MOLECULAR STUDIES ON IgA1 PROTEASE AND NEURAMINIDASE FROM *Streptococcus pneumoniae*

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Miguel María Cámara García
Department of Microbiology
University of Leicester

December 1992

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

A MIS PADRES

(To my parents)

" Pneumococcus is altogether an amazing cell. Tiny in size, simple in structure, frail in make-up, it possesses physiological functions of great variety, performing biochemical feats of extraordinary intricacy and, attacking man, sets up a stormy disease so often fatal that it must be reckoned as one of the foremost causes of human death. "

Benjamin White (1938)

STATEMENT

The accompanying thesis submitted for the degree of Ph.D. entitled "Molecular Studies on IgA1 Protease and Neuraminidase from *Streptococcus pneumoniae*" is based on work conducted by the author in the Department of Microbiology of the University of Leicester mainly between September 1988 and December 1992.

All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed:..... Date:16/12/92

ABSTRACT

Streptococcus pneumoniae is an important human pathogen responsible for lower respiratory tract infections, septicemia and meningitis. It produces IgA1 protease and neuraminidase which may be of relevance in the pathogenesis of pneumococcal disease.

A LambdaEMBL301 genomic library from *S. pneumoniae* strain R36A was made. A stable clone expressing IgA1 protease activity was isolated from this library. Preliminary hybridization studies suggested the pneumococcal authenticity of this IgA1 protease recombinant.

From the same pneumococcal genomic library a gene expressing neuraminidase activity was also isolated. Restriction mapping and DNA hybridization studies revealed that this gene is distinct from another pneumococcal neuraminidase gene previously cloned. Both genes were found simultaneously present in several different pneumococcal serotypes.

The DNA sequence of the pneumococcal neuraminidase gene plus flanking regions was determined. The predicted amino acid sequence of this neuraminidase corresponds to a mainly hydrophilic protein. This neuraminidase presents a putative signal peptide near the N-terminus. It also contains a typical surface anchor motif in the C-terminus which is common to other surface proteins from Gram-positive organisms. N-terminal to this motif there is a region containing twenty amino acids tandemly repeated three times. Regions of homology between this protein and other neuraminidases were found.

Electron microscopy studies using immunogold staining showed the presence of pneumococcal neuraminidase associated to the surface of the pneumococcus.

Analysis of the neuraminidase DNA flanking regions revealed the presence of DNA repetitive elements (BOXES). Databases analysis showed the presence of some of these elements in the vicinity of pneumococcal genes related with virulence and competence.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Professor Graham Boulnois for his help, interest, reassurance and for his constructive criticism during the completion of this work.

Words would not be enough to thank Dr Tim Mitchell for his supervision, encouragement and invaluable support, in particular, during all those darkest moments.

In addition I would like to thank Dr Peter Andrew for his motivation, excellent advice and sincerity.

The help and friendship of the members of the Department of Microbiology in the University of Leicester has been of great support to me. They have provided many lighthearted moments.

I would also like to thank Dr James Paton from Australia for the donation of a neuraminidase clone, antibodies and bacterial strains that have been the basis of some of the experiments described in this thesis. Thanks also to Stephan Hyman for his help in the electron microscopy studies, and to Steve Gordon for his generous assistance with the CLUSTALV analysis.

My friends have given me continuous and infinite support and have always been there to cheer me when most needed. I am indebted to them always.

My most precious thank you goes to my parents and Yolanda. They have given me the strength and support to overcome the worst moments during the formation of this thesis.

ABBREVIATIONS

BHI	Brain heart infusion
bp	Base pair
BSA	Bovine serum albumin
CETAB	Hexadecyltrimethylammonium bromide
CFU	Colony forming units
CIP	Calf intestinal phosphatase
CSF	Cerebral spinal fluid
DTT	Dithiothreitol
EDTA	Diaminoethantetra-acetic acid
Fabα	Antigen binding portion of IgA
Fcα	Crystallisable fraction of IgA
HRP	Horse radish peroxidase
IPTG	Isopropyl-ß-D-thiogalactosylpyranoside
J	Joining chain
Kb	Kilobase pair
Kd	Kilo daltons
MU	4-methylumbelliferone
MUAN	2'-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid
NANA	N-acetylneuraminic acid
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PMN (PMNL)	Polymorphonuclear leucocytes
SC	Secretory component
SDS	Sodium dodecyl sulphate
sIgA	Secretory immunoglobulin A
TEMED	N,N,N,N'-Tetramethylethylethylenediamine
X-Gal	5-bromo-4-chloro-3-indoyl-ß-D-galactopyranoside
UV	Ultraviolet
V	Volts
W	Watts

INDEX

PART III: CLONING OF AN IgA1 PROTEASE GENE FROM *Streptococcus pneumoniae*

PART IV: CLONING OF A NEURAMINIDASE GENE FROM *Streptococcus pneumoniae*

PART V : DNA SEQUENCE DETERMINATION OF THE NEURAMINIDASE GENE IN LambdaEMBL301-Neu1 CLONE

PART I

INTRODUCTION

1.1. *Streptococcus pneumoniae*

Streptococcus pneumoniae is a Gram-positive bacterium from the family of Streptococcaceae, genus Streptococcus. It grows characteristically in pairs, appearing often as lancet-shaped diplococci and short chains. It has a mucoid colonial morphology and α-hemolytic properties when grown on the surface of blood agar. It ferments inulin, autolyses when exposed to detergents (bile solubility) and is sensitive to ethylhydrocupreine HCl (optochin) [Austrian 1984].

The history of this bacterium started more than one hundred years ago, in 1880, when it was simultaneously first isolated from saliva, by Stenberg in the United States [Stenberg 1881] and Pasteur in France [Pasteur 1881]. Over the next ten years this organism was shown to be an important pathogen of man [reviewed by Austrian 1984] . Still, after so many years, pneumococcal infection persists as a major cause of pneumonia, meningitis and otitis media. Pneumococcal pneumonia is estimated to be responsible for 150,000 - 570,000 cases per year in the USA, that is 10 - 25 % of all

pneumonias with a mortality rate of 5 % [Williams *et al* 1988]. Also pneumococcal bacteraemia and meningitis are responsible respectively for 15 - 19 and 1 - 2 infections per annum per 100,000 population with a fatality rate of 40 and 55 % respectively [Lee *et al* 1991].

Central to the epidemiology of pneumococcal disease is the carriage of this organism in the upper respiratory tract by a significant number of healthy individuals: 59 % of the population carry the organism by the twelfth postnatal day [Riley & Douglas 1981]. It has been said about the pneumococcus: "Its ecological niche is the nasopharynx; and, when localized to that area, its enjoys a commensal relationship with the subject it inhabits" [Austrian 1986]. When this balanced relationship is lost the organism can cause invasive disease. There are several factors that contribute to this: the extremes of age, immunological deficiencies, metabolic abnormalities and chronic respiratory diseases [Gillespie 1989].

There are measures for both prophylaxis and treatment of pneumococcal disease, but both of these have shortcomings. Prophylaxis involves the use of a twenty three valent polysaccharide vaccine, but this has some limitations (see below). Treatment of pneumococcal disease involves the use of antibiotics which has the drawback of a high increase in antibiotic resistance of the pneumococcus during the last few years [Klugman 1990]. Therefore, there is no completely effective way of preventing or eliminating pneumococcal disease.

1.1.1. *Streptococcus pneumoniae* IN DISEASE

Streptococcus pneumoniae is one of the major causes of pneumonia, meningitis, otitis media and bacteraemia. The understanding of the pathogenesis of this disease has been a major challenge for many scientists working in this field, but after many decades it still remains unclear.

I. PNEUMONIA

The pneumococcus accounts for up to 70 % of the total number of cases of human pneumonia in the United States [Farr & Mandell 1983].

Pneumococcal pneumonia is believed to arise as a result of aspiration of virulent pneumococci from the upper respiratory tract into a previously injured lower respiratory tract. Damage of the respiratory tract by different factors such as viral infection, or by chemical agents, can cause a pulmonary oedema, generating a lowering of the efficiency of clearance mechanisms, thus allowing bacterial multiplication and development of pneumonitis [Harford & Hara 1950].

