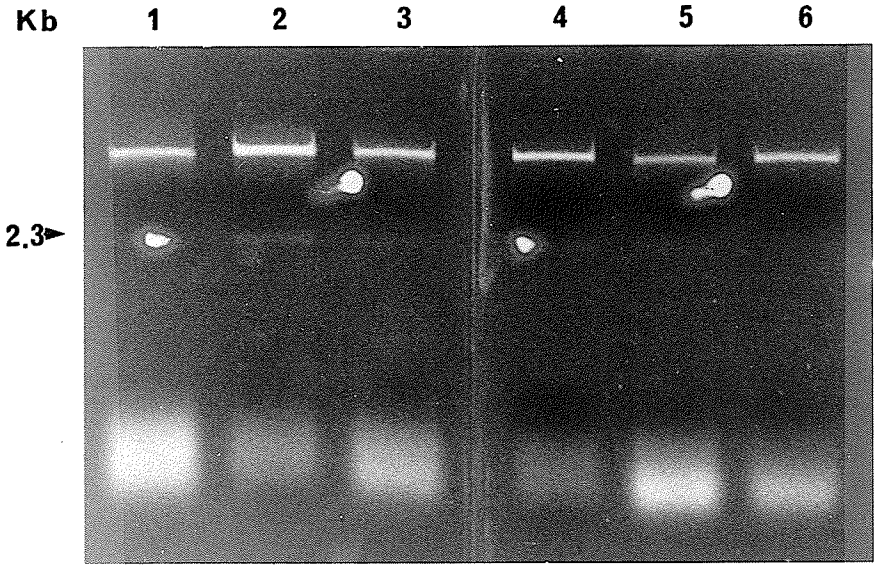
A: Silver stained gel; lanes 1 to 5:lysogens of the EDR recombinants, arrow points out the 150kd fusion protein; lane 6:lambda gt11 lysogen; lane 7: Y1089(r-); lane 8: C. trachomatis, arrow points out the 75kd antigen.

Figure 8. 12.5% SDS-PAGE Analysis of EDR Recombinants



B: Western blot, probed with monoclonal antibody 14-E6; lanes 1 to 5:lysogens of the EDR recombinants, arrow points out the 150kd fusion protein; lane 6:lambda gt11 lysogen; lane 7: Y1089(r-); lane 8: C. trachomatis, arrow points out the 75kd antigen.

Figure 9. Size Determination of EDR Recombinants



Lanes 1 to 5: Recombinants EDR-1 to EDR-5, digested with EcoRI; lane 6: lambda gtll, linearized with EcoRI.

Arrow points to the ca. 2300 nucleotide fragment.

respectively. Restriction analysis was carried out on the recombinant plasmids, pERW-1, pERW-8, pEU4-2, pEU3-1, and pEU6-1. Plasmids pERW-1 and pERW-8, were derived from pWR590-1. Plasmids pEU4-2, pEU3-1, and pEU6-1, were derived from pUC19. The recombinant pERW-1 was immunoblot positive, while pERW-8 was immunoblot negative.

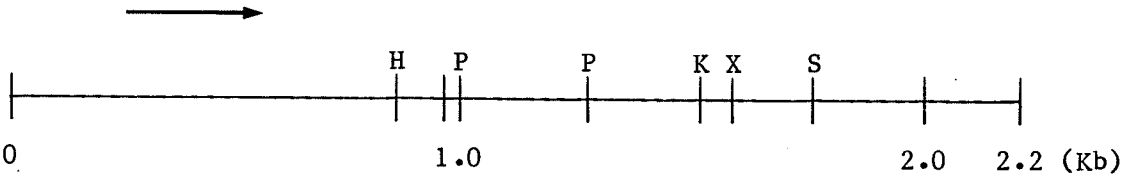
The analysis revealed that the insert was closer to 2.2kb in length and consisted of the restriction sites shown in Figure 10. The arrow above the map indicates the correct direction of transcription of the gene, as determined by the orientation of the restriction sites in plasmids ERW-1 and ERW-8. Restriction enzymes SmaI, BamHI, and SphI did not cut within the insert.

B) Search for complete gene of the 75Kd protein

(i) Construction of a Sau3AI partial genomic library

In order to locate the complete gene encoding the 75Kd antigen, a library containing random, partially digested chlamydial genomic DNA was prepared and screened. The restriction enzyme Sau3AI, cuts double stranded DNA at the four base-pair palindrome GATC. Because this sequence is short, a macromolecule of DNA will contain the palindrome approximately once every 4^4 = 256 bases. A partial digest, limited in either time or enzyme concentration, would result in a relatively random digest of the DNA, generating fragments of various size. Optimal conditions for generating DNA fragments of the size range desired can be determined by creating a series of experiments varying

Figure 10. Restriction Map of the 2.2 Kb Eco RI Fragment



H = Hind III
P = Pst I
K = Kpn I
X = Kba I
S = Sac I

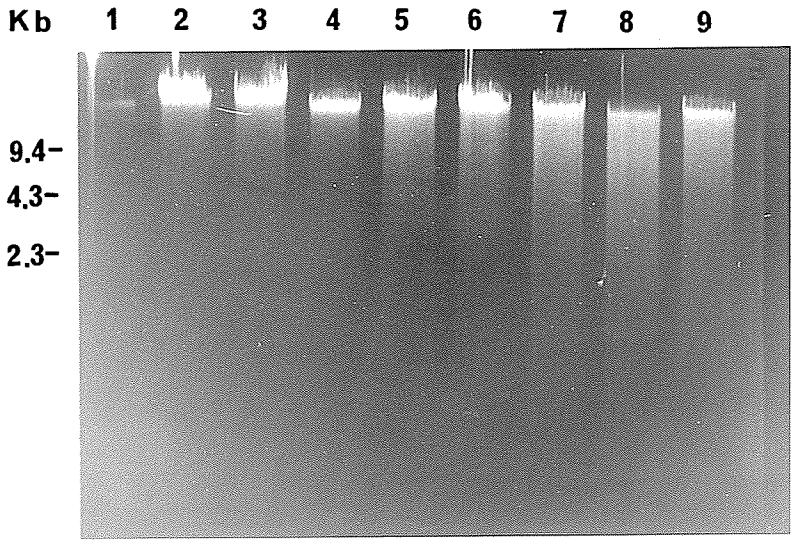
enzyme concentration or reaction time.

A series of Sau3AI digestions in which the concentration of enzyme added was varied and the time of reaction kept constant was carried out with chlamydial genomic DNA. The optimal conditions for the partial digest was considered to be that which generated the most fragments from 5 to 7.5kb. The length of the 75Kd protein gene is expected to be about 2kb. A random digest generating fragments of 5 to 7.5kb should produce fragments which contain the gene, possibly including its regulatory elements.

The Sau3AI concentration was varied from 0.047 to 0.236 units/ug DNA. The concentration of DNA in the reaction mixtures was kept constant at 0.1059 mg/ml. The length of reaction time was kept constant at 60 mins. at 37°C. The results of the various Sau3AI digests are shown in the agarose gel of Figure 11. It was determined that the optimal condition for generating the most fragments in the size range of 5 to 7.5kb was a Sau3AI enzyme concentration of 0.12 to 0.16 units /ug DNA.

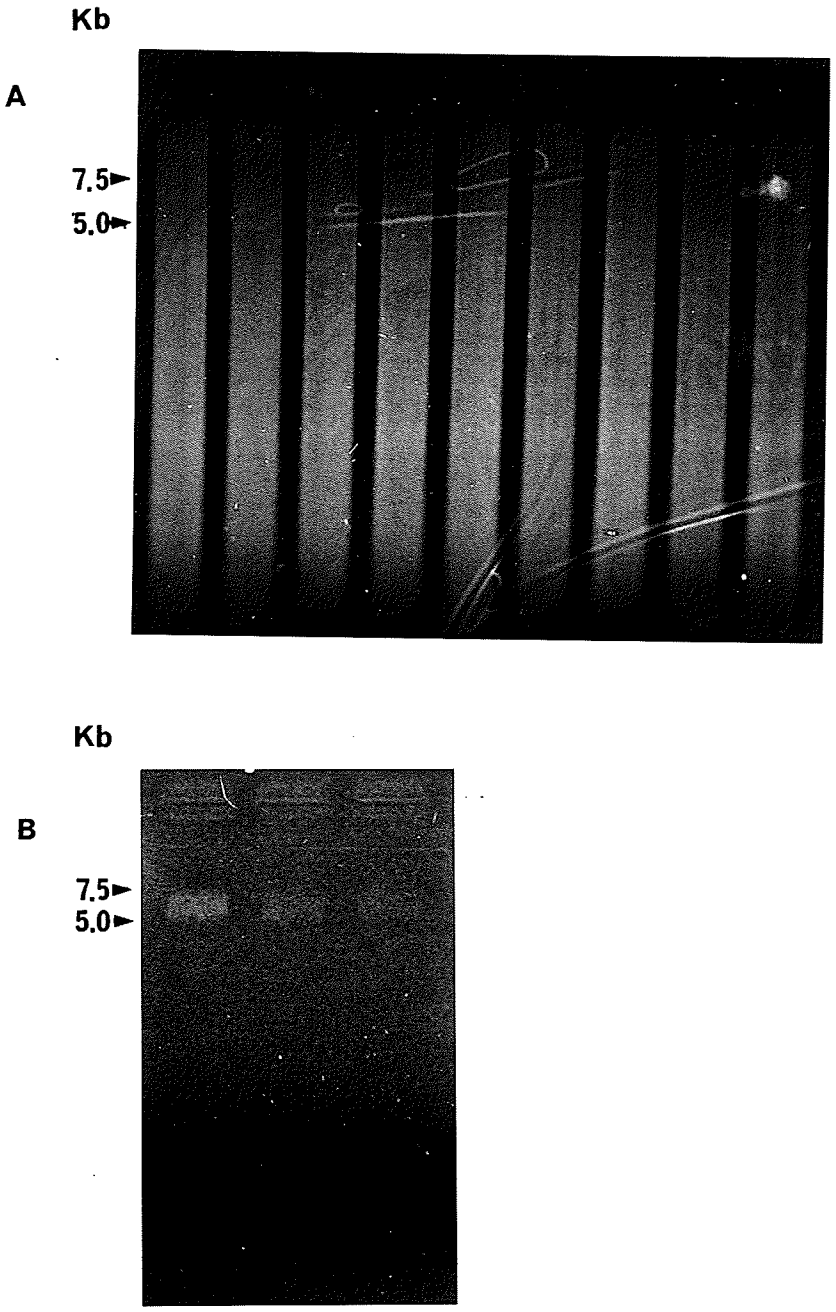
Three conditions (0.12, 0.144, and 0.16 units/ug) from the Sau3AI digest were used for a large scale digestion of chlamydial DNA (25 ug). The reaction time was kept at 60 mins and a temperature of 37°C. The digested DNA was combined and electrophoresed in low melting temperature agarose. The DNA fragments in the size range of 5 to 7.5kb were cut out and purified from the gel. Figure 12 shows the

Figure 11. Digestion with Varying Concentrations of Sau3A1



lane 1: uncut Chlamydial genomic DNA
lane 2: 0.047 units/ug DNA
lane 3: 0.075 units/ug DNA
lane 4: 0.100 units/ug DNA
lane 5: 0.120 units/ug DNA
lane 6: 0.144 units/ug DNA
lane 7: 0.160 units/ug DNA
lane 8: 0.200 units/ug DNA
lane 9: 0.236 units/ug DNA

Figure 12. Large Scale Digest with Sau3Al



A: All lanes: Combined Sau3Al digests; arrows point out the 5 to 7.5kb region that was recovered from the gel.

B: All lanes: The purified DNA of 5 to 7.5kb recovered from the gel.

combined digestions after electrophoresis through low melting temperature agarose, and the resultant purified DNA. The region removed from the gel is indicated by the arrows. The purified DNA fragments were ligated to the BamHI site (compatible sticky ends to Sau3AI) of the vector pUC18 and transformed into competent E. coli DH5-alpha cells.

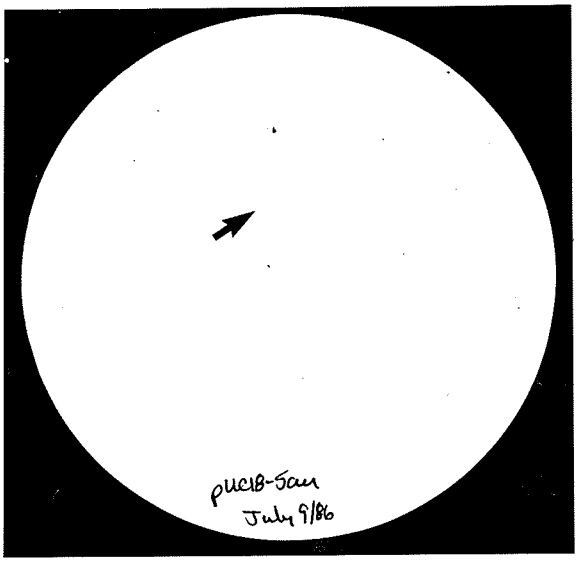
(ii) Identification of pERU-S1

The transformed colonies obtained were initially screened by immunoblot, using the monoclonal antibody 14-E6. This resulted in the identification of only one recombinant, ERU-S1, expressing the 75Kd protein. Figure 13 shows the immunoblot with the one positive reaction.

The transformants were also screened by colony DNA hybridization using the EcoRI fragment of the lambda recombinant EDR4-2 as the probe. The probe hybridized to the recombinant pERU-S1, as well as other recombinants which were negative by immunoblot assay.

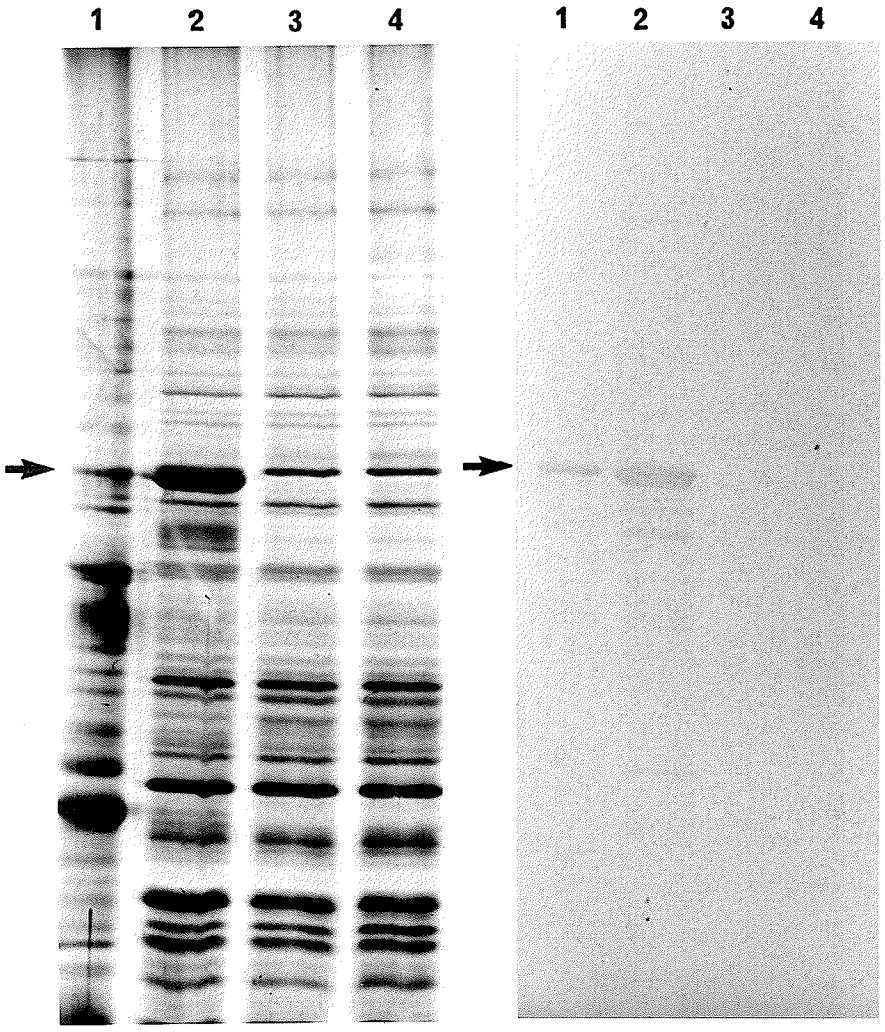
Protein profiles of the transformant ERU-S1 were analyzed by SDS-PAGE. Figure 14 shows the silver stain and western blot of the same gels. A novel protein of approximately 75Kd is present in the protein profile (Figure 14A) of the transformant (lane 2) while this protein is absent in both the wild type E. coli DH5-alpha (lane 3) and also E. coli DH5-alpha transformed with pUC18 (lane 4). The novel protein expressed by pERU-S1 co-migrated with the 75Kd protein of C. trachomatis (lane 1), which is pointed out by an arrow in Figure 14A.

Figure 13. Immunoblot of Plasmid Library



The arrow points to the one positive reaction. This clone was named pERU-S1.

Figure 14. 10% SDS-PAGE Analysis of pERU-S1



A: Silver Stain
B: Western Blot, probed with monoclonal antibody 14-E6.

For both figures A and B; lane 1: C. trachomatis; lane 2: pERU-S1 transformant; lane 3: E. coli DH5-alpha; lane 4: pUC18 transformant. Arrow points to the 75kd antigen.

Figure 14B contains a western blot of the protein profiles discussed in the previous paragraph. The novel protein expressed by pERU-S1 (lane 2) reacts specifically with the monoclonal antibody 14-E6. Again, no reaction has occurred in any of the other lanes except those containing chlamydial proteins. The protein which reacted co-migrated with the 75Kd protein of C. trachomatis (lane 1).

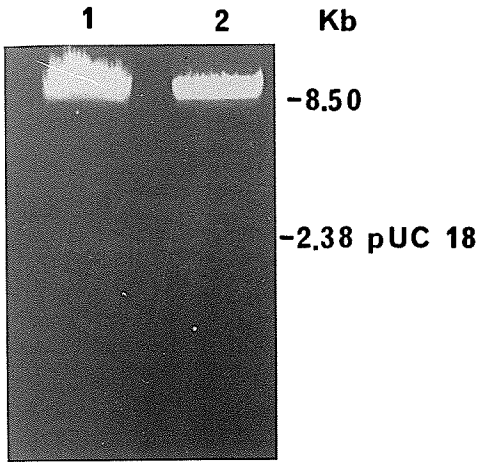
(iii) DNA Analysis of pERU-S1

Plasmid DNA of pERU-S1 was isolated by CsCl dye-buoyant-density gradient centrifugation. The size of the insert expressing the 75Kd protein was determined by digesting the plasmid with the restriction enzymes SalI and SmaI, independently. Both enzymes cut the recombinant molecule once within the insert. The fragment was resolved in 0.7% agarose and sized by comparison to HindIII digested lambda DNA fragments (Figure 15). The insert size was determined by subtracting the known length of the vector pUC18 (2.38kb) from the size of pERU-S1. The size of pERU-S1 was determined to be 8.5kb in length. Therefore the resulting insert size was 6.1kb in length. This insert is of sufficient size to contain the 75Kd protein gene. A limited restriction enzyme study of pERU-S1 was done. The restriction enzyme map generated from these digestions is diagrammed in Figure 16A.

(iv) Construction of pERU-S2

With the aid of the restriction map, the 75Kd gene was localized to a 2.9kb region of the pERU-S1 insert.

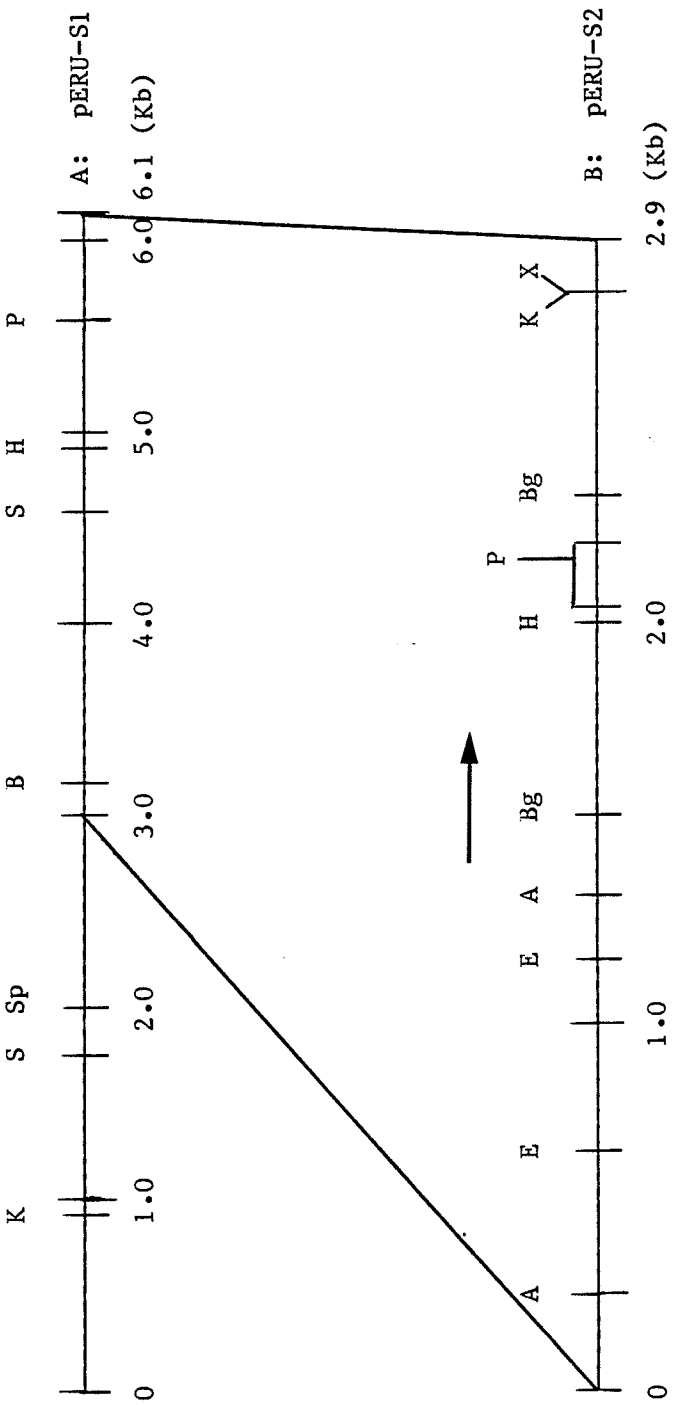
Figure 15. Size Determination of pERU-S1 Insert



lane 1: SalI digestion of pERU-S1

lane 2: SmaI digestion of pERU-S1

Figure 16. Restriction Maps of pERU-SI and pERU-S2



Arrow refers to direction of transcription.

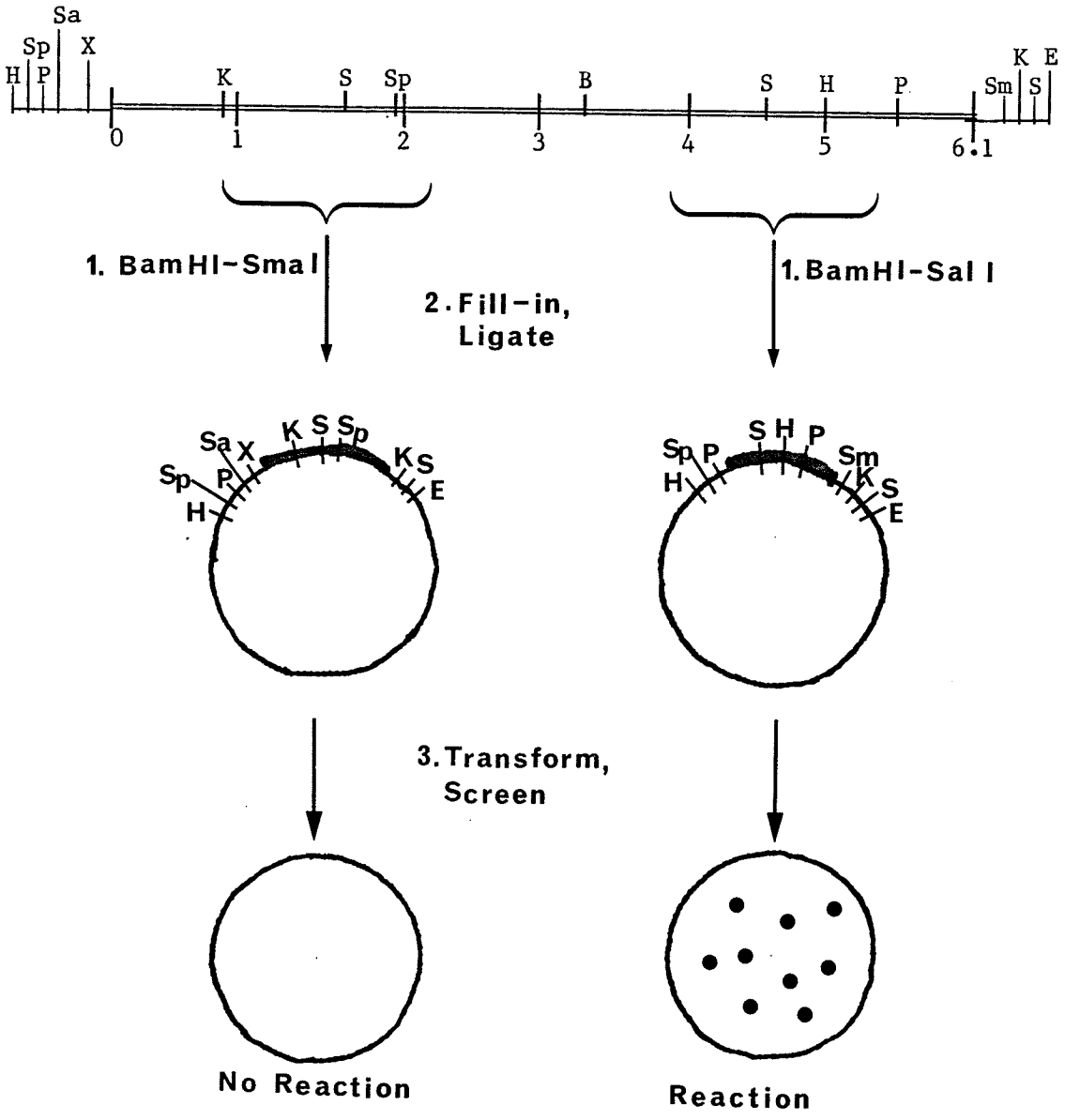
A = Asu II
B = Bam HI
Bg = Bgl II
E = EcoRI
H = Hind III
K = Kpn I
P = Pst I
S = Sac I
Sp = Sph I
X = Xba I

Localization of the gene was done by removing fragments of the insert, re-ligating and analyzing the resultant molecule for expression of the 75Kd antigen. Figure 17 diagrams the strategy taken. The pERU-S1 recombinant was cut with BamHI and a second enzyme that had a unique restriction site within the polycloning region of the vector. When the enzyme SalI was used, everything upstream of the BamHI site and downstream of the SalI site was deleted. When the enzyme SmaI was used, everything upstream of SmaI and downstream of BamHI was deleted.

The new constructions were screened for expression of the 75Kd antigen by immunoblot. Only transformants of the SalI-BamHI construction reacted with 14-E6, and no reaction was observed with transformants of the SmaI-BamHI construction.