Studies done by intratracheal inoculation of rats with pneumococci reveals pulmonary changes that closely resemble those of human pneumococcal pneumonia. The infection starts with a lesion consisting of alveoli filled with fluid and bacteria. Soon after, numerous PMNs migrate into the alveoli phagocytosing the pneumococci. This is followed by the appearance of macrophages which remove the debris of the infection, resulting in the ultimate resolution of the process, with the restoration of the lung to its normal state [Austrian 1981; Yoneda & Coonrod 1980].

Complement has been shown to be involved in the clearance of pneumococci from the lung by opsonizing bacteria and generating chemotactic factors that can increase the number of PMNs in this organ [Guckian *et al* 1980; Coonrod and Yoneda 1982]. Pneumococci may enter the blood stream as a result of tissue damage and increase of permeability at the site of infection, leading to systemic bacteraemia [Coonrod & Yoneda 1981].

II. MENINGITIS

The pneumococcus persists as one of the three principal causes of bacterial meningitis with a high fatality rate [Wenger *et al* 1990]. Meningitis is defined as an inflammation of the leptomeninges, which are the two membranes that surround the brain and the spinal column. The gap between the two layers is filled by cerebrospinal fluid (CSF). Meningitis arises after infection of CSF which, in humans, usually follows bacterial colonization of the respiratory tract [Moxon 1981]. This is commonly associated with persistence of bacteremia [Moxon 1977], especially when endocarditis has developed [reviewed by Austrian 1984]. Meningitis may also result as an

extension of the infection from the middle ear or paranasal sinuses to the bony structures of the skull. Alternatively, after basal fractures of the skull, the organisms colonizing the upper respiratory tract can gain direct access to the meninges by reaching the subarachnoid space [reviewed by Austrian 1984].

After bacterial colonization of CSF, there is a vigorous inflammatory response, possibly caused by pneumococcal cell wall components [Tuomanen *et al* 1985]. This response involves a series of reactions in CSF that results in the generation of chemotactic factors, appearance of complement-mediated opsonic activity, immunoglobulins and PMNs. All these factors improve the local host defense mechanisms [Zwahlen *et al* 1982] although they have some detrimental effects, as they can cause oedema. Among of these effects are the loss of neurological function [Dodge & Swartz 1965] and the impairment of cerebral blood flow [McMenamin & Volpe 1984], generated by a heightening in intracranial pressure. These two effects can rapidly cause death.

III. OTITIS MEDIA

Middle ear infection is one of the most common illnesses in young children. The pneumococcus is the cause of 50-67 % of bacterial otitis media in the United States [Austrian 1984]. It has been shown in a chinchilla model that influenza virus infection of animals already colonized by intranasal inoculation with *Streptococcus pneumoniae*, resulted in a substantial increase in the incidence of otitis media [Giebink 1981]. The influenza infection generates a decrease in eustachian tube pressure causing obstruction of normal drainage of the tube which may facilitate the colonization by the pneumococcus from the respiratory tract. It has been shown that cell wall components released during bacterial lysis may contribute to chronic otitis media with effusion in humans [Ripley-Petzoldt *et al* 1988]. The presence of anti-pneumococcal capsular antibodies in middle ear effusions during acute otitis media, seems to be beneficial for the resolution of the pneumococcal middle ear infection [Karjalainen *et al* 1991].

IV. BACTERAEMIA

Pneumococcal bacteremia has a lower incidence than pneumonia or meningitis but a 40 % fatality rate [Lee *et al* 1991]. Several routes exist for pneumococci to gain access to the blood. First, bacteraemia can occur as a sequel to invasion of lower respiratory tract. Second, it may also arise as a complication of meningitis of cephalic origin in which the meninges are infected by direct extension from a focus of infection in the head. In this case the pneumococcus enters the blood system via the venous sinuses of the head [Austrian 1984]. Finally, pneumococci may spread directly into the blood vessels via the highly vascular nasopharyngeal mucosa, or alternatively, via lymphatic spread [Moxon 1981]. Termination of bacteraemia and recovery correlates with the spontaneous appearance of type specific antibodies which have an opsonic effect via classical pathway complement activation [Brown *et al* 1982].

1.1.2. VIRULENCE DETERMINANTS

Infections with *Streptococcus pneumoniae* cause significant morbidity and mortality. This organism produces a capsule that has been very well studied and which is involved in virulence of the pneumococcus (see below). Recently pneumolysin, a protein toxin made by these bacteria, has been described as a virulence factor. This organism also possesses several structures and proteins, some of them with enzymatic activity, that could be involved in virulence and which will be described later in this section.

I. CAPSULE

Streptococcus pneumoniae possesses a capsule composed of acidic and neutral polysaccharide. This polysaccharide forms a hydrophilic layer on the surface of the organism [Austrian 1981]. The capsule is important in disease since decapsulated organisms show a reduction in virulence of several orders of magnitude [Avery & Dubos 1931]. Early studies showed that the virulence of the pneumococcus can be influenced by the quantity of the polysaccharide [MacLeod & Krauss 1950].

The capsule is produced maximally by the pneumococcus in the logarithmic

phase of growth. Most capsular polysaccharides are covalently bound to the peptidoglycan [Sørensen *et al* 1990]. There are 83 serologically defined pneumococcal types based on the polysaccharides, although there is an extensive cross-antigenicity and cross-immunogenicity among many types within a group [Lee *et al* 1981].

Capsule presents a mechanical barrier preventing complement, deposited on the cell wall, from being opsonically active. Capsular polysaccharide has also been found to inhibit phagocytosis of opsonised bacteria by PMNLs [reviewed by Gillespie 1989]. The pneumococcal capsule has recently been shown to be involved in virulence by using a capsule negative mutant of a type 3 pneumococcus, constructed by transposon mutagenesis, for challenge experiments in mice. The LD_{50} increased from 1 CFU in the wild type to $>5 \times 10^7$ CFU in the acapsular mutant [Watson & Musher 1990].

Detection of capsular polysaccharide in serum by ELISA can be used for diagnosis for pneumococcal infection although the level of antigenaemia varies independently to the severity of the infection [Schaffner *et al* 1991].

Although capsule is an essential virulence determinant of the pneumococcus, it is not toxic to animals or humans. This fact, together with the existence of 83 different capsular serotypes in *S.pneumoniae*, most of them virulent, has forced many scientists to look for other factors that, being involved in virulence, are also present in all the different serotypes. These factors can be classified into two different groups: protein virulence factors and non-protein virulence factors. These groups are described below.

II. PROTEIN VIRULENCE FACTORS

a/ Pneumolysin

Pneumolysin is a thiol-activated toxin and a member of a family of toxins produced by a diverse genera of gram positive bacteria, that share common physical and biological properties [Smyth & Duncan 1978].

Pneumolysin has a cytoplasmic location in the peumococcus [Johnson 1977] and is believed to be released by autolysis [Walker *et al* 1987b]. It binds to eukaryotic membranes through cholesterol and is thought to oligomerize to form transmembrane channels that lead to lysis of cells [Bhakdi *et al* 1985]. Low doses of pneumolysin inhibit respiratory burst, chemotaxis and antimicrobial activities of human PMNs and macrophages [Mandoskar *et al* 1985, Paton & Ferrante 1983]. A higher dose of the toxin activates the classical pathway of complement and considerably reduces its opsonic activity for the pneumococcus [Paton *et al* 1984].

Mice immunized with purified pneumolysin survived significantly longer when challenged with virulent *S. pneumoniae* [Paton *et al* 1983]. This was the first direct evidence for the involvement of pneumolysin in pneumococcal pathogenicity. Also, after the cloning and sequencing of the pneumolysin gene [Walker *et al* 1987], a pneumolysin negative isogenic mutant has been used to show reduced virulence in mouse challenge experiments by an increase in LD_{50} between 10^1 and 10^2 fold. This mutant was also cleared from the blood earlier than the wild-type [Berry *et al* 1989a]. These facts indicate a role in virulence for pneumolysin.

Pneumolysin has been shown to cause slowed ciliary beating and disruption of the surface integrity of human respiratory epithelium in organ culture. This could perturb the mucociliary defense mechanism of the respiratory epithelium during infection, allowing the bacterium time to proliferate and spread within the respiratory tract [Steinfort *et al* 1989]. The injection of pneumolysin into the apical lobe bronchus of rat lung was associated with the development of a severe lobar pneumonia with histologic features identical to those of pneumococcal infection. This developement may be due to its ability to activate the complement cascade and therefore induce inflammation [Feldman *et al* 1991].

Genetically engineered pneumolysin toxoids have been considered as a potential protein carrier for pneumococcal capsular polysaccharide in a conjugate vaccine since they increase the immunogenicity of the

polysaccharide thus inducing immunological memory to the polysaccharide and the toxin [Mitchell *et al* 1992].

b/ Hyaluronidase

Hyaluronidase is an enzyme which depolymerises hyaluronic acid, a component of connective tissue and the extracellular matrix. It has also been called a spreading factor since it may facilitate spreading of the organism through tissues [Russel & Sherwood 1948]. This enzyme is produced by the pneumococcus [Humphrey 1944], although very little is known about its putative role in virulence.

c/ Surface Protein A

Surface protein A (PspA) is a surface exposed, highly immunogenic protein that is found on all isolates of *Streptococcus pneumoniae* [Crain *et al* 1990]. This protein has been shown to be required for full virulence of the pneumococcus. Immunization with $PspA^+$ pneumococci, but not with their $PspA^-$ isogenic mutants, protects *xid* mice (which only show immunologic responses to proteins but not capsule) from fatal challenge with virulent pneumococci [McDaniel *et al* 1987]. Mice immunized with a purified N-terminal fragment of PspA were protected from fatal pneumococcal challenge [Talkington *et al* 1991].

PspAs from different pneumococcal isolates have variable molecular weights, ranging from 67 Kd to 99 Kd, and have been shown to exhibit antigenic variability. This variation was also observed within strains of the same capsular type [Waltman *et al* 1990]. Despite this variability, enough common epitopes may be present to allow a single, or at least a small number of PspAs to elicit protection against most *Streptococcus pneumoniae* types [Yother *et al* 1991].

d/ Neutrophil elastase inhibitor

This factor has only been studied in the context of pneumococcal pneumonia. During the course of this disease there is a massive influx of PMNLs releasing toxic products, however there is rarely any derangement of alveolar

structure after recovery. This lack of tissue damage may be associated with neutrophil elastase inhibitors. Pneumococci produce two types of these inhibitors. One is a negatively charged substance of low molecular weight that depends on an ionic interaction with the cationic elastase. It is unlikely to be important *in vivo* since it is substantially inhibited by NaCl at physiological concentrations [Gillespie 1989]. The second inhibitor, located outside the cell membrane, has a molecular weight of 140 Kd [Lee *et al* 1985]. It specifically inhibits the elastase from human neutrophils and is not inactivated by NaCl. A pneumococcal strain (type III) that is associated with higher risk of lung necrosis during infection, contains less of the inhibitor per unit cell protein than does a strain (type I) that is rarely implicated in tissue damage [Vered *et al* 1985]. Therefore the neutrophil elastase inhibitor is thought to play a role in pneumococcal infection.

e/ Autolysin

Autolysins are enzymes that hydrolyze the cell wall of Gram-positive bacteria. *Streptococcus pneumoniae* produce these enzymes. The pneumococcal autolysin gene (*lyt*A) has been cloned and sequenced giving a predicted molecular weight for the protein of 35 Kd [García *et al* 1985]. The activity of this enzyme is inhibited by the pneumococcal Forssman antigen [Briles & Tomasz 1973]. The loss of this regulatory antigen results in the activation of the pneumococcal autolysin, causing cell wall degradation [Horne & Tomasz 1985]. A LytA negative mutant has been constructed which shows a normal growth rate but fails to lyse in the stationary phase of growth [Sánchez-Puelles *et al* 1986]. An autolysin negative mutant generated by insertion-duplication mutagenesis showed markedly reduced virulence for mice compared with the wild type. In addition, mice immunized with purified recombinant autolysin survived significantly longer than control mice after intranasal challenge with the pneumococcal strain D39. It has been speculated that autolysin could be responsible for the release, during infection, of potentially lethal toxins from the pneumococcus [Berry *et al* 1989b]. In contrast to these results, Tomasz *et al* (1988) made pneumococcal Lyt$^-$ insertional mutants that showed the same degree of virulence as did the isogenic Lyt$^+$ parent strain. Therefore the role of this enzyme still remains

unclear.

f/ IgA1 protease

IgA1 protease hydrolyzes human IgA1 in the hinge region of the immunoglobulin at a site located in one half of the duplicated octapeptide region [Plaut 1983]. A more detail review of this enzyme is given in Section 1.2.

g/ Neuraminidase

Neuraminidase is an enzyme that cleaves N-acetylneuraminic acid from mucins, glycoproteins and gangliosides [Kelly *et al* 1967, Scalon *et al* 1989]. The current knowledge of this enzyme is reviewed in section 1.3.

III. NON-PROTEIN VIRULENCE DETERMINANTS

a/ C polysaccharide

The C-polysaccharide, common to all the 83 different serotypes, is the major cell wall component of the pneumococcus. Native C polysaccharide is a ribitol teichoic acid containing phosphorylcholine and galactosamine. This antigen is covalently bound to the peptidoglycan via the muramic acid residue [Jennings *et al* 1980].

Not much is known about the role of the C polysaccharide. It has been seen that it can activate alternative complement pathway, because removal of teichoic acid from peptidoglycan diminishes complement activation [Winkelstein & Tomasz 1978]. Also, rabbit polyclonal and mouse monoclonal anti-phosphorylcholine antibodies have been shown to protect mice from fatal challenge [Yother *et al* 1982], although anti-C polysaccharide antibodies from vaccinated humans fails to opsonise capsulated strains [Musher *et al* 1986]. It may be that there is a fundamental difference in the immune response between man and mouse as suggested by Gillespie (1989).

b/ F antigen

The F antigen (Forssman) is an amphiphatic molecule structurally similar to

C-polysaccharide, containing choline with additional covalently attached lipid material. The choline from this antigen represents 15% of the total choline incorporated into the organism [Briles & Tomasz 1973].

The F antigen is found to be uniformly distributed in the pneumococcal plasma membrane anchored to this through the lipid region of the molecule with the C polysaccharide parts exposed on the surface [Briles & Tomasz 1973].

Although the role F antigen plays in the pneumococcus is still unknown, it has been shown to be a powerful inhibitor of the pneumococcal autolysin. All pneumococci are prone to lysis during the stationary phase of growth, a state that is preceded by the release of the F antigen [Horne & Tomasz 1985].

1.1.3. PNEUMOCOCCAL VACCINE.

The history of the pneumococcal vaccine starts in 1911 when, in an attempt to eradicate pneumonia caused by *S. pneumoniae*, Wright made a vaccine of killed pneumococcus without regard to serotype by using very few organisms. This vaccine was not efficient in conferring protection against the pneumococcus [reviewed by Gillespie 1989].

A more elaborate vaccine was created by Lister. This vaccine included cresol-killed organisms of serotypes 1, 2 and 5, that were responsible for 70% of the pneumonias detected in South African miners. With the use of this vaccine the incidence of pneumonia fell by 50% [reviewed by Musher *et al* 1990].

In the late 1940s a 6-valent vaccine of purified pneumococcal polysaccharide became commercially available, but its introduction coincided with the emerging view that penicillin was effective in dealing with pneumococcal infection. Thus, it was taken off the market [reviewed by Bruyn & Furht 1991].

Many years later, in 1977, a 14-valent pneumococcal polysaccharide vaccine

was commercially available which, after six years, was replaced by the one currently used which includes 23 different serotypes [reviewed by Shapiro *et al* 1991]. The use of this vaccine has been recommended especially in the elderly, since they are a high risk population for pneumococcal disease [Mufson *et al* 1991].