The SalI-BamHI recombinant was named pERU-S2. A transformant of pERU-S2 was isolated in pure culture and analyzed by SDS-PAGE so that the reactive protein of pERU-S2 could be resolved and its size determined. The polyacrylamide gel was done in duplicate and one was silver stained while the other was used for western blotting. It was observed in the silver stain and western blot (Figure 18) that pERU-S2 was expressing a protein of 75Kd in size which co-migrated with the 75Kd protein of C. trachomatis. The western blot confirmed that this 75Kd protein reacted with the monoclonal antibody 14-E6. These results suggest that pERU-S2 most likely contains the complete gene encoding

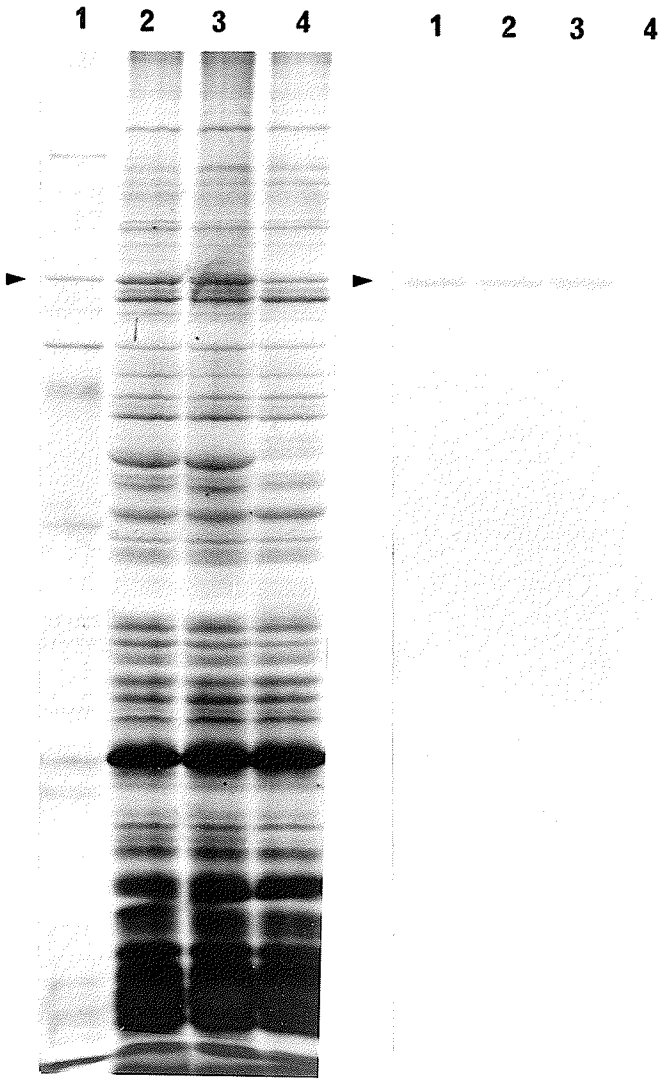
Figure 17. Strategy of Locating the 75 Kd Antigen Gene in pERU-S1



H = Hind III
Sp = SphI
P = PstI
Sa = SalI
X = XbaI

K = KpnI
B = Bam HI
Sm = SmaI
S = SacI
E = EcoRI

Figure 18. 12.5% SDS-PAGE Analysis of pERU-S2



A: Silver Stain
B: Western Blot, probed with monoclonal antibody 14-E6.

For both figures A and B; lane 1: C. trachomatis; lanes 2 and 3: transformant of pERU-S2; lane 4: E. coli DH5-alpha. Arrows point out the 75kd antigen.

the 75Kd antigen of C. trachomatis.

Plasmid DNA of pERU-S2 was isolated by CsCl dye-buoyant-density gradient centrifugation. Restriction enzyme analysis of pERU-S2 resulted in the restriction map of Figure 16B. Common restriction sites between this map and the others, Figures 10 and 16A, were found. The orientation of the common restriction sites found between EDR4-2 and pERU-S2 suggested that the direction of transcription of the 75Kd antigen gene in pERU-S2 is as shown by the arrow above the map in Figure 16B.

C) Gene Sequencing

(i) Sequencing Strategy

The strategy of locating the 75kd antigen gene and sequencing it involved three different approaches. First, the EcoRI insert purified from the EDR4-2 lambda recombinant was subcloned into M13mp19. Deletions of the M13 recombinant ERM-1 (ERM series) was performed using the DNA modifying enzyme, T4 DNA polymerase. The second method employed for sequencing the 75kd protein gene involved creating deletions of the insert in pERU-S2. Deletions in this situation were created with the use of the DNA modifying enzyme, Exonuclease III. The third approach taken involved isolation of specific fragments of pERU-S2, and subcloning them into M13mp18, and M13mp19.

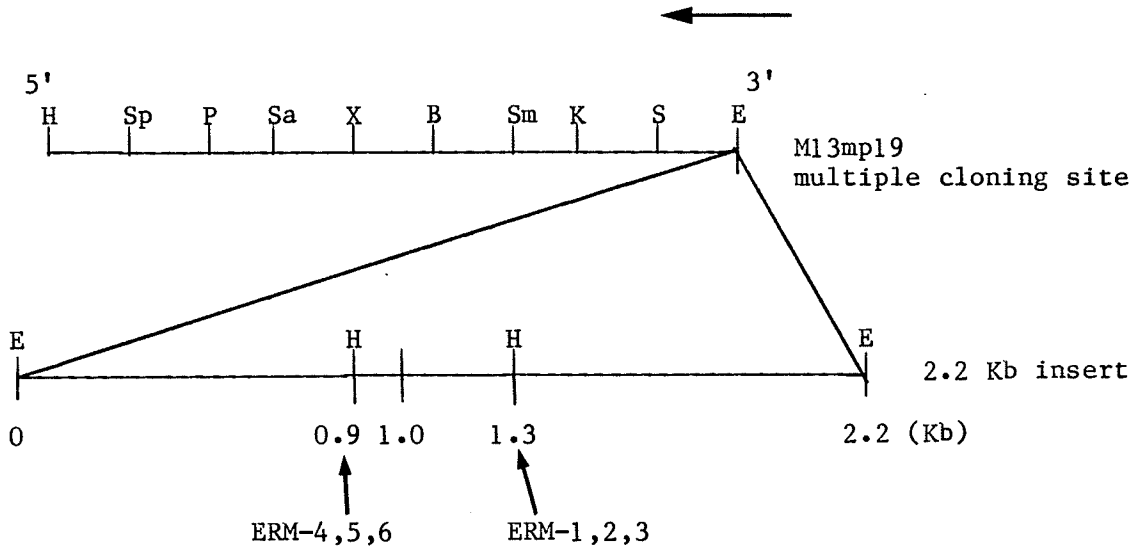
(ii) Sequence Derived from EDR4-2 Subclones

The insert of the phage recombinant, EDR4-2, was subcloned into pUC-19 for amplification. The 2.2kb EcoRI

insert was isolated from the recombinant plasmid by resolving in 0.7% LMT gel and purifiying. The purified fragment was subcloned into the EcoRI site of M13mp19. A total of six of the M13 recombinants (ERM clones) were picked and analyzed. Prior knowledge of the EcoRI 2.2kb fragment restriction map (Figure 10) and digestion of the ERM clones with the restriction enzyme HindIII revealed that three recombinants (ERM-1, -2, and -3) contained inserts in one orientation while the other three (ERM-4, -5, and -6) were in the opposite orientation. ERM-1, -2, and -3 generated HindIII fragments of approximately 1.3kb, while ERM-4, -5, and -6 generated fragments of only approximately 0.9kb. Figure 19 depicts this situation of the two orientations. Comparison of the restriction site to previous restriction data indicated that ERM-1, -2, and -3 were in the proper orientation for obtaining sequence data within the coding region of the insert. The arrow above the multiple cloning site of M13mp19 in Figure 19 indicates the direction of DNA synthesis and the correct direction of transcription of the 75kd protein gene.

In order to obtain sequence data throughout the insert, deletions were generated using the enzyme T4 DNA polymerase. Both ERM-1 and ERM-3 were used, at separate times, as the progenitor from which the deletions were generated. The deletions obtained by this method are listed in Table V, along with the length of their remaining insert and sequence obtained from them.

Figure 19. Orientaton Determinaton of ERM Recombinants.



Arrow refers to direction of DNA synthesis with the universal primer (Pharmacia) and the direction of transcription of the 75 Kd antigen.

H = Hind III
Sp = Sph I
P = Pst I
Sa = Sal I
X = Xba I
B = Bam HI
Sm = Sma I
K = Kpn I
S = Sst I
E = EcoRI

TABLE V: DELETIONS OF ERM-1

CLONE	INSERT LENGTH (KB)	SEQUENCE OBTAINED (BP)
ERM-1	2.0	404
ERM-3	2.0	350
ERM-7	1.39	326
ERM-8	1.189	358
ERM-9	1.37	209
ERM-10	1.26	208
ERM-11	----	ND*
ERM-12	----	ND
ERM-13	1.42	ND
ERM-14	1.58	312
ERM-15	0.40	252
ERM-16	0.318	184

*ND = Not Done

The ERM deletions were chosen for sequencing based on the length of the remaining insert and whether or not they would generate useful sequence data within the gene or around it. Figure 20 is a schematic representation of the deletions used for sequencing and the length of sequence obtained from them.

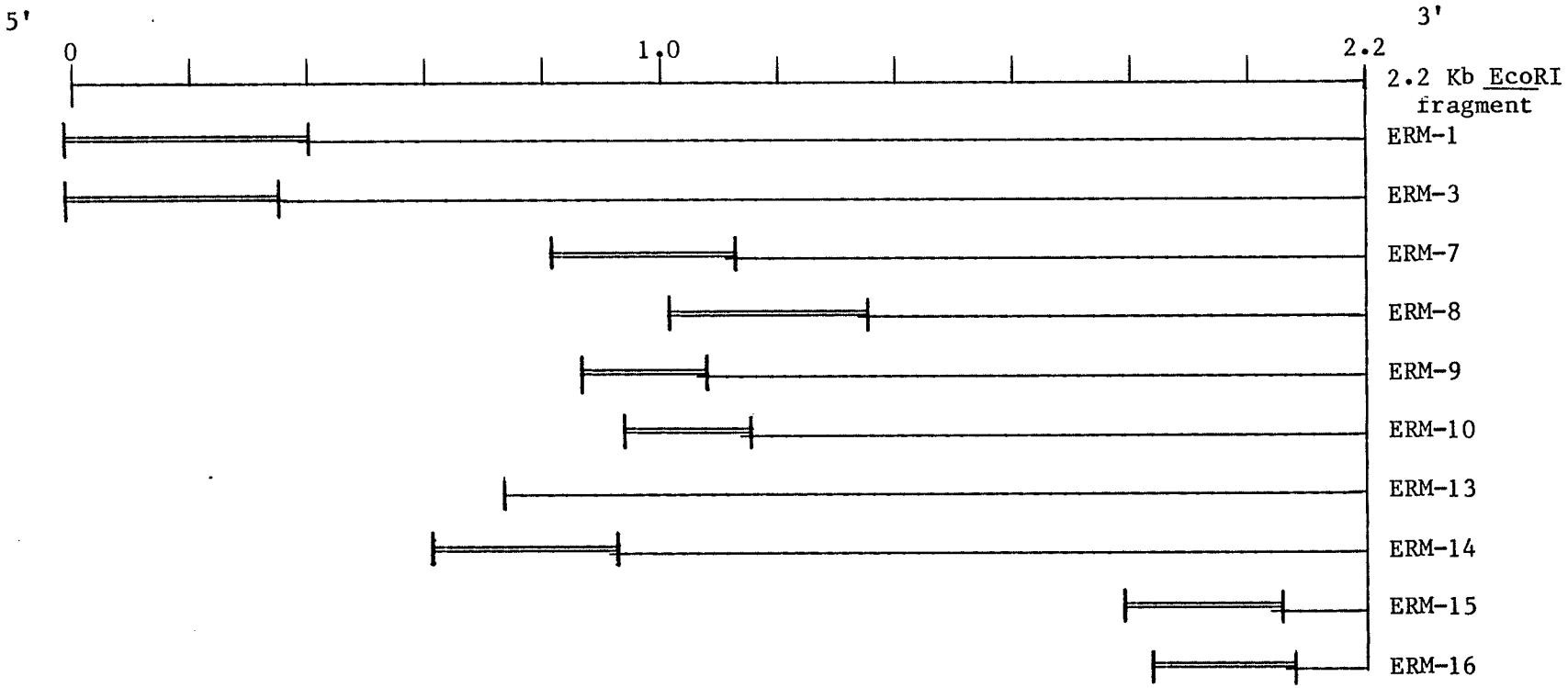
(iii) Sequence Derived from pERU-S2 Subclones

a) Exonuclease III Deletions

The recombinant molecule pERU-S2 contained an insert of 2.9kb which encoded the 75kd protein of C. trachomatis. In order to obtain sequence data from this insert, that was not provided from the ERM series of deletions, a similar method of creating deletions was used with pERU-S2. The plasmid was digested with the restriction enzymes SphI and SalI. Both of these enzymes are located within the multiple cloning site of pUC18 and do not cut within the insert. The enzyme SphI results in a 3' overhang and SalI results in a 5' overhang upon digestion of double stranded DNA (see Table II). The linear molecules were then subjected to digestion with Exonuclease III to create a series of deletions throughout the insert.

Using a crude analysis of the resultant deletion sizes, mutants which appeared to contain inserts that would provide sequence data within the gene were isolated and their plasmids were purified. The various deletion mutants chosen for further analysis by dideoxy sequencing are listed in Table VI along with the length of the remaining insert and

Figure 20. Schematic Representation of ERM-1 Deletions[1] and Sequence Derived From Them[2].



1. Length of the remaining insert within the deletion mutants is represented by the length of the line next to that mutant designation on the right.
2. Sequenced portions are represented by the double-lined regions.

TABLE VI: DELETIONS OF pERU-S2

CLONE	INSERT LENGTH (KB)	SEQUENCE OBTAINED(BP)
ERU-S2	2.9	245/175 *
ERU-S10	2.0	160
ERU-S21	0.78	315
ERU-S39	0.225	251
ERU-S61	0.55	314
ERU-S70	0.52	200
ERU-S74	1.32	242
ERU-S77	0.58	356
ERU-S85	1.5	236
ERU-S93	2.3	259

* second number refers to sequence obtained by reverse sequencing

the sequence obtained from them.