The use of this 23-valent pneumococcal vaccine has several limitations. Firstly, it only protects against 23 serotypes out of the 83 described. It is not very efficient in infants (children under 2 years of age), since they exhibit different degrees of responsiveness to different pneumococcal polysaccharides. It is also very inefficient in immunocompromised patients [Bruyn & Furth 1991]. Furthermore, some perfectly healthy individuals tend to respond well and others to respond poorly to all pneumococcal serotypes [Bardardottir *et al* 1990].

This 23 valent vaccine has been designed for the population of the U.S.A. considering the more frequent pneumococcal serotypes in that country. The types included in the 23-valent pneumococcal vaccine are considerably less frequent in Asia than the U.S.. The proportion of these types is 62.9 % in Taiwan, 27.9% in Japan, 79.9% in China, as compared to 88.2 % in the U.S.A. [reviewed by Lee *et al* 1991]. Additionally, although presently it covers most cases in the U.S.A., variation of serotypes with time has been observed [Austrian 1984] which may decrease its effectiveness in that country.

Finally, capsular polysaccharides are T-cell independent type 2 (TI-2) antigens, therefore they are unable to generate immunological memory. Also, anti TI-2 antibody responses are generally not detectable until 18-24 months of age, while T-cell dependent antibody responses can be induced from birth onwards [reviewed by Peeters *et al* 1991].

The shortcomings in the use of the current polysaccharide vaccine, together with the difficulties in the treatment of the pneumococcal disease because of increased antibiotic resistance [Klugman 1990] has opened new lines of

research in the improvement of a vaccine. Initial studies were done by the use of polysaccharides conjugated to protein carriers in immunization experiments against *Haemophilus influenzae* type b (Hib) [Schneerson *et al* 1980]. In these experiments, conjugation to one of several proteins changed the nature of anti-polysaccharide antibody from a T-independent to a T-dependent response. Currently, there are several Hib vaccines using the capsular polysaccharide (polyribosyl phosphate, PRP) conjugated to one of several protein carriers. These vaccines, have been shown to be very effective in the prevention of the disease caused by Hib [Reviewed by Moxon & Rappuoli 1990].

In the case of the pneumococcus, capsular polysaccharide from a type 12F was conjugated to either diphtheria or tetanus toxoid [Lue *et al* 1990; Sarnaik *et al* 1990] resulting in an increase in immunogenicity of the polysaccharide. It was thought that polysaccharide could be conjugated to a protein involved in the virulence of pneumococci to get a response to this factor in addition to the polysaccharide. Pneumolysin has been the protein of choice as it is produced by all pneumococcal serotypes and has been shown to be important in virulence [Berry *et al* 1989]. Conjugation of this toxin to type 19F polysaccharide, which is normally a poor immunogen in children and mice, converted this polysaccharide into an antigen capable of inducing a booster effect and increased immunogenicity. The conjugate also gave modest responses to pneumolysin, a non-type specific antigen [Paton *et al* 1991].

Incorporation of other pneumococcal protein factors, common to all serotypes, to the current vaccine could contribute to the eradication of the pneumococcal disease. Among these factors, IgA1 protease [Pratt & Boulnois 1987] and neuraminidase [Cámara *et al* 1991] should be considered. In the case of neuraminidase, it has been shown that the inactivated enzyme can act as an immunogen, stimulating the production of specific antibodies that can be protective during pneumococcal infection [Lock *et al* 1988]. Therefore, neuraminidase, which is one of the main subjects of this thesis, could be considered as another candidate for a carrier in the development of new pneumococcal vaccines.

1.2. IgA1 PROTEASE

Streptococcus pneumoniae is one of a multitude of organisms producing an enzyme capable of specifically splitting IgA1 into Fcα and Fabα fragments. In order to understand the nature of these enzymes as well as the role they could play in pathogenesis, it is fundamental to review the structure and function of their substrate.

1.2.1. IMMUNOGLOBULIN A STRUCTURE

Immunoglobulin A is a relatively minor component of serum, comprising about 10 - 15 % of the total serum immunoglobulins [Tomasi 1984]. However, this immunoglobulin is the predominant antibody in the secretions bathing mucous membranes [Plaut 1983]. Secretory IgA is produced by mucosal plasma cells, having a more complex structure than the IgA from serum [FIGURE 1.1.]. It primarily consists on a monomer of two heavy chains (alpha chains) linked to two light chains by disulphide bonds, giving a sedimentation coefficient of 7S [Heremans 1974]. Two monomers are linked by a polypeptide (J chain) attached to the penultimate cysteine in the alpha chains, giving the dimeric form found in the secretions [Halpern & Koshland 1970, Mulks 1983]. During transport of the secretory IgA (sIgA) to the

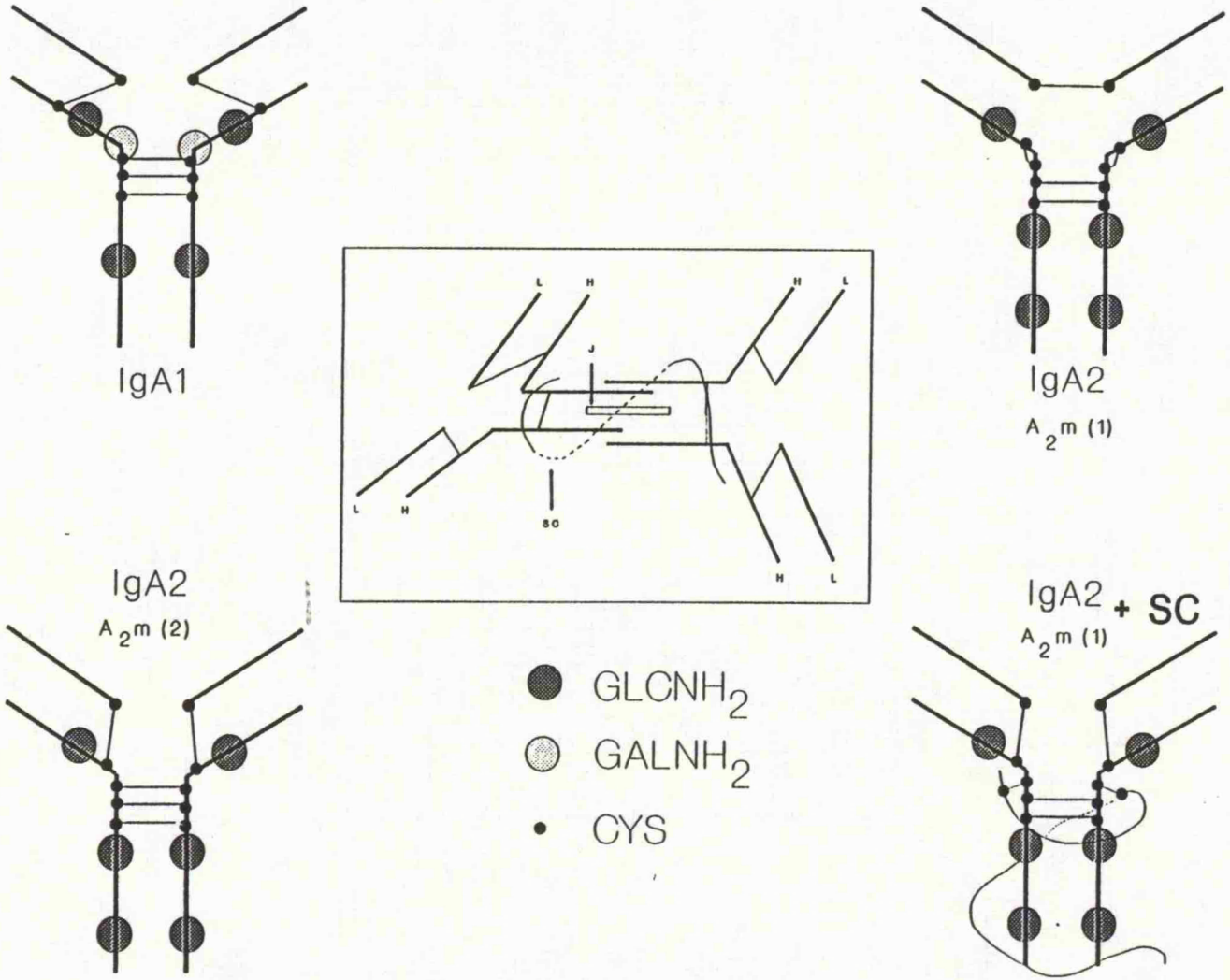


FIGURE 1.1. Schematic illustration of the forms of human IgA1. Central figure represents the the dimeric form with secretory component (SC) and J chain.

surface of the mucosal epithelial cells, a secretory component (SC), consisting on a glycoprotein of 70 Kd, is added. Brandtzaeg proposed that SC, present on the surface of the epithelial cells, serves as a receptor for polymeric IgA [Brandtzaeg 1974]. Attachment of this SC to the dimeric backbone of the sIgA molecule is accompanied by a significant increase in resistance of this backbone to digestion by both trypsin and pepsin [Lindh 1975].

Serum and secretory IgA contains two isotypes of IgA that are designated IgA1 and IgA2 [Vaerman & Heremans 1966, Feinstein & Franklin 1966]. Their structures are similar. The main difference lies in the hinge region, where IgA1 contains a polypeptide repeated twice; thirteen amino acids of that repeat are absent in IgA2 [Wolfenstein-Todel *et al* 1972]. IgA1 accounts for 80-90 % of serum IgA, whereas in external secretions IgA1 constitutes only 50-74 % of the total IgA [Delacroix *et al* 1982]. In contrast to IgA2, IgA1 is enriched in *O*-glycosidically-linked oligosaccharide chains in the hinge region; some of which carry sialic acid [Pierce-Cretel *et al* 1981].

There is a genetic polymorphism described only in the case of IgA2, of which two allotypes are decribed. These allotypes are designated A2m(1) and A2m(2) [Figure 1.1]. It has been shown that the disulphide bond linkage between the heavy and the light chain is absent in the A2m(1) allotype but is present in the A2m(2) allotype which resembles the structure of the IgA1 isotype [Jerry *et al* 1970]. The A2m(1) allotype is prevalent in caucasians and the A2m(2) allotype is predominant in negroids and mongoloids [Kunkel *et al* 1969].

1.2.2. ROLE OF IgA1

1.2.2.1. DIRECT EFFECTOR FUNCTION OF IgA

a/ Adherence

Bacteria colonizing an exposed mucous surface must attach to the surface in order to avoid being washed away by bathing secretions. It has been shown

that preparations of sIgA, isolated from human parotid fluid, specifically inhibited adherence of Streptococcus strains to epithelial cells, thus limiting bacterial colonization [Williams & Gibbons 1972].

Studies with sIgA-sensitized *Salmonella typhimurium* showed this immunoglobulin increases the affinity of this bacterium for the mucus belt of the intestine. Thus, it was postulated that sIgA has the tendency to form complexes with high molecular weight glycoproteins (mucus) in secretions which are likely to potentiate the agglutinating capacity of sIgA antibody and thereby, antigen disposal [Magnusson & Stjernström 1982].

Many bacteria present adhesins that allow them to attach to host cells. One of the mechanisms of neutralizing the action of these bacterial adhesins is by production of sIgA antibodies directed against them. In gonococcal infections, pili facilitate the attachment of bacteria to susceptible mucosal surfaces. Rabbit antisera, directed against the gonococcal adhesins, have been shown to block attachment of *N. gonorrhoeae* to epithelial cells [Tramont *et al* 1980].

b/ Toxin and enzyme neutralization

One of the direct actions of sIgA antibodies, is the ability to neutralize microbial and other environmental toxins and enzymes, by sterically blocking their binding to target cells or substrates. Gilbert *et al* (1983) observed that gonococcal and meningococcal IgA1 proteases were inhibited by exposure to secretory IgA.

Also, secretory IgA from saliva can inhibit the firm establishment of *Streptococcus mutants* on tooth surfaces by inactivating the glycosyltransferases produced by this organism, which catalyze the synthesis of extracellular polysaccharides required for the accumulation of bacteria [McGhee & Michaelek 1981; Smith *et al* 1982].

1.2.2.2. INDIRECT EFFECTOR FUNCTION OF IgA

a/ Complement

Complement is an important defense mechanism and is present at the mucosal surface [Robertson *et al* 1976]. Complement activation can occur by two different routes: the classical or the alternative pathways. It has been shown that IgA cannot activate complement by the classical pathway [Heremans 1974]. On the other hand, complexed IgA activates the alternative complement pathway in mice and guinea pig, while it does not affect human complement. It was suggested that IgA immunocomplexes do not supply a suitable surface for C3 binding, therefore the rest of the steps of the complement-mediated amplification steps do not occur [Pfaffenbach *et al* 1982]. It can be concluded that complement, although possibly important in IgA-mediated protection in some mammals, does not seem to play this role in the human.

b/ Cell mediated

IgA can interact with functionally diverse cells, including monocytes, macrophages and polymorphonuclear neutrophils through receptors for IgA Fcα domain (RFcα) found on these cells. However, the role of IgA in the various effector functions of these cells is unclear. Early studies showed that human IgA has an inhibitory effect on phagocytosis by PMNs [Van Epps *et al* 1978]. In later studies, target cells coated with IgA antibodies have enhanced phagocytosis by human oral PMNs, but not those from blood [Fanger *et al* 1983]. Finally, some investigations have shown that the phagocytic index for PMNs incubated with bacteria opsonized with sIgA is higher than for bacteria opsonized with IgA, as described in the case of *Staphylococcus aureus* [Gorter *et al* 1987].

Immunoglobulin A has been shown to be involved in the defense of the lung against pneumococcal infections. Natural antibacterial activity of purified lung lymphocytes against *Streptococcus pneumoniae* have been found to be mediated by preexisting IgA bound to the lymphocyte surfaces. This activity can be inhibited by incubation of these lymphocytes with anti-IgA antibodies,

and enhanced by incubation with purified specific IgA against phosphorylcholine, a component of the outer wall of *S. pneumoniae* [Sestini *et al* 1987; Sestini *et al* 1988].

Polymeric IgA isolated from patients with IgA myeloma inhibits normal human neutrophil and eosinophil chemotaxis [Reed *et al* 1979]. The Fcα domain of IgA has been shown to be responsible for this activity. Removal of this domain by treatment of human IgA with IgA-specific protease from *Neisseria gonorrhoeae* reduces the neutrophil chemotactic inhibitory activity of IgA [Van Epps *et al* 1980].

Fcα receptors, responsible for chemotaxis inhibition have been found on PMNs. They have been seen to stimulate PMN random migration (chemokinesis) [Sibille *et al* 1987]. In alveolar macrophages the number of Fc receptors for IgA greatly increase upon activation of these cells. This is coincident with an increase of phagocytic activity [Gauldie *et al* 1983] for which IgA is responsible. This has been shown in the case of monocyte-mediated anti-meningococcal activity, where purified IgA was able to induce phagocytosis in the absence of complement [Lowell *et al* 1980].

There is a hypothetical model which postulates a possible mechanism of regulation of the immune response by IgA. This model uses some of the findings mentioned in the last two sections. IgA can coat infecting bacteria, inhibiting killing by either complement, as IgA immunocomplexes would not supply suitable surface for C3 binding, or PMNs through chemotaxis inhibition. This would allow macrophages to phagocytose the microorganism, processing the antigen for its presentation with subsequent triggering of the immune response [Griffiss 1983].

c/ Lactoferrin, lactoperoxidase, and lysozyme

Secretory IgA1 and myeloma IgA1 and IgA2 have been shown to potentiate the effect of some of the nonspecific antibacterial factors in exocrine secretions. In contrast to IgG and IgM, IgA significantly enhances the antibacterial efficiency of the lactoperoxidase system against *Streptococcus*

mutants [Tenovuo *et al* 1982]. It is thought that IgA strengthens the action of lactoferrin and lysozyme but there is no clear evidence to confirm this idea [Kilian *et al* 1988].