The deletions were sequenced using the double stranded dideoxy sequencing method. The sequence data were analyzed for homology with sequence data from other mutants expected to contain regions of overlap. In some cases the data were also compared with sequence from the ERM series of deletions.

b) Subclones in M13 vectors

Since the deletions generated from both ERM-1 and pERU-S2 did not span the complete insert, fragments of pERU-S2 were subcloned into M13mp18 and M13mp19, in order to obtain these regions of nucleotide sequence. Table VII lists the ERM recombinants obtained from subcloning. The recombinants ERM-20 to ERM-25 were all constructed by shotgun cloning of EcoRI fragments from pERU-S2 into M13mp19. These ligations generated two clones which contained an insert of approximately 1.65kb (ERM-20 and ERM-21) and four clones which contained inserts of approximately 0.55kb (ERM-22 to 25). Upon analysis of sequence data, it was determined that ERM-22 and ERM-24 contained the same insert, but in opposite orientation.

The recombinants ERM-26 to ERM-29 were constructed by insertion of purified pERU-S2 fragments into M13mp18. ERM-26 contains an EcoRI fragment insert of approximately 1.65kb. Recombinants ERM-27, -28, and -29 all contain a 0.7kb insert isolated from a SphI-EcoRI digest of pERU-S2. Sequence data indicated that the two recombinants ERM-28 and

TABLE VII: ERM SUBCLONES OF pERU-S2

CLONE	INSERT LENGTH (KB)	SEQUENCE OBTAINED (BP)
IN M13mp19:		
ERM-20	1.65	ND*
ERM-21	1.65	ND
ERM-22	0.55	346
ERM-23	0.55	ND
ERM-24	0.55	320
ERM-25	0.55	ND
IN M13mp18:		
ERM-26	1.65	ND
ERM-27	0.70	ND
ERM-28	0.70	323
ERM-29	0.70	353

*ND = Not Done

ERM-29 were inserted in opposite orientation from each other.

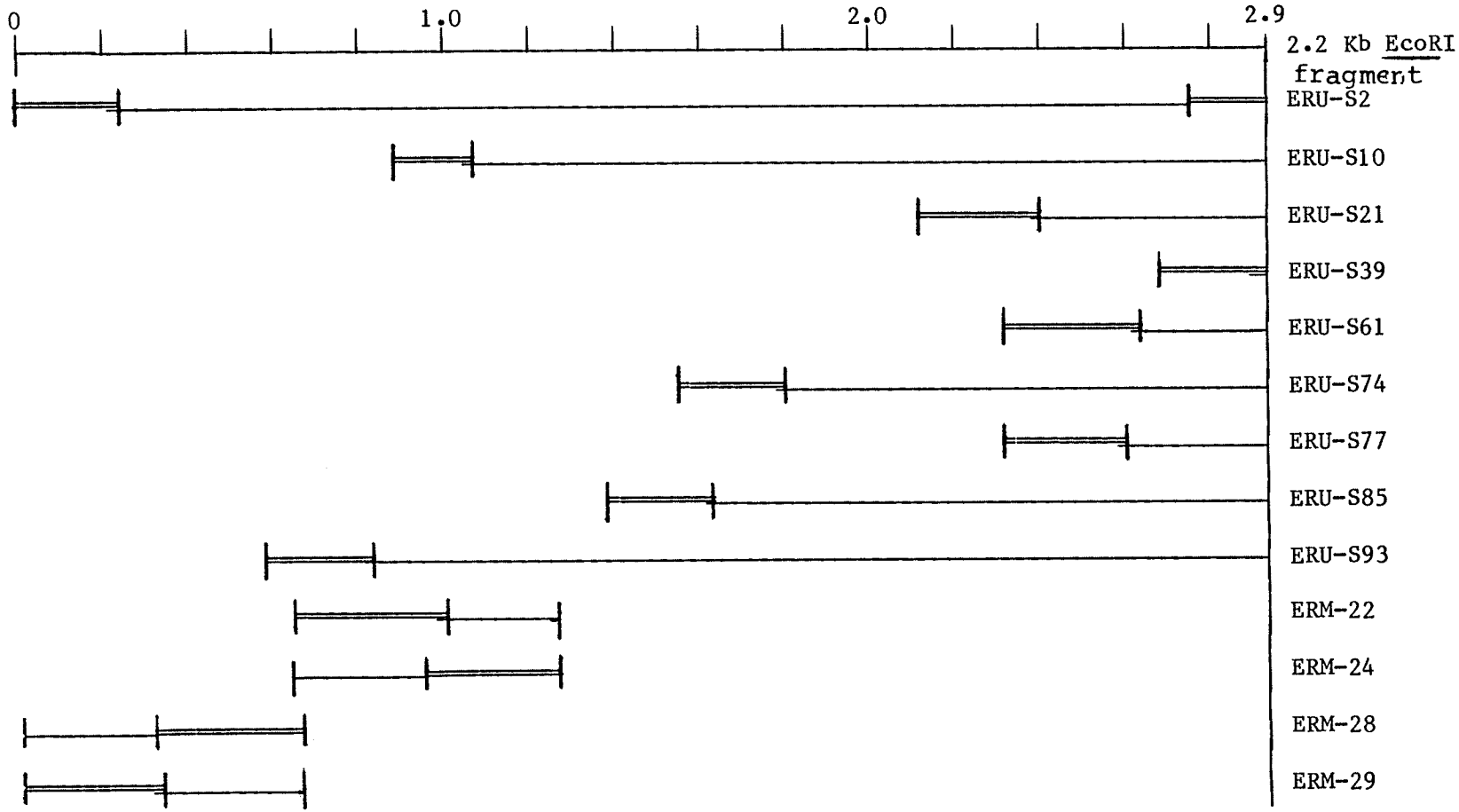
Figure 21 is a composite schematic representation of the sequence data generated from pERU-S2, deletions of pERU-S2, and M13 subclones of pERU-S2 fragments.

(iv) Nucleotide Sequence of the 75Kd protein gene

By comparison of overlapping regions of nucleotide sequence obtained from all of the various recombinants and resequencing regions of discrepancy, the complete nucleotide sequence encoding the 75kd protein gene was determined. The complete sequence is shown in Figure 22. The gene consists of an open reading frame of 1956 nucleotides coding for a protein of 652 amino acids. Table VIII is a distribution of the amino acid sequence encoded by this gene. The molecular weight of the protein was determined to be 71Kd, based on the amino acids derived from the nucleotide sequence.

Upstream of the translation start site, the Shine-Dalgarno sequence, or ribosome binding site (RBS) is indicated by the boxed 5bp sequence beginning at -7 in Figure 22. The designation of this sequence was determined by comparison of sequence upstream of the start codon to a 13bp sequence, complementary to the 3' end of E. coli 16S rRNA. It is believed that the 13bp sequence of 16S rRNA is involved in translation initiation by annealing to the mRNA (Stormo, 1986). The thirteen bases around the indicated RBS in Figure 22 had the greatest homology to the 16S rRNA complementary sequence, shown in Figure 23. This potential

Figure 21. Schematic Representation of pERU-S2 Deletions[1] and Subclones[1] and Sequence Derived From Them[2].



1. Length of the insert in the various clones is indicated by the length of the lines to the left of the clone designation.
2. Sequenced portions are shown by the double-lined regions.

Figure 22. Nucleotide and Amino Acid Sequence of the 75kd Protein Gene

```
-150         -140         -130         -120         -110         -100
 *            *            *            *            *            *
AGA AGG CTA TCT TGA GAG TTT ACT AAA GGT TAT AAG ATA GGA GAT CGT CCT ATT CGC

   -90          -80          -70          -60          -50          -40
    *            *            *            *            *            *
GTA GCC AAA GTG AAA GTA GCA AAA CTT CCT GCT AAA GAA ACT CAG ACA GTA ACG AAG

       -30          -20          -10            1           10           20
        *            *            *             *           *            *
AAA AAG AAT AAT TAT CCT AGA AAT TAG GTG TTC AAT ATG AGC GAA AAA AGA AAG TCT
                                                Met Ser Glu Lys Arg Lys Ser

          30           40           50           60           70
           *            *            *            *            *
AAC AAA ATT ATT GGT ATC GAC CTA GGG ACG ACC AAC TCT TGC GTC TCT GTT ATG GAA
Asn Lys Ile Ile Gly Ile Asp Leu Gly Thr Thr Asn Ser Cys Val Ser Val Met Glu

 80           90          100          110          120          130
  *            *            *            *            *            *
GGT GGC CAA CCT AAA GTT ATT GCC TCT TCT GAA GGA ACT CGT ACT ACT CCT TCT ATC
Gly Gly Gln Pro Lys Val Ile Ala Ser Ser Glu Gly Thr Arg Thr Thr Pro Ser Ile

    140          150          160          170          180          190
     *            *            *            *            *            *
GTT GCT TTT AAA GGT GGC GAA ACT CTT GTT GGA ATT CCT GCA AAA CGT CAG GCA GTA
Val Ala Phe Lys Gly Gly Glu Thr Leu Val Gly Ile Pro Ala Lys Arg Gln Ala Val

        200          210          220          230          240
         *            *            *            *            *
ACC AAT CCT GAA AAA ACA TTG GCT TCT ACT AAG CGA TTC ATC GGT AGA AAA TTC TCT
Thr Asn Pro Glu Lys Thr Leu Ala Ser Thr Lys Arg Phe Ile Gly Arg Lys Phe Ser

250          260          270          280          290          300
 *            *            *            *            *            *
GAA GTC GAA TCT GAA ATT AAA ACA GTC CCC TAC AAA GTT GCT CCT AAC TCG AAA GGA
Glu Val Glu Ser Glu Ile Lys Thr Val Pro Tyr Lys Val Ala Pro Asn Ser Lys Gly

   310          320          330          340          350          360
    *            *            *            *            *            *
GAT GCG GTC TTT GAT GTG GAA CAA AAA CTG TAC ACT CCA GAA GAA ATC GGC GCT CAG
Asp Ala Val Phe Asp Val Glu Gln Lys Leu Tyr Thr Pro Glu Glu Ile Gly Ala Gln

       370          380          390          400          410          420
        *            *            *            *            *            *
ATC CTC ATG AAG ATG AAG GAA ACT GCT GAG GCT TAT CTC GGA GAA ACA GTA ACG GAA
Ile Leu Met Lys Met Lys Glu Thr Ala Glu Ala Tyr Leu Gly Glu Thr Val Thr Glu

            430          440          450          460          470
             *            *            *            *            *
GCA GTC ATT ACC GTA CCA GCT TAC TTT AAC GAT TCT CAA AGA GCT TCT ACA AAA GAT
Ala Val Ile Thr Val Pro Ala Tyr Phe Asn Asp Ser Gln Arg Ala Ser Thr Lys Asp

 480          490          500          510          520          530
  *            *            *            *            *            *
GCT GGA CGT ATC GCA GGA TTA GAT GTT AAA CGC ATT ATT CCT GAA CCA ACA GCG GCC
Ala Gly Arg Ile Ala Gly Leu Asp Val Lys Arg Ile Ile Pro Glu Pro Thr Ala Ala

     540          550          560          570          580          590
      *            *            *            *            *            *
GCT CTT GCT TAT GGT ATT GAT AAG GAA GGA GAT AAA AAA ATC GCC AGT CTT CGA CTT
Ala Leu Ala Tyr Gly Ile Asp Lys Glu Gly Asp Lys Lys Ile Ala Ser Leu Arg Leu
```

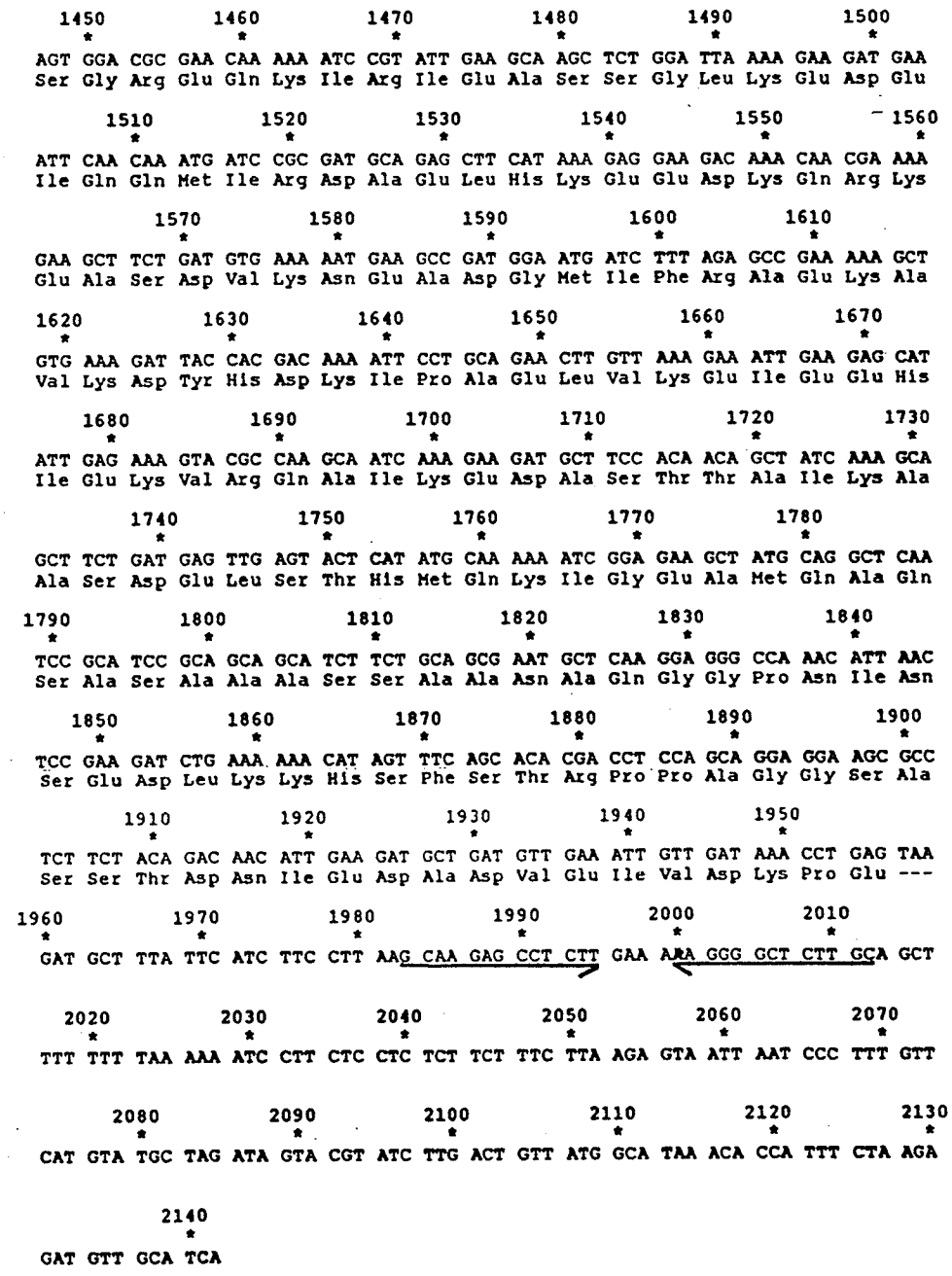
Continued....

Figure 22, Cont'd.

```
          600           610           620           630           640
           *             *             *             *             *
AGG AGG AGG AAC TTT CGA TAT TTC TAT CTT GGA AAT CGG TGG ATC GGA GTT TTT GAA
Arg Arg Arg Asn Phe Arg Tyr Phe Tyr Leu Gly Asn Arg Trp Ile Gly Val Phe Glu

 650           660           670           680           690           700
  *             *             *             *             *             *
GTT CTC TCA ACC AAC GGG GAT ACT CAC TTG GGA GGA GAC GAC TTC GAT GAA TTC AAA
Val Leu Ser Thr Asn Gly Asp Thr His Leu Gly Gly Asp Asp Phe Asp Glu Phe Lys

     710           720           730           740           750           760
      *             *             *             *             *             *
AAA CAA GAA GGC ATT GAT CTA AGC AAA GAT AAC ATG GCT TTG CAA AGA TTG AAA GAT
Lys Gln Glu Gly Ile Asp Leu Ser Lys Asp Asn Met Ala Leu Gln Arg Leu Lys Asp

          770           780           790           800           810
           *             *             *             *             *
GCT GCT GAA AAA GCA AAA ATA GAA TTG TCT GGT GTA TCG TCT ACT GAA ATC AAT CAG
Ala Ala Glu Lys Ala Lys Ile Glu Leu Ser Gly Val Ser Ser Thr Glu Ile Asn Gln

820           830           840           850           860           870
 *             *             *             *             *             *
CCA TTC ATC ACT ATC GAC GCT AAT GGA CCT AAA CAT TTG GCT TTA ACT CTA ACT CGC
Pro Phe Ile Thr Ile Asp Ala Asn Gly Pro Lys His Leu Ala Leu Thr Leu Thr Arg

    880           890           900           910           920           930
     *             *             *             *             *             *
GCT CAA TTC GAA CAC CTA GCT TCC TCT CTC ATT GAG CGA ACC AAA CAA CCT TGT GCT
Ala Gln Phe Glu His Leu Ala Ser Ser Leu Ile Glu Arg Thr Lys Gln Pro Cys Ala

        940           950           960           970           980           990
         *             *             *             *             *             *
CAG GCT TTA AAA GAT GCT AAA TTG TCC GCT TCT GAC ATT GAT GAT GTT CTT CTA GTT
Gln Ala Leu Lys Asp Ala Lys Leu Ser Ala Ser Asp Ile Asp Asp Val Leu Leu Val

           1000          1010          1020          1030          1040
            *             *             *             *             *
GGC GGA ATG TCC AGA ATG CCT GCG GTA CAA GCA GTT GTA AAG AGA TCT TTG GTA AAG
Gly Gly Met Ser Arg Met Pro Ala Val Gln Ala Val Val Lys Arg Ser Leu Val Lys

1050          1060          1070          1080          1090          1100
 *             *             *             *             *             *
AGC CTA ATA AAG GCC GTC AAT CCA GAT GAA GTT GTA GCG ATT GGA GCT GCT ATT CAG
Ser Leu Ile Lys Ala Val Asn Pro Asp Glu Val Val Ala Ile Gly Ala Ala Ile Gln

    1110          1120          1130          1140          1150          1160
     *             *             *             *             *             *
GGT GGT GTC CTC GGC GGA GAA GTG AAA GAC GTT CTG TTG TTG GAT GTG ATT CCC CTC
Gly Gly Val Leu Gly Gly Glu Val Lys Asp Val Leu Leu Leu Asp Val Ile Pro Leu

          1170          1180          1190          1200          1210
           *             *             *             *             *
TCT TTA GGA ATT GAG ACT CTA GGT GGG GTC ATG ACT CCT TTG GTA GAG AGA AAC ACT
Ser Leu Gly Ile Glu Thr Leu Gly Gly Val Met Thr Pro Leu Val Glu Arg Asn Thr

1220          1230          1240          1250          1260          1270
 *             *             *             *             *             *
ACA ATC CCT ACT CAG AAG AAG CAA ATC TTC TCT ACA GCC GCT GAC AAT CAG CCA GCA
Thr Ile Pro Thr Gln Lys Lys Gln Ile Phe Ser Thr Ala Ala Asp Asn Gln Pro Ala

   1280          1290          1300          1310          1320          1330
    *             *             *             *             *             *
GTG ACT ATC GTC GTT CTT CAA GGT GAA CGG CCT ATG GCG AAA GAC AAT AAG GAA ATT
Val Thr Ile Val Val Leu Gln Gly Glu Arg Pro Met Ala Lys Asp Asn Lys Glu Ile


          1340          1350          1360          1370          1380
           *             *             *             *             *
GGA AGA TTT GAT CTA ACA GAC ATT CCT CCT GCT CCT CGC GGC CAT CCA CAA ATT GAG
Gly Arg Phe Asp Leu Thr Asp Ile Pro Pro Ala Pro Arg Gly His Pro Gln Ile Glu

1390          1400          1410          1420          1430          1440
 *             *             *             *             *             *
GTA ACC TTC GAT ATT GAT GCC AAC GGA ATT TTA CAC GTT TCT GCT AAA GAT GCT GCT
Val Thr Phe Asp Ile Asp Ala Asn Gly Ile Leu His Val Ser Ala Lys Asp Ala Ala
```

Continued....

Figure 22, Cont'd

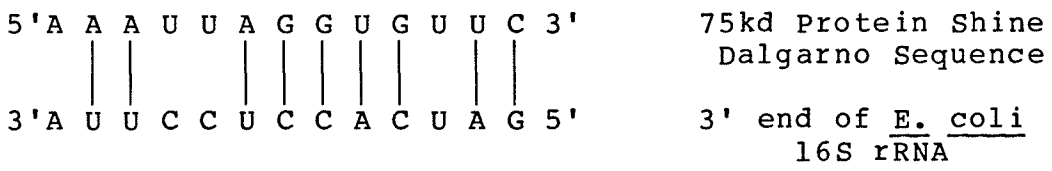


Boxed regions upstream of the translation start site, centered at -136 and -117, represent the putative promoter regions -35 and -10, respectively.
The five base, boxed sequence, located 7 bases upstream of the translation start site represents the ribosome binding site.
The transcription terminator is shown by the inverted arrows located 24 bases downstream of the translation stop codon.

TABLE VIII: 75KD PROTEIN AMINO ACIDS

RESIDUE	NUMBER PRESENT	MOLE PERCENT
ALANINE	71	10.890
CYSTEINE	2	0.307
ASPARTIC ACID	44	6.748
GLUTAMIC ACID	54	8.282
PHENYLALANINE	17	2.607
GLYCINE	46	7.055
HISTIDINE	10	1.534
ISOLEUCINE	51	7.822
LYSINE	56	8.589
LEUCINE	44	6.748
METHIONINE	13	1.994
ASPARAGINE	21	3.221
PROLINE	29	4.448
GLUTAMINE	27	4.141
ARGININE	30	4.601
SERINE	46	7.055
THREONINE	38	5.828
VALINE	44.	6.748
TRYPTOPHAN	1	0.153
TYROSINE	8	1.227

Figure 23. Complementarity of the 5' sequences of the 75Kd protein gene to 3' end of E. coli 16S rRNA



Shine-Dalgarno sequence also adheres to other criteria for a typical RBS. On the average, RBS's are 5 ± 1 bases in length and located 7 ± 2 bases upstream of the start site and should contain at least two guanine residues (Stormo, 1986). The RBS chosen fits into these guidelines.

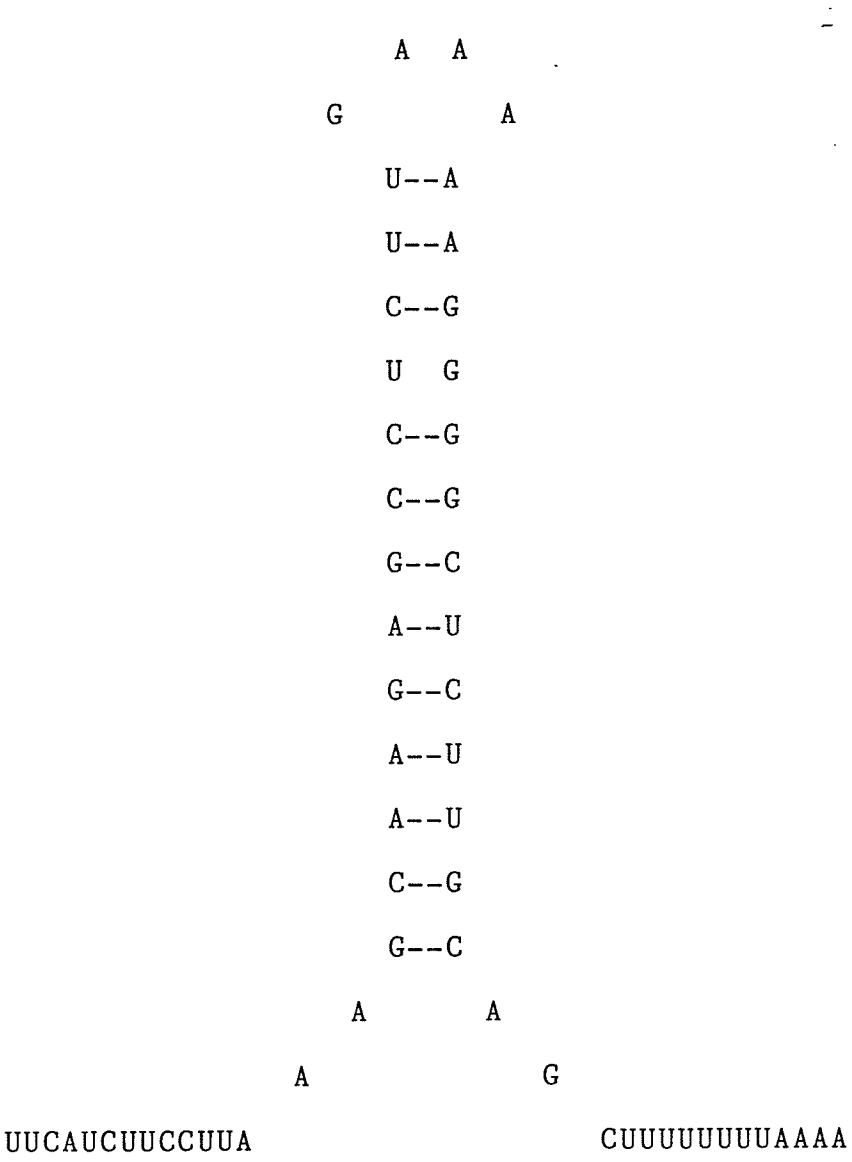
Farther upstream, potential transcription promoters were identified by their homology to typical E. coli promoters. The transcription promoter regions of the 75Kd protein gene are centered around -136 and -117, upstream from the translation start site. These regions are boxed in Figure 22. The potential promoter elements were located by a homology search of 490 bases upstream from the translation start site of the 75Kd protein gene, to the -35 and -10 consensus sequences of E. coli promoters. The two regions identified showed the greatest homology at 83%, independently, as depicted in Figure 24.

Downstream from the translational stop codon, a typical rho-independent terminator, was found. The terminator consisted of a 13 base dyad, indicated by arrows in Figure 22, beginning at base 1983 and ending at base 2012. The 3' end of the dyad is followed by eight thymidine residues in the DNA. Transcription of this region would result in folding of the mRNA into the secondary structure shown in Figure 27. This stem and loop structure has a calculated least free energy of -24.4kcal (Zuker and Stiegler, 1981).

Figure 24. Homology to Prokaryotic Promoters

```
5' T T G A C A 3'      -35 Consensus Sequence
   | | | |   |
5' T T G A G A 3'      -139 to -134 of 75kd gene sequence


5' T A T A A T 3'      -1Ø Consensus Sequence
   | | | | |
5' T A T A A G 3'      -12Ø to -115 of 75kd gene sequence
```

Figure 25. Secondary Structure of Transcription Terminator



Free energy = -24.4 kcal

Discussion

Genital infection with Chlamydia trachomatis is a serious public health problem. Studies of women with salpingitis, ectopic pregnancy, and tubal infertility, have implicated C. trachomatis as the etiologic agent in many cases of these diseases. It has been estimated that approximately 20% of infected women will be rendered infertile (Westrom, 1980).

Studies on the humoral immune response to chlamydial infection in women have suggested that the presence of antibody to certain surface exposed antigens may provide protective immunity against ascending infection (Brunham et al., 1987). The proteins which appear to play a role in eliciting protective immunity include the 100kd, 75kd, 60kd, 57kd, 32kd, and 29kd polypeptides (Brunham et al., 1987).

Cloning of chlamydial antigenic proteins that appear to be important in the infection process is an invaluable method for obtaining information on the mechanism of pathogenesis of these organisms. Cloned genes expressing antigen can provide an abundance of the protein for examination of its biological properties and potential for vaccine use. Characterization of the gene and its regulation can provide information on the developmental biology of chlamydial species, of which little is known at this time (Allan, 1986).

This project was undertaken to isolate the gene of one of the chlamydial proteins important in the immune response

to infection. Isolation of the gene was desired to facilitate genetic characterization and potential use of the recombinant gene product for studying the pathogenesis of this organism.

Preliminary experiments involved the production of a chlamydial genomic library in Lambda gt11. This cloning system was chosen for its ease of library preparation and the nature of the host bacteria. The lysogenic host, E. coli Y1089(r-), contains a variety of mutations which make it a suitable host for expression of foreign gene products without degradation by host proteases.