1.2.3. IgA1 PROTEASE PRODUCERS

IgA1 protease is an enzyme produced by many organisms and is capable of hydrolyzing human IgA1 to produce intact Fabα and Fc fragments. The first evidence of bacterial enzymes with the ability to attack IgA1 was presented by Müller (1971) when he found a substance in *Neisseria gonorrhoeae* culture supernatants which was capable of altering IgA migration rate when incubated with serum proteins. Mehta *et al* (1973) identified intact Fcα fragment and demonstrated that an enzyme could be isolated from human faeces that cleaved myeloma IgA1 into two fragments. They found that the origin of this enzyme was bacterial because a cell free fluid from an aerobic culture of human faeces yield IgA1 protease activity. From these cultures *E. coli* was isolated. The production of the protease activity was attributed to this bacterium. However, later studies have failed to detect any IgA1 protease activity from this organism. Therefore, it is not clear whether the IgA1 protease activity detected in those cultures was from *E. coli* or from a different bacterium present in the human faeces.

IgA1 protease was first isolated from cultures filtrates of *Streptococcus sanguis*, an organism normally found in the oral cavity and intestinal tract [Plaut *et al* 1974]. This enzyme was also found in culture supernatants from *Neisseria gonorrhoeae* and *Neisseria meningitidis* [Plaut *et al* 1977]. The gonococcal enzyme has also been detected in vaginal washes from women with gonorrhoea [Blake *et al* 1979]. Both *Haemophilus influenzae* and *Streptococcus pneumoniae* produce IgA1 protease [Male 1979].

The organisms mentioned are among some of the most studied for production of IgA1 protease activity. There are many other bacteria that also produce proteins with this activity, most of which are human pathogens. There are also a large number of bacteria which have been tested for IgA1 protease activity giving negative results. Some of these IgA1 protease negative

microorganisms, which are important human pathogens, are shown in the following list [reviewed by Pratt 1988]:

Bacteroides	*fragilis*
Clostridium	*perfringens, tetani*
Corynebacterium	*diphtheriae*
Legionella	*pneumophila*
Listeria	*monocytogenes*
Mycobacterium	*tuberculosis*
Salmonella	*typhimurium*
Vibrio	*cholerae*

There is a large diversity of IgA1 proteases. Most of them have specificity for a certain peptide bond within the hinge region of IgA1 [Figure 1.2]. Also there are some organisms that produce more than one protease. *Neisseria meningitidis* produces two distinct IgA1 proteases, with different specificities, called type 1 and type 2, that are mutually exclusive among the meningococcal strains [Mulks *et al* 1980]. The same situation occurs with the proteases from *Neisseria gonorrhoeae* [Simpson *et al* 1988]. *Haemophilus influenzae* has also been shown to produce at least two serologically different IgA1 proteases with two particular specificities. Each strain makes one or both of these types of enzymes [Mulks *et al* 1982, Kilian *et al* 1983]. In contrast, there are other organisms that only produce one protease, as in the case of *Streptococcus pneumoniae* and *Streptococcus sanguis* [Kilian *et al* 1980, Gilbert *et al* 1991].

Among the immunoglobulin proteases produced by different organisms, there are some that have more than one specificity, and therefore are able to cleave other immunoglobulins apart from IgA1. Most of the organisms responsible for periodontal disease cleave only IgA1, but there are some that also cleave IgA2 and IgG [Kilian 1981]. Some studies with IgA protease from *Clostridium sp.* revealed a dual substrate specificity, in that it cleaved both IgA1 and IgA2 of the A2m(1) allotype [Fujiyama *et al* 1985]. Analysis of substrate specificity of the protease from *Serratia marcescens* showed a double specificity of this enzyme for both IgA and IgG [Molla *et al* 1988].

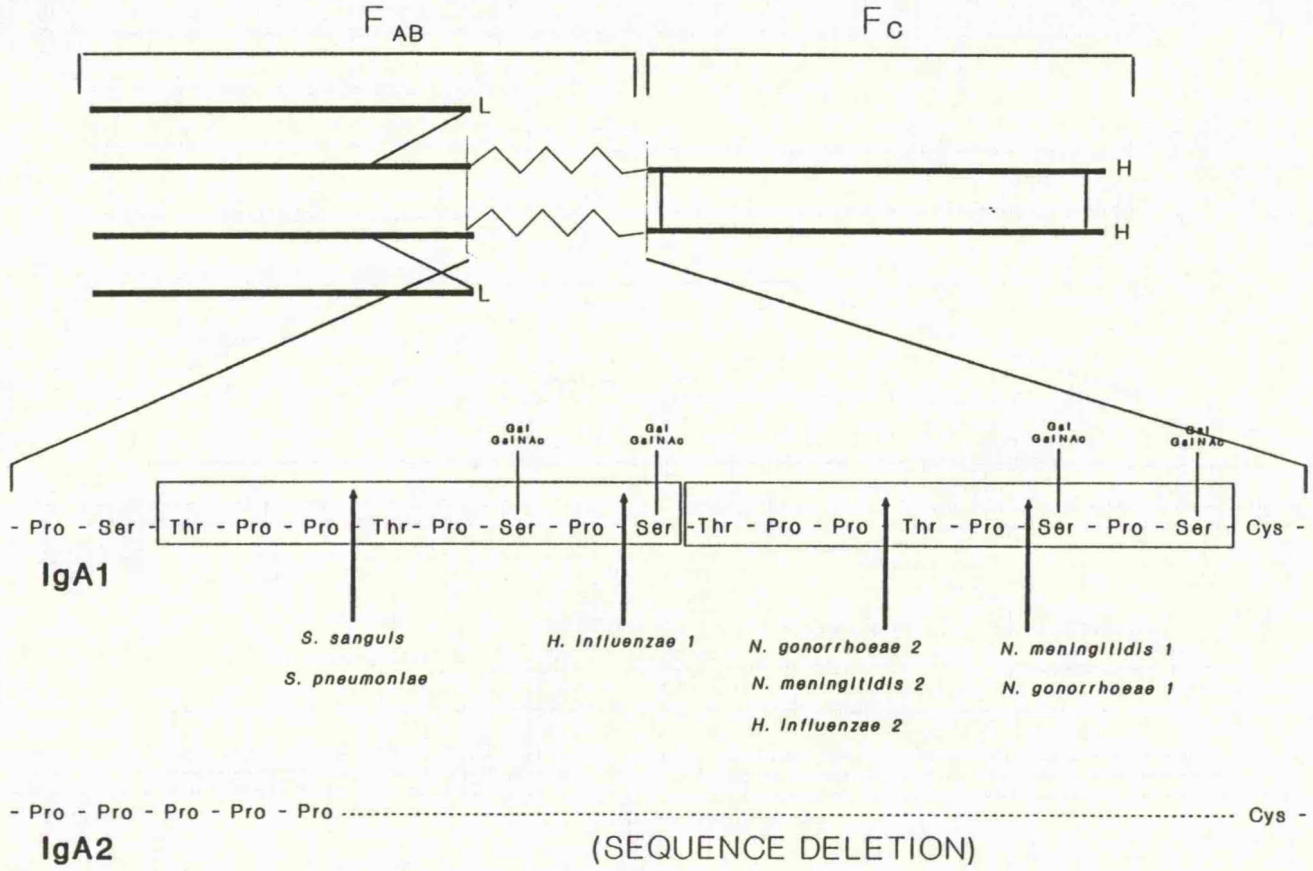


FIGURE 1.2. Representation of the hinge region sequences of both human IgA1 and IgA2 showing the locations of the specific peptide bonds cleaved by some bacterial IgA1 proteases.

Imunoglobulins are not the only substrates that have been found for some of these proteases. Shoberg and Mulks (1991) have shown that type 2 IgA1 protease from *N. gonorrhoeae* is able to cleave two outer membrane proteins and three minor cytoplasmic proteins from this organism. In addition, treatment of outer membranes from *E. coli* with gonococcal IgA1 protease caused the hydrolysis of several proteins.

1.2.4. ROLE OF IgA1 PROTEASE IN PATHOGENICITY

Although many studies have been done to determine the role of IgA1 protease, no conclusive results have been achieved. There are some studies showing a possible role for these enzymes in virulence. Deagglutination of IgA1 aggregated cells can be achieved by treatment of these cells with IgA1 proteases from *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Streptococcus sanguis* indicating that IgA1 bound to antigen is also susceptible to enzyme cleavage [Plaut *et al* 1977].

IgA1 proteases can also play a role in adherence. Mulks *et al* (1983) showed that cleavage of sIgA by gonococcal IgA1 protease reduces the ability of human secretory IgA to inhibit bacterial adherence.

It has been shown that Fabα fragments derived from treatment of sIgA with IgA1 protease from *Streptococcus sanguis* retain their specificity for antigen [Mallet *et al* 1984]. The same effect was observed when myeloma IgA1 was treated with the IgA1 protease from *Haemophilus influenzae*. Additionally, it was found that the neutralizing activity associated with the free Fabα fragment was not significantly different from that of the IgA1 protein [Mansa & Kilian 1986]. It has been postulated that Fabα fragment bound to the epitopes on the bacterial surface could protect the bacterium from the immune system by blocking access of intact antibody molecules of the same or other isotypes. Also, any mechanism of IgA-immune complex-elimination that depends on the Fcα fragment would be eliminated [Kilian & Reinholdt 1987].

One of the main problems when trying to understand the role of these IgA1

proteases in virulence is the lack of a satisfactory animal model. The specificity of IgA1 proteases for human IgA1 has made *in vivo* studies impossible. One strategy to develop an animal model would be to identify species with IgA1 susceptible to cleavage by IgA1 protease from human pathogens. Cole & Hale (1991) have found that chimpanzee sIgA can be cleaved at the hinge region by *Haemophilus influenzae* IgA1 protease into Fabα and $(Fc\alpha)_2$.SC fragments. This finding would suggests the chimpanzee as the basis of an animal model to determine the importance of IgA1 protease in pathogenesis. Several reasons such as the large number of animals needed for these experiments and legal difficulties in using primates in experimental studies, among others, have made the use of this model impractical. Therefore a search for new models of infection is being pursued.

1.2.5. CLONED IgA1 PROTEASE GENES

As discussed above, the role of IgA1 protease in pathogenesis is still uncertain. One strategy to investigate this matter is to clone the IgA1 protease genes (*iga*). This would allow the study of the presence of this gene in different bacterial isolates by the use of gene probes; and also facilitate the study of homology between enzymes with the same activity. Also, production of synthetic peptide substrates for the IgA1 protease, as described for *Neisseria gonorrhoeae* type 2 IgA1 protease [Wood & Burton 1991], would allow molecular studies of the active sites of these proteases as well as their specificities. Finally, if a suitable animal model is found, construction of IgA1 protease negative isogenic mutants would facilitate the study of the role this enzyme plays during infection.

Several *iga* genes from different organisms have been already cloned. The type 2 IgA1 protease gene from *Neisseria gonorrhoeae* F62 was the first to be cloned in *E. coli* by using a cosmid vector [Koomey *et al* 1982]. The DNA sequence of this gene has been determined [Pohlner *et al* 1987]. Another type 2 *iga* gene has been isolated from *N. gonorrhoeae* MS11 [Halter *et al* 1984]. Also, the type 1 IgA from *N. gonorrhoeae* 32819 has been cloned in *E. coli*. Greater that 99% of the enzyme produced by *E. coli* was secreted into the growth medium in soluble form [Fishman *et al* 1985].

Bricker *et al* (1984) cloned the type 1 IgA1 protease from *Haemophilus influenzae*. This enzyme was not secreted in *E. coli* and was found in the periplasmic space. Also from the same organism, the type 2 IgA1 protease has been cloned [Grundy *et al* 1987]. A considerable amount of sequence homology has been found, not only between the *iga* genes of the various serotypes of *Haemophilus influenzae* [Bricker *et al* 1985], but also with those from the gonococcus [Koomey & Falkow 1984], especially around the cleavage site determinant for a specific immunoglobulin bond [Grundy *et al* 1990].

Another *iga* gene was isolated from *Streptococcus sanguis* [Gilbert *et al* 1988]. Probes made from this gene do not hybridize with chromosomal DNA from *Streptococcus pneumoniae*, although the IgA1 proteases of the two organisms cleaved the identical peptide bond in the human IgA1 heavy chain hinge region [Gilbert *et al* 1988]. DNA sequence of the *iga* gene of *Streptococcus sanguis* has revealed no significant homology with the type 1 and type 2 IgA1 protease forms from the gonococcus and *H. influenzae* respectively [Gilbert *et al* 1991].

Finally, an attempt to clone the *iga* gene from *Streptococcus pneumoniae* has been made by using a cosmid vector. Problems found with the stability of the recombinant DNA made this cloning unsuccessful [Pratt & Boulnois 1987].

1.3. NEURAMINIDASE

Neuraminidase in an enzyme responsible for the cleavage of N-acetyl neuraminic acid (sialic acid) from mucins, glycoproteins and gangliosides. In order to get a better understanding of the possible role of this enzyme in pneumococcal disease, the structure and function of its substrate will be discussed followed by an examination of neuraminidase.

1.3.1. STRUCTURE OF SIALIC ACIDS

The sialic acids are a family of *N*- and *O*-substituted derivatives of neuraminic acid isolated from natural materials. They are commonly called neuraminic acids. These comprise all the *N*-acetylneuraminic acid (NANA) derivates and also those derived from *N*-glycolylneuraminic acid [Schauer 1982]. The structure of the parent molecule of a sialic acid is shown in Figure 1.3. [Varki & Diaz 1983] in which "R_1, R_2, R_3" represent the multiple substitutions in the sialic acid molecule at the carbons indicated with numbers.

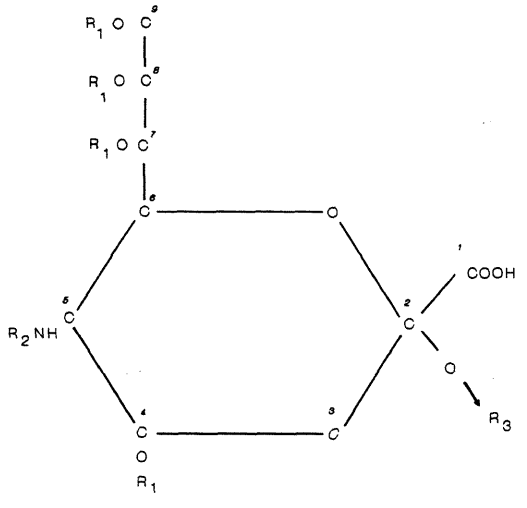


FIGURE 1.3 **Structure of sialic acids**

Representation of a molecule of sialic acid. Substitutions are represented by R_1-R_3 in the carbons indicated with numbers. The possible substitutions are as follows:

R_1 = H, Acetyl (4,7,8,9), Lactyl (9), Methyl (8), or Sulphate (8)

R_2 = Acetyl or Glycolyl

R_3 = Gal, GalNAc, GlcNAc or Sialic acid

In natural compounds, sialic acids are α-glycosidically linked, with the exception of their CMP glycosyl esters. The names, abbreviations, and nature and position of the *N*- and *O*-substituents of natural sialic acids are reviewed by Schauer 1982.

1.3.2. DISTRIBUTION OF SIALIC ACIDS

1.3.2.1. Viruses

Sialic acids have been found in several viruses such as Rous sarcoma and stomatitis virus. They have been shown to play a role in infectivity of the virions [Schloemer & Wagner 1974].