Only one of the desired chlamydial proteins was expressed in the Lambda gt11 library. These recombinants produced a protein which contained an epitope found on the 75kd antigen of C. trachomatis. The recombinant gene product has a molecular weight of 150kd, which specifically binds to a monoclonal antibody directed against the 75kd antigen of C. trachomatis.

The recombinant molecule is expressed as a fusion protein of B-galactosidase (B-gal) and the 75Kd outer membrane protein; therefore, it is expected that the recombinant protein would be larger than the native chlamydial antigen. The EcoRI insertion site in Lambda gt11 is located near the 3' end of the B-gal gene. When a fusion protein is expressed, most of the B-gal gene is transcribed and translated. The fusion protein was 150kd in size and B-

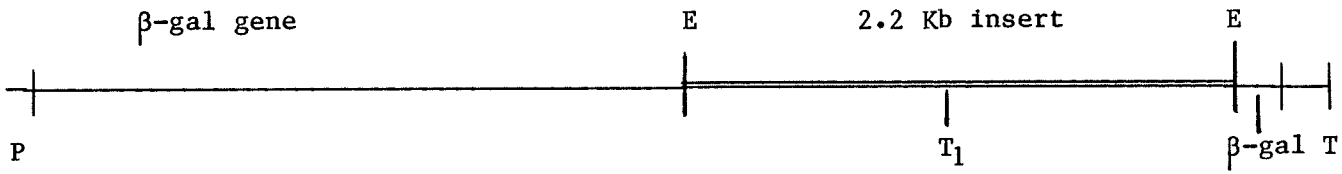
gal has a molecular weight of 116Kd, therefore approximately 38Kd of the fusion protein must be chlamydial protein product. Consequently, approximately half of the 75Kd protein gene is located within the insert of the immunopositive Lambda gt11 (EDR) recombinants.

Assuming an average molecular weight for an amino acid of 110 daltons, a protein of 75Kd must consist of 680 amino acids. A protein of this size would require a gene of at least 2000 nucleotides in length to code for that number of amino acids. Since the fusion protein contains only half of the 75Kd protein, the foreign DNA insert of the recombinant must contain at least half of the gene and be at least 1000 nucleotides long. The actual insert size is 2.2kb in length.

Since only half of the gene was present, the insert most likely contains the latter half, including the stop codon. If the fusion protein was terminated by the B-gal stop codon the size of the protein product would be in the order of 190Kd. Figure 26 is a schematic representation of the chlamydial DNA insert in Lambda gt11.

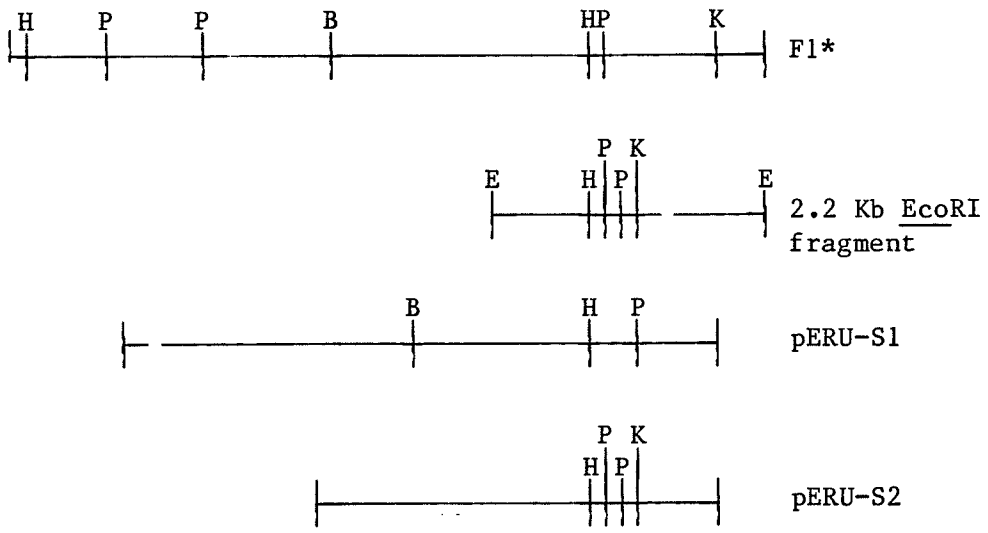
Comparison of the restriction map generated from the 2.2kb fragment and the restriction map of a DNA fragment expressing a 74kd protein (Kaul and Wenman, 1985) showed similarities, see Figure 27. The common restriction sites suggest that the antigen referred to here as the 75kd antigen may in fact be the same as the 74kd antigen cloned

Figure 26. Schematic Representation of the 2.2 Kb Insert in Lambda gt11



P = β-gal promotor
T_1 = approximate termination site of fusion protein, or 75 Kd antigen gene
T = β-gal terminator
E = Eco RI

Figure 27. Comparison of restriction maps of the 75 Kd protein gene recombinants and the 74 Kd protein gene recombinant F1.



*F1 adapted from Kaul and Wenman (1985).

H = HindIII
P = PstI
B = BamHI
K = KpnI
E = EcoRI

by other investigators (Kaul and Wenman, 1985; Bavoil et al., 1986; Palmer and Falkow, 1986b).

Subclones of the 2.2kb fragment in M13mp19 were constructed for the purpose of obtaining DNA sequence of the gene. In order to obtain the DNA sequence throughout the 2.2kb fragment, deletions of the insert in M13mp19 were created. A number of attempts were made to produce adequate deletion mutants evenly spaced over the insert to obtain sequence data of the coding region. Relatively few deletion mutants were generated that produced useful sequence data. A possible reason for this problem may lie in the size of the chlamydial DNA insert. In general, inserts of greater than 1000 nucleotides tend to be unstable in M13 vectors (Chen and Seeburg, 1985). Perhaps, the low number of deletion mutants created was due to spontaneous rearrangement of the ERM recombinants, prior to, during, or following the deletion experiments.

Since recombinants obtained from the Lambda gt11 library did not contain the complete gene, another attempt, using a different cloning approach was utilized. This procedure involved a partial digestion of C. trachomatis genomic DNA with the restriction endonuclease Sau3AI. This cloning approach has proven fruitful for the isolation of a number of chlamydial genes by other investigators.

The gene bank constructed in pUC18 produced a single recombinant plasmid, pERU-S1, that expressed the 75kd

antigen epitope and hybridized to the 2.2kb EcoRI fragment of the EDR recombinants.

Colony immunoblotting revealed only the one recombinant, pERU-S1. However colony DNA-DNA hybridizations revealed that many recombinants hybridized to the 2.2kb EcoRI fragment. Even though these recombinants contained DNA which hybridized to the 2.2kb probe, absence of reaction of these recombinants by immunoblot, was a result of lack of expression of the 75kd protein. The absence of expression could be due to a number of reasons;

1. These recombinants may not contain the complete 75Kd protein gene, and were missing the portion coding for the epitope recognized by the monoclonal antibody.
2. The recombinants may contain DNA homologous to only a small portion of the probe or to a region of the probe in which the 75Kd gene is not found.
3. Assuming that the recombinant protein expressed is a fusion with B-galactosidase, the 75Kd gene may not be in frame with the B-gal promoter, resulting in a frame-shift mutant protein that would not react with the monoclonal antibody 14-E6.

Subsequent analysis of the immunoblot-positive gene product of pERU-S1 revealed that the protein was 75kd in size, co-migrating with the 75kd antigen of C. trachomatis. It was also obvious from SDS-PAGE analysis of this

recombinant that the 75kd gene product was abundantly produced by the host E. coli, accounting for one of the most highly expressed proteins in the cell.

Analysis of pERU-S1 with restriction endonucleases revealed an insert size of 6.1kb. Limited restriction enzyme analysis of the insert revealed common restriction sites with the 2.2kb EcoRI fragment in Lambda gt11 and to the recombinants isolated by Kaul and Wenman (1985) (Figure 27). The length of this insert is certainly adequate to contain the gene encoding the 75kd protein of C. trachomatis.

These observations suggested that the protein expressed by pERU-S1 was not a fusion protein with B-gal, and that pERU-S1 contained the complete gene encoding the 75Kd protein, including its own regulatory elements. If this is true, the chlamydial regulatory elements must be functional in E. coli and therefore should be similar to other prokaryotic regulatory sequences. The extent to which the protein is expressed suggested that the promoter elements of this gene must have a high affinity for the transcriptional and/or translational machinery of E. coli.

In order to locate the 75kd protein gene within the 6.1kb insert of pERU-S1, different portions of the insert were removed and the resultant recombinants analyzed by colony immunoblot for antigen expression. Deletion of

approximately 3kb of the pERU-S1 insert led to the construction of the subclone pERU-S2. This subclone retained full expression and immunoreactivity of the 75kd gene product, within an insert of 2.9kb. Restriction enzyme analysis again revealed similarities to the 2.2kb EcoRI fragment of EDR4-2 and the clones of Kaul and Wenman (1985) (Figure 27).

For the purpose of obtaining sequential DNA sequence data of the gene in pERU-S2, deletions of the insert were constructed. Although the deletion mutants obtained consisted of a variety of deletions, again they were not spaced well enough to provide sequence of the complete insert or the gene. One of the limitations of creating deletions for sequencing in a plasmid vector is the extent of sequence data that can be obtained from a single recombinant. Double stranded sequencing does not tend to generate as much sequence as single stranded sequencing; therefore, many deletions are required which are closely spaced.

The regions of pERU-S2 from which sequence data were not yet generated, were subcloned into M13mp18 and M13mp19 vectors. Single stranded dideoxy sequencing of these subclones, in both directions, generated the absent sequence data.

Analysis of the sequence data by comparing and combining all of the overlapping regions generated the

complete 75kd protein gene sequence. The sequence consists of a 1956 nucleotide open reading frame, beginning with an ATG start codon and ending with a TAA stop codon. The open reading frame codes for 652 amino acids, producing a protein with a molecular weight of 70,558 daltons. Since SDS-PAGE sizing of proteins is only an estimation, the discrepancy between the molecular weight derived by SDS-PAGE and DNA sequence is not significant. Another possibility is that the 75kd protein is a glycoprotein and therefore would electrophorese at a slower rate than expected, resulting in an inaccurate estimate of molecular weight.

Preceding the start codon of the gene is a Shine-Dalgarno sequence of 5'AGGTG 3'. The sequence is 5 bases in length and begins 7 bases upstream from the start codon. The sequence was arrived at by comparing the upstream sequence to a 13 base sequence complementary to the 3' end of E. coli 16S rRNA. The 3' end of chlamydial 16S rRNA has not been determined, therefore the E. coli sequence was used. Since the cloned gene was expressed, the 75kd protein gene must have utilized the E. coli ribosomes for translation initiation. The 3' end of 16S rRNA is believed to be involved in direct interaction and attachment to mRNA transcripts for translation initiation (Stormo, 1986). The presence of conserved sequence in the upstream region of the 75kd protein gene suggests that chlamydial 16S rRNA subunit may be similar to that of E. coli. Indeed, Palmer and

Falkow (1986b) have sequenced 90% of the C. trachomatis 16S rRNA gene and found 70% sequence homology to the E. coli 16S rRNA gene.

Farther upstream, centered around -136 and -117, putative transcriptional promoter sequences analogous to the -35 and -10 regions of E. coli, were identified. These sequences were discovered through a homology search of 490 upstream bases of the 75Kd protein gene, to the consensus sequences of the -10 and -35 regions. The search for each was carried out separately and independently from the other. Each sequence was 83% homologous to the analogous consensus sequence, with five out of the six bases being identical. No other sequences with 70% or greater homology to the -35 consensus sequence were found in the 490 upstream bases. Two sequences showed greater than 70% homology to the -10 consensus sequence. Both sequences were 83% homologous, however, one was located too close to the translation start codon to be the correct sequence. The other sequence was located optimally in conjunction with the previously identified putative promoter, analogous to the -35 region.

The 75Kd protein gene promoter, analogous to the -35 sequence identified in E. coli, consisted of the sequence 5' TTGAGA 3'. The guanine residue in the fifth position is different from the E. coli -35 sequence which contains a cytosine residue in this position, as in 53% of the promoters analyzed by Hawley and McClure (1983).

The second promoter identified in the 75kd protein

gene, and analogous to the -10 sequence in E. coli, consisted of the sequence 5' TATAAG 3'. In this situation the difference occurs in the sixth position where a thymidine residue of the consensus sequence in E. coli is a guanine residue in the chlamydia promoter sequence. The presence of a thymidine residue in the sixth position had an occurrence of 96% in the Hawley and McClure study (1983).

The spacing between the -10 and -35 regions of promoters is also important. Reznikoff and McClure (1986) state that the consensus sequences should be spaced by 17 $\pm$1 bases, again based on the compilation by Hawley and McClure (1983). The spacing between the putative promoter regions of the 75kd protein gene was 13 bases. Since the two regions were identified independently of each other, and no other compatible sequences were identified, these sequences, centered around -136 and -117, most likely represent the RNA polymerase recognition sequences for the 75kd protein gene.

The minor differences in promoter sequences and spacing is probably a result of divergence between chlamydial and Enterobacteriaceae.

The upstream regulatory elements identified in chlamydial genes are compiled in Table IX. There does not appear to be any significant homology between the promoter elements, except for the fact that they were all identified by comparison to the -10 and -35 consensus sequences derived

TABLE IX: CHLAMYDIAL GENE REGULATORY ELEMENTS

GENE	RBS[1]	-35	-10	SPACE	REFERENCE
MOMP L1	AGAGGT	TTGCAA	ATTAATT	15	Pickett et al (1987)
MOMP L2	GAGGT	NI[2]	NI	--	Stephens et al., (1986)
18kd protein	AGGA	NI	NI	--	Kaul et al., (1987)
rRNA P1	--	TGCAGA	TAAGAT	17	Engel, 1987
P2	--	TTTACG	AACAAT	20	Engel, 1987
75kd protein	AGGTG	TTGAGA	TATAAG	13	This work

1. Ribosome Binding Site

2. NI = Not Identified

from E. coli. This may reflect differences in the regulation of expression of genes in chlamydial species, or simply divergence through evolution.

The most divergent promoter elements are those of the P1 and P2 elements of the rRNA operon (Engel and Ganem, 1987). The 16S and 23S rRNA genes should consist of strong promoters since these gene products are required in abundance during the reticulate body stage of the life cycle.

The lack of identification of promoter elements upstream of the 18Kd protein gene may reflect its regulation and/or the possibility that it is present in an operon within the chlamydial genome, as suggested for the 60Kd and 12Kd protein genes (Bavoil et al., 1986).

The promoter elements identified in the MOMP and 75Kd protein gene upstream sequences show a small degree of similarity in sequence and spacing in the -10 and -35 promoter regions of these genes. Both of these proteins are expressed early in the reticulate body stage (Palmer and Falkow, 1986a; Hatch et al., 1984); therefore the similarity in their promoter elements may reflect involvement of a similar regulatory mechanism.

The Shine-Dalgarno sequences of the chlamydial structural genes sequenced to date show some homology. The sequences identified in the MOMP L1 and L2 serovars are identical to each other; the upstream sequences of these two serovars are identical until -133. The analogous sequence

in the 75Kd protein gene is identical with the exception of one base. The Shine-Dalgarno sequence of the 18Kd protein gene is the least homologous with only three bases identical to either the MOMP or 75Kd protein putative Shine-Dalgarno sequences.

The differences and similarities between the regulatory sequence elements implies that regulation at the transcriptional level is important in chlamydial gene expression.

Transcription termination in the 75Kd protein gene occurs through the rho-independent mechanism, as determined by sequence data. A typical terminator of this type was identified 24 bases downstream of the stop codon. This type of terminator has also been found in the MOMP gene sequence, but not in the 18Kd protein gene sequence. Therefore, termination of transcription must occur via the rho-independent mechanism with some chlamydial genes.

The presence of both transcriptional promoters and terminators bordering the 75Kd protein gene sequence stronly suggests that expression of this protein is independently regulated within chlamydial species.

In conclusion, the gene encoding the 75Kd protein of C. trachomatis consists of an open reading frame of 1956 nucleotides. The open reading frame codes for 652 amino

acids, producing a protein with a molecular weight of 70,558 daltons. The coding region of the gene is preceded by putative promoter elements and a ribosome binding site and followed by a rho-independent transcription terminator. Therefore, the regulatory sequences of the 75Kd protein gene strongly suggest that it is present in the genome as a distinct entity, not within an operon.

References

Adhya, S. and M. Gottesman. 1978. Control of transcription termination. Ann. Rev. Biochem. 47: 967-996.

Allan, I., T. M. Cunningham, and M. A. Lovett. 1984. Molecular cloning of the major outer membrane protein of Chlamydia trachomatis. Infect. Immun. 45: 637-641.

Allan, I. 1986. Chlamydial antigenic structure and genetics. pp. 73-80. In D. Oriel, G. Ridgway, J. Schachter, D. Taylor-Robinson and M. Ward (eds). Chlamydial Infections. Cambridge University Press. Cambridge.

Barbour, A. G., K. I. Amano, T. Hackstadt, L. Perry, and H. D. Caldwell. 1982. Chlamydia trachomatis has penicillin-binding proteins but not detectable muramic acid. J. Bacteriol. 151: 420-428.

Bavoil, P., A. Ohlin, and J. Schacter. 1984. Role of disulfide bonding in outer membrane structure and permeability in Chlamydia trachomatis. Infect. Immun. 44:479-485.

Bavoil, P., L. Palmer, S. Falkow, and D. Gump. 1986. Characterization and cloning of cysteine-rich proteins from Chlamydia trachomatis: a preliminary report. pp. 97-100. In, D. Oriel, G. Ridgway, J. Schachter, D. Taylor-Robinson and M. Ward (eds). Chlamydial Infections. Cambridge University Press. Cambridge.

Brunham, R. C., I. W. Maclean, B. Binns, and R. Peeling. 1985. Chlamydia trachomatis: its role in tubal infertility. J. Infect. Dis. 152: 1275-1282.

Brunham, R. C., B. Binns, J. McDowell, and M. Paraskevas. 1986. Chlamydia trachomatis infection in women with ectopic pregnancy. Obstet. Gynecol. 67: 772-776.

Brunham, R. C., R. Peeling, I. Maclean, J. McDowell, K. Persson, and S. Osser. 1987. Post abortal Chlamydia trachomatis salpingitis: Correlating risk with antigen-specific serological responses and with neutralization. J. Infect. Dis. 155: 749-755.

Byrne, G. I. 1976. Requirements for ingestion of Chlamydia psittaci by mouse fibroblasts (L cells). Infect. Immun. 14:645-651.

Byrne, G. I. and J. W. Moulder. 1978. Parasite-specified phagocytosis of Chlamydia psittaci and Chlamydia trachomatis by L and HeLa cells. Infect. Immun. 19: 598-606.

Caldwell, H. D., J. Kromhout, and J. Schachter. 1981. Purification and partial characterization of the major outer membrane protein of Chlamydia trachomatis. Infect. Immun. 31:1161-1176.

Caldwell, H. D., and J. Schachter. 1982. Antigenic analysis of the major outer membrane protein of Chlamydia spp. Infect. Immun. 35: 1024-1031.

Chen and Seeburg, 1985. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA. 4: 165-170.

Clarke, I. N. and C. Hatt. 1986. In vitro transcription/translation analysis of cloned plasmid DNA from Chlamydia trachomatis serovar L1. pp. 85-88. In D. Oriel, G. Ridgway, J. Schachter, D. Taylor-Robinson and M. Ward (eds). Chlamydial Infections. Cambridge University Press. Cambridge.

Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18S rDNA. Plasmid. 13: 31-40.

Eissenberg, L. G., P. B. Wyrick, C. H. Davis, and J. W. Rumpp. 1983. Chlamydia psittaci elementary body envelopes: Ingestion and inhibition of phagolysosome fusion. Infect. Immun. 40: 741-751.

Engel, J. N., and D. Ganem. 1987. Chlamydial rRNA operons: Gene organization and identification of putative tandem promoters. J. Bacteriol. 169: 5678-5685.

Garrett, A. J., M. J. Harrison, and G. P. Manire. 1974. A search for the bacterial mucopeptide component, muramic acid, in chlamydia. J. Gen. Microbiol. 80: 315-318.

Gorini, L. and E. Kataja. 1964. Phenotypic repair by streptomycin of defective genotypes in E. coli. Proc. Nat. Acad. Sci. U.S.A. 51:487-493.

Grunstein, M., and D. Hogness. 1975. Colony Hybridization: A method for isolation of cloned DNAs that contain a specific gene. Proc. Nat. Acad. Sci. U.S.A. 72: 3961-3965.

Guo, L., R. C. A. Yang, and R. Wu. 1983. An improved strategy for rapid direct sequencing of both strands of long DNA molecules cloned in a plasmid. Nucl. Acids Res. 11: 5521-5540.

Guo, L. H., P. P. Stepien, J. Y. Tso, R. Brousseau, S. Narang, D. Y. Thomos, and R. Wu. 1984. Synthesis of human insulin gene VIII: Construction of expression vectors for fused proinsulin production in Escherichia coli. Gene. 29: 251-254.

Hackstadt, T., and H. D. Caldwell. 1985. Effect of proteolytic cleavage of surface-exposed proteins on infectivity of Chlamydia trachomatis. Infect. Immun. 48: 546-551.

Hackstadt, T. 1986. Identification and properties of chlamydial polypeptides that bind eucaryotic cell surface components. J. Bacteriol. 165: 13-20.

Harrison, M. J. 1970. Enhancing effect of DEAE-dextran on inclusion counts of an ovine chlamydia (Bedsonia) in cell culture. Aust. J. Exp. Biol. Med. Sci. 48: 207-213.

Hatch, T. P., I. Allan, and J. H. Pearce. 1984. Structural and polypeptide differences between envelopes of infective and reproductive life cycle forms of chlamydia sp. J. Bacteriol. 157: 13-20.

Hatch, T. P., M. Miceli, and J. E. Sublett. 1986. Synthesis of disulfide-bonded outer membrane proteins during the developmental cycle of Chlamydia psittaci and Chlamydia trachomatis. J. Bacteriol. 165:379-385.

Hawley, D. K. and W. R. McClure. 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. Nuc. Acids Res. 11: 2237-2249.

Henikoff, S. 1984. Unidirectional digestion with Exonuclease III creates targeted breakpoints for DNA sequencing. Gene. 28: 351-359.

Holmes, K. K. 1981. The Chlamydia epidemic. JAMA. 245: 1718-1723.

Horn, J. E., M. L. Hanner, S. Falkow, and T. C. Quinn. 1986. Detection of Chlamydia trachomatis in tissue culture and cervical scrapings by in situ DNA hybridization. J. Infect. Dis. 153: 1155-1159.

Joseph, T., F. E. Nano, C. F. Garon, and H. D. Caldwell. 1986. Molecular characterization of Chlamydia trachomatis and Chlamydia psittaci plasmids. Infect. Immun. 51: 699-703.

Jenkin, H. M. 1966. The continuous passage of agents of trachoma in cell culture. J. Infect. Dis. 116: 390-399.

Kaul, R., and W. M. Wenman. 1985. Cloning and expression in Escherichia coli of a species-specific Chlamydia trachomatis outer membrane antigen. FEMS Microbiol. Lett. 27: 7-12.

Kaul, R., K. L. Roy, and W. M. Wenman. 1987. Cloning, expression, and primary structure of a Chlamydia trachomatis binding protein. J. Bacteriol. 169: 5152-5156.

Kuo, C. C., S. P. Wang, and J. T. Grayston. 1972. Differentiation of TRIC and LGV organisms based on enhancement of infectivity by DEAE-dextran in cell culture. J.Infect. Dis. 125: 313-317.

Kuo, C. C., S. P. Wang, and J. T. Grayston. 1973. Effect of polycations, polyanions and neuraminidase on the infectivity of trachoma-inclusion conjunctivitis and lymphogranuloma venereum organisms in HeLa cells. Infect. Immun. 8: 74-79.

Kuo, C. C. and J. T. Grayston. 1976. Interaction of Chlamydia trachomatis organisms and HeLa 229 cells. Infect. Immun. 13:1103-1109.

Laemlli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227: 680-685.

Legerski, R. S. and D. L. Robberson. 1985. Analysis and optimization of recombinant DNA joining reactions. J. Mol. Biol. 181: 297-312.

Levy, N. J. 1979. Wheat germ agglutinin blockage of chlamydial attachment sites: Antagonism by N-acetyl-D-glucosamine. Infect. Immun. 25: 946-953.

Levy, N. J., and J. W. Moulder. 1982. Attachment of cell walls of Chlamydia psittaci to mouse fibroblasts (L cells). Infect. Immun. 37: 1059-1065.

Lewin, B. M. (ed). 1985. Termination and Antitermination. pp206-218. In, Genes II. John Wiley and Sons. New York.

Maniatis, T., E. F. Fritsch, J. Sambrook (eds). 1982. Molecular Cloning: A Laboratory Manual. New York. Cold Spring Harbor Laboratory.

Manire, G. P., and A. Tamura. 1967. Preparation and chemical composition of the cell walls of mature infectious forms of meningopneumonitis organisms. J. Bacteriol. 94: 1178-1183.

Matsumoto, A., and G. P. Manire. 1973. Electron microscope observations on the effects of polymixin B sulfate on cell walls of Chlamydia psittaci. J. Bacteriol. 101: 278-285.

McClenaghan, M., A. J. Herring, and I. D. Aitken. 1984. Comparison of Chlamydia psittaci isolates by DNA restriction endonuclease analysis. Infect. Immun. 45: 384-389.

Morrisey, J. H. 1981. Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. Anal. Biochem. 117: 307-310.

Nano, F. E., P. A. Barstad, L. W. Mayer, J. E. Coligan, and H. D. Caldwell. 1985. Partial amino acid sequence and molecular cloning of the encoding gene for the major outer membrane protein of Chlamydia trachomatis. Infect. Immun. 48: 372-377.

Nano, F. E., and H. D. Caldwell. 1985. Expression of the chlamydial genus-specific lipopolysaccharide epitope in Escherichia coli. Science. 228: 742-744.

Narita, T., P. B. Wyrick, and G. P. Manire. 1976. Effect of alkali on the structure of cell envelopes of Chlamydia psittaci elementary bodies. J. Bacteriol. 125: 300-307.

Newhall, W. J., B. Batteiger, and R. B. Jones. 1982. Analysis of the human serological response to proteins of Chlamydia trachomatis. Infect. Immun. 38: 1181-1189.

Newhall, W. J., and R. B. Jones. 1983. Disulfide-linked oligomers of the major outer membrane protein of chlamydiae. J. Bacteriol. 154: 998-1001.

Newhall, W. J. 1987. Biosynthesis and disulfide cross-linking of outer membrane components during the growth cycle of Chlamydia trachomatis. Infect. Immun. 55: 162-168.

Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligodeoxyribonucleotide-directed mutagenesis. Gene. 26: 101-106.

Palmer, L., and S. Falkow. 1986a. A common plasmid of Chlamydia trachomatis. Plasmid. 16: 52-62.

Palmer, L., and S. Falkow. 1986b. Characterization of cloned genes from Chlamydia trachomatis. pp. 91-95. In, Microbiology '86. American Society for Microbiology.

Palmer, L., S. Falkow, and L. Klevan. 1986. 16S ribosomal RNA genes of Chlamydia trachomatis. pp. 89-92. In, D.

Oriel, G. Ridgway, J. Schachter, D. Taylor-Robinson and M. Ward (eds). Chlamydial Infections. Cambridge University Press. Cambridge.

Pickett, M. A., M. E. Ward, and I. N. Clarke. 1987. Complete nucleotide sequence of the major outer membrane protein gene from Chlamydia trachomatis serovar L1. FEMS Microbiol. Lett. 42: 185-190.

Platt, T., and D. G. Bear. 1983. Role of RNA polymerase, rho factor, and ribosomes in transcription termination. pp. 123-161. In, J. Beckwith, J. Davies, and J. A. Gallant (eds). Gene Function in Prokaryotes. Cold Spring Harbor, New York.