1.3.2.2. Bacteria

Sialic acids are present in capsular material of both Gram-negative and Gram-positive bacteria.

a) Gram-negative:

Escherichia coli K1	[Troy II 1992]
Neisseria meningitidis	[Troy II 1992]
Salmonella spp	[Irani & Ganapathi 1962; Kedzierska *et al* 1968]
Pasteurella haemolytica A2	[Adlam *et al* 1987]
Moraxella nonliquefaciens	[Adlam *et al* 1987]
Bordetella pertusis	[Irani & Ganapathi 1962]
Pseudomonas sp.	[Irani & Ganapathi 1962]

b) Gram-positive:

Staphylococcus aureus	[Irani & Ganapathi 1962]
Streptococcus faecalis	[Irani & Ganapathi 1962]
Corynebacterium spp.	[Luppi & Cavazzini 1966]
TypeIII Group B Streptococcus	[Baker & Kasper 1976]

1.3.2.3. Mammals

Sialic acids in mammals can be found as free molecules through to constituents of many polysaccharides, proteins and lipids.

a) Free sialic acid: Free N-acetylneuraminic acid has been found in human cerebrospinal fluid [Uzman & Rumley 1959]. However, in most mammalian tissues, the amount of free sialic acid is low. In rat brain, free N-acetyl neuraminic acid represents only about 3% of the total sialic acid content [Ng & Dain 1976].

b) Heterosaccharides: In all of these compounds, sialic acids are ketosidically bound to either D-galactose, N-acetyl-D-galactosamine or to another sialic acid residue as shown in FIGURE 1.3. They have been found in milk during early lactation and in soluble urinary oligosaccharides [Ng & Dain 1976].

c) Glycoproteins: In most animal tissues, bound sialic acids are mainly associated with proteins. This is the case for erythrocytes, histocompatibility antigens, many serum glycoproteins and a large list of enzymes [Ng & Dain 1976]. They are also present in some hormones like active follicle-stimulating hormone, human chorionic gonadotropin, and erythropoietin, where they contribute to the native active structural conformation required for interaction with the hormone receptors [PapKoff 1966]. Finally, sialic acids have been found in all classes of immunoglobulins in man, rabbit, horse, guinea pig and sheep [Clamp & Johnson 1972].

d) Gangliosides: Two thirds of the brain sialic acids are associated with gangliosides. They are constituents of nerve endings, and presumably are essential for the normal functioning of the neuronal membranes [Ng & Dain 1976]. They also serve as receptors for some toxins like tetanus, diphtheria, botulinus and cholera [Schauer 1982].

1.3.3. ROLE OF SIALIC ACIDS

Sialic acids have been the main focus of much research since they are involved in the regulation of a great variety of biological phenomena. They usually have a protective role in living cells and organisms. This appears to be due to their peripheral position in the glycoconjugates and, correspondingly, to their frequent external location in the cell membrane.
This section will study the main roles of sialic acids in different processes.

1.3.3.1. Cell to cell interaction

Given the presence of negatively charged sialic acids residues on the cell surface, it may be expected that these compounds strongly influence the behavior of cells. There are several lines of evidence that support this theory.

Sialic acids can play a role in adhesion. It has been shown that removal of sialic acids from cultured cells, by neuraminidase treatment, provokes detachment of these cells from glass [Weiss 1963].

Cell aggregation is intimately involved in cell differentiation and cancer since, in both cases, cell disaggregation releases cells that migrate and adhere again to each other or to different cells. Studies carried out by Deman and Bruyneel (1973) have stressed the role of sialic acids in these processes. They found an increase in aggregation of HeLa cells after removal of the sialic acids by neuraminidase treatment. Similar effects have been shown in the generation of hybridomas between B cells and myeloma cells. After neuraminidase treatment, not only mouse-mouse cell fusion, but also human-mouse hybridization produced significantly more clones. [Igarashi 1990].

Sialic acids have been shown to play diametrically opposed roles depending on the cell. In some cells like blood platelets, erythrocytes, and carcinoma cells in culture, sialic acids prevent aggregation due to electrostatic repulsion, whereas, in others, for example, chick embryonic muscle-cells aggregation is facilitated [Schauer 1982].

Various basic polymers are known to agglutinate erythrocytes. It has been shown in several investigations that removal of sialic acids from the surface of those cells by neuraminidase treatment, causes a decrease in hemagglutination mediated by those polymers [Jeanloz & Codington 1976]. This fact again emphasizes the importance of the sialic acids in the determination of surface negative charges.

1.3.3.2. Contribution to macromolecular structures

Sialic acids contribute markedly to the high degree of viscosity of many mucous secretions through their presence in many glycoproteins. This effect is considered to be due to the mutually repelling sialic acid residues extending oligosaccharide chains from the protein core, thus giving these molecules a rod-like structure, and facilitating gel-formation in water [Schauer 1982].

The influence of sialic acids the macromolecular conformation seems to be one the reasons for the proteolytic resistance of several glycoproteins. One example of this is IgA1. It has been shown that removal of sialic acids from

this immunolgobulin by the action of neuraminidase, increases the susceptibility to protease cleavage [Reinholdt *et al* 1990]. Interestingly, the pneumococcus produces both IgA1 protease and neuraminidase. These two enzymes may act in combination during pneumococcal disease. Neuraminidase could potentiate the action of IgA1 protease by releasing sialic acids from the hinge region of IgA1.

A role has been postulated for sialyl residues in the conformation of some enzymes. Removal of sialic acids from such enzymes can cause changes in their affinity towards some substrates, or even optimal pH for activity, as has been shown in the case of acetylcholinesterase [Brodbeck *et al* 1973].

Sialic acids are also responsible for the rigidity of cell surfaces. It has been shown that cells from Sarcoma 37 and Ehrlich murine ascites were very easily deformed after neuraminidase treatment [Weiss 1965].

1.3.3.3. Anti-Recognition effect

The anti-recognition effect is one of the most studied roles of sialic acids. This effect was initially observed by Morell *et al* (1971) and Ashwell & Morell (1974). They discovered that fully sialylated proteins survive normally in the circulation, whereas even partial desialylation, results in prompt destruction of the modified protein by the liver. The terminal galactose residues, exposed after neuraminidase treatment, were identified as the major determinants for hepatic recognition and clearance. The same result has been observed in the removal of hormones from the circulation [Rosenberg & Schengrund 1976]. This effect has also been observed in blood cells. When injected back into the rat, neuraminidase-treated rat lymphocytes were rapidly taken up by the liver [Woodruff & Gesner 1968]. Also partially desialylated rat erythrocytes are phagocytosed *in vitro* by macrophages. This effect is assumed to be mediated by immunoglobulins and complement [Schauer 1982]. Stenberg *et al* (1991) showed that injection of neuraminidase from *Clostridium perfringens* into mice, induced thrombocytopenia, since platelets lacking sialic acid on their membranes are recognized as foreign and therefore removed from the circulation.

In the case of tumores, Simmons *et al* (1971) found that a total regression of firmly established methylcholanthrene fibrosarcoma can be induced in mice by inoculation of living tumor cells treated with *Vibrio cholerae* neuraminidase. This effect was due to an increased immunogenicity of the treated cells.

1.3.3.4. Sialic acids as receptors

Sialic acids act as receptors. They interact specifically with biologically active compounds and micro-organisms, playing a role as receptors for toxins, such as cholera, diptheria and botulinus. In these cases, gangliosides are able to act as toxin receptors in the cell membranes [Clowes *et al* 1972].

Whether or not sialic acids are involved in the interaction of several hormones with their receptors is not clear. In the particular case of ACTH (adrenocorticotropic hormone), it has been suggested that sialic acid may play a role in transmitting the signal from the ACTH-receptor interaction to the catalytic unit of adenyl cyclase [Haskar *et al* 1974].

Sialic acids are essential constituents of some immunoglobulin receptors. This has been shown for human lymphocytes in culture. After sialidase treatment, of these cells, their IgM receptors were completely inactivated. This is in contrast to IgG receptors, the expression of which was increased by treatment with this enzyme [Itoh & Kumagai 1980].

Although for some bacteria it has been shown that removal of sialic acid from receptors is required for attachment [Andersson *et al* 1983; Guzman *et al* 1990], for others the presence of this molecule is essential. One example of the latter group is *Streptococcus sanguis* which possesses a ligand that interacts with sialic acids from glycoprotein receptors. Neuraminidase treatment of salivary glycoproteins abolishes aggregating activity against this organism, suggesting that sialic acids might be an essential component in this reaction [Demuth *et al* 1990].

Studies of viral infection have revealed that for some viruses, the presence