Pribnow, D. 1975a. Nucleotide sequence of and RNA polymerase binding site at an early T7 promoter. Proc. Nat. Acad. Sci. U.S.A. 72:784-789.

Pribnow, D. 1975b. Bacteriophage T7 early promoters: Nucleotide sequences of two RNA polymerase binding sites. J. Mol. Biol. 99: 419-443.

Rata, T. R., and R. L. Nichols. 1971. Infection of cell cultures by trachoma agent: Enhancement by DEAE-dextran. J. Infect. Dis. 124: 419-421.

Reznikoff, W. S. and W. R. McClure. 1986. E. coli promoters. pp. 1-33. In W. S. Reznikoff and L. Gold (eds). Maximizing Gene Expression. Butterworths. New York.

Roberts, J. W. 1969. Termination factor for RNA synthesis. Nature. 224:1168-1174.

Rosenberg, M. and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 13:319-353.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. U.S.A. 74:5463-5467.

Schachter, J. 1978. Chlamydial Infections. New Eng. J. Med. 298: 428-435, 490-495, and 540-549.

Schachter, J. 1985. Overview of Chlamydia trachomatis infection and the requirements for a vaccine. Rev. Infect. Dis. 7: 713-716.

Schaller, H., C. Gray, and K. Herrmann. 1975. Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. Proc. Nat. Acad. Sci. U.S.A. 72: 737-741.

Seeburg, P. H., C. Nusslein, and H. Schaller. 1977. Interaction of RNA polymerase with promoters from bacteriophage fd. Eur. J. Biochem. 74:107-113.

Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc. Nat. Acad. Sci. U.S.A. 71:1342-1346.

Soderlund, G. and E. Kihlstrom. 1983. Effect of methylamine and monoansylcadaverine on the susceptibility of McCoy cells to Chlamydia trachomatis infection. Infect. Immun. 40: 534-541.

Stefano, J. E., and J. D. Gralla. 1982. Mutation-induced changes in RNA polymerase-lac p promoter interactions. J. Biol. Chem. 257:13924-13929.

Stephens, R. S., C. C. Kuo, G. Newport, and N. Agabian. 1985. Molecular cloning and expression of Chlamydia trachomatis major outer membrane protein antigens in Escherichia coli. Infect. Immun. 47: 713-718.

Stephens, R. S., G. Mullenbach, R. Sanchez-Pescador, and N. Agabian. 1986. Sequence analysis of the major outer membrane protein gene from Chlamydia trachomatis serovar L2. J. Bacteriol. 168: 1277-1282.

Stormo, G. D., T. D. Schneider, and L. Gold. 1982. Characterization of translational initiation sites in E. coli. Nucl. Acids Res. 10: 2971-2996.

Stormo, G. D. 1986. Translation initiation. pp. 195-224. In W. S. Reznikoff and L. Gold (eds). Maximizing Gene Expression. Butterworths. New York.

Storz, J., and L. A. Page. 1971. Taxonomy of the chlamydiae: Reasons for classifying organisms of the genus Chlamydia, family Chlamydiaceae, in a separate order, Chlamydiales ord. nov. Int. J. Syst. Bacteriol. 21: 332-334.

Struhl, K. 1983. Direct selection for gene replacement events in yeast. Gene. 26: 231-242.

Struhl, K. 1985. A rapid method for creating recombinant DNA molecules. Biofeedback. 3: 452-453.

Sweet, R. L., J, Schachter, and M. O. Robbie. 1983. Failure of B-lactam antibiotics to endometrium despite apparent clinical cure of acute salpingitis. J. Am. Med. Assoc. 250: 2641-2645.

Tabor, S. and C. C. Richardson. 1987. DNA sequence analysis

with a modified bacteriophage T7 DNA polymerase. Proc. Nat. Acad. Sci. U.S.A. 84: 4767-4771.

Takanami, M., K. Sugimoto, H. Sugisaki and T. Okamoto. 1976. Sequence of promoter for coat protein of bacteriophage fd. Nature. 260:297-302.

Tamura, A., and G. P. Manire. 1967. Preparation and chemical composition of cell membranes of developmental reticulate forms of meningopneumonitis organisms. J. Bacteriol. 94: 1184-1188.

Tamura, A., and G. P. Manire. 1968. Cytochrome C reductase activity of meningopneumonitis organisms at different stages of development. Proc. Soc. Exp. Biol. Med. 129: 390-393.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. PNAS. 76: 4350-4354.

Ward, M. E., and A. Murray. 1984. Control mechanisms governing the infectivity of Chlamydia trachomatis for HeLa cells: Mechanisms of endocytosis. J. Gen. Microbiol. 130: 1765-1780.

Weiss, E., and H. R. Dressler. 1960. Centrifugation of rickettsia and viruses onto cells and its effect on infection. Proc. Soc. Exp. Biol. Med. 103:691-695.

Wenman, W. M., and M. A. Lovett. 1982. Expression in E. coli of Chlamydia trachomatis antigen recognized during human infection. Nature. 296: 68-70.

Wenman, W. M., and R. U. Meuser. 1986. Chlamydia trachomatis elementary bodies possess proteins which bind to eucaryotic cell membranes. J. Bacteriol. 165: 602-607.

Westrom, L. Incidence, prevalence, and trends of acute pelvic inflammatory disease and its consequences in industrialized countries. 1980. Am. J. Obstet. Gynecol. 138: 880-892.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene. 33: 103-119.

Young, R. A. and R. W. Davis. 1983a. Efficient isolation of genes using antibody probes. Proc. Nat. Acad. Sci. U.S.A. 80: 1194-1198.

Young, R. A. and R. W. Davis. 1983b. Yeast RNA polymerase II

genes: Isolation with antibody probes. Science. 222: 778-782.

Young, R. A. and R. W. Davis. 1985. Immunoscreening lambda gt11 recombinant DNA expression libraries. In J. K. Setlow, and A. Hollaender (eds). Genetic Engineering: Principles and Methods, vol. 7. Plenum Publishing Corp. New York.

Zuker, M., and P. Stiegler. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. 9:133-140.

Appendix A: Growth Media and Solutions

Chlamydial Cultures

Hank's Balanced Salt Solution (10 x stock) (HBSS)

80 g NaCl
10 g Dextrose
4 g KCl
2 g $MgSO_4 \cdot 7H_2O$
0.6g KH_2PO_4
0.48g Na_2HPO_4

Dissolve each in 900ml ddH_2O. Add 100 ml 0.2% phenol red and 1.4g $CaCl_2.2\ H_2O$ (for 1.1g $CaCl_2$)

Eagle's Minimal Essential Medium (MEM)

700ml 1 x Hank's BSS
16ml 50 x amino acids (Gibco)
20ml 0.2N NaOH
8ml 100 x Vitamins (Gibco)
8ml 100 x 200M Glutamine (Gibco)
80ml Fetal Calf Serum (Gibco)
4-6ml 5.6% $NaHCO_3$
8ml 100ug/ml Vancomycin (Eli Lilly)
8ml 25ug/ml Gentamycin (Schering)

Sucrose, Phosphate, and Glutamate (SPG)

60g Sucrose
0.415g KH_2PO_4
0.976g Na_2HPO_4
0.576g Glutamic acid

Dissolve in 800ml ddH_2O.
pH to 7.4-7.6
Filter sterilize.

E. coli cultures

LB (Luria-Bertani) Media

10g Tryptone
5g Yeast extract
10g NaCl

In one liter ddH_2O. Adujust pH to 7.5 with sodium hydroxide. Autoclave to sterilize.

2 YT Media

16g Tryptone
10g Yeast Extract
5g NaCl

Dissolve in one liter ddH_2O. Autoclave to sterlize.

M9 Medium

6g Na_2HPO_4
3g KH_2PO_4
0.5g NaCl
1g NH_4Cl
adjust to pH 7.4, autoclave, cool and add:

2ml 1M $MgSO_4$
10ml 20% glucose
0.1ml 1M $CaCl_2$

For agar plates, 20 g agar was added before autoclaving.

Ampicillin Stock

Ampicillin (sodium salt) was dissolved in ddH_2O to 25mg/ml. Sterilized by filtration. Aliquoted and stored at -20C. Added to cooled media to final concentration of 50 ug/ml.

IPTG, 2.0% Solution, (Sigma)

0.2g dissolved in 10mls ddH_2O. Filter sterilize.

X-gal, 2.0% Solution, (Boeringer Mannehiem)

0.2g dissolved in 10mls N,N'-dimethyl formamide.

Appendix B: Nucleic Acid Extraction Solutions and Buffers

TE/TES

10mM Tris-HCl (pH 7.5)
1mM EDTA (pH 8.0)

For TES
add NaCl to 10mM

Chlamydial Lysis Solution

50mM Tris-HCl (pH 7.5)
20mM EDTA
20% Sucrose (wt/vol)
0.7% Sarkosyl

add Proteinase K to final 200ug/ml

SM

5.8g NaCl
2.0g $MgSO_4.7H_2O$
50ml 1M Tris-HCl (pH 7.5)
2ml 2% gelatin

Suspend in 1 liter ddH_2O

E. coli Lysis Solution

25% Sucrose (wt/vol)
50mM Tris-HCl (pH 7.5)
40mM EDTA

add Lysozyme to final 2mg/ml

Triton X-100 Solution

1.0% Triton X-100
50mM Tris-HCl (pH 7.5)
50mM EDTA

Appendix C: Restriction Endonuclease Buffers

10x High Salt

1M NaCl
500mM Tris-HCl (pH 7.5)
100mM $MgCl_2$
10mM DTT (dithiothreitol)

10x Medium Salt

500mM NaCl
100mM Tris-HCl (pH 7.5)
100mM $MgCl_2$
10mM DTT

10X Low Salt

100mM Tris-HCl (pH 7.5)
100mM $MgCl_2$
10mM DTT

SmaI Buffer

200mM KCl
100mM Tris-HCl (pH 7.5)
100mM $MgCl_2$
10mM DTT

ClaI Buffer

BRL React 1 buffer
500mM Tris-HCl (pH 8.0)
100mM $MgCl_2$

Appendix D: Electrophoresis Buffers and Solutions

Agarose Gels for DNA

Loading Dye

0.25% bromophenol blue
0.25% xylene cyanol
15% Ficoll in ddH_2O

50x ELFO Buffer

2M Tris-HCl
50mM EDTA pH 7.9

Dilute to 1x ELFO for gel and running buffer. This buffer is used instead of the common Tris-acetate, due to its minimal salt concentration; allows the use of the gel ligation procedure.

Sequencing Gels

10x Tris-HCl-Borate-EDTA (TBE), pH 8.3

108gm Tris Base
55gm Boric acid
9.3gm Na_2-EDTA
Bring volume up to 1 liter with ddH_2O

40% Acrylamide Stock
38gm Acrylamide (BIO-RAD Electrophoresis grade)
2gm BIS-acrylamide
Bring volume up to 100ml with ddH_2O

6% Gel Mix
75ml 40% Acrylamide
250gm Urea (BRL Pure Grade)
50ml 10x TBE
175ml ddH_2O

10% Ammonium Persulfate (APS)
1gm APS in 9ml ddH_2O
Store at 4°C for up to 2 weeks.

SDS-PAGE Gels

Running Buffer, pH 8.3
15.0gm Tris Base
72.0gm Glycine
5gm Sodium dodecyl sulfate (SDS)
Bring volume to 5 liters with ddH_2O

Stacking Gel Buffer
6.0gm Tris Base
Bring volume to 100mls with ddH_2O, adjust pH to 6.8 with concentrated HCl.

Stacking Gel Mix

30% acrylamide-:	0.83ml
stacking buffer:	1.25ml
10% SDS--------:	50ul
TEMED----------:	5ul
ddH_2O----------:	2.6ml
APS------------:	0.25ml

Discontinuous Gel Buffer
36.3gm Tris Base
Bring pH down to 8.8 and volume up to 100mls with ddH_2O

Discontinuous Gel Mix

	10%	12.5%
30% acrylamide------:	5.0ml	6.25ml
discontinuous buffer:	3.7ml	3.7ml
10% SDS----------:	0.15ml	0.15ml
TEMED----------:	15ul	15ul
ddH_2O----------:	4.1ml	5.4ml
APS----------:	0.75ml	0.75ml

30% Acrylamide
30gm Acrylamide
0.8gm BIS-acrylamide
Bring volume to 100mls with ddH_2O

Solubilizing Buffer
0.1M Tris-HCl (pH 6.8)
2.5% SDS
20% 2-mercapto ethanol
0.001% bromophenol blue

Appendix E: DNA Modifying Enzyme Buffers and Solutions

10x Ligation Buffer

600mM Tris-HCl (pH 7.2)
100mM $MgCl_2$
100mM DTT
10mM ATP

10x Nick Translation Buffer

500mM Tris-HCl (pH 7.2)
100mM $MgSO_4$
1mM DTT
500ug/ml BSA

10x Annealing Buffer

100mM Tris-HCl, pH 7.4
100mM $MgCl_2$

10x T4 DNA Polymerase Buffer

330mM Tris-acetate, pH 7.8
660mM K-acetate
100mM Mg-acetate

10x Exonuclease III Buffer

660mM Tris-HCl, pH 8.0
50mM $MgCl_2$
770mM NaCl
100mM DTT

10x S1 Nuclease Buffer

0.5M Na-acetate, pH 4.0
0.5M NaCl
60mM $ZnSO_4$

Appendix F: Hyridization Buffers and Solutions

Blot Buffer (for Western blotting), pH 8.3
3.03gm Tris-HCl (25mM)
14.41gm Glycine (192mM)
200ml Methanol
Bring volume up to 1 liter with ddH_2O

Phosphate Buffered Saline (PBS), 25x Solution

27.4g Na_2HPO_4
7.87g $NaH_2PO_4.H_2O$
Dissolve in 1000ml ddH_2O.

For working solution;
80ml 25x
17g NaCl
Dissolve in 2000ml ddH_2O.
Adjust pH to 7.2-7.4.
May be autoclaved.

Tris Buffered Saline (TBS)

9.0g NaCl
1.21g Tris Base
Dissolve in 1000mls ddH_2O

DNA-DNA Hybridization Solution

0.1gm SDS
1.0gm Dextran Sulfate
2.0ml 5M NaCl
Dissolve in 10mls ddH_2O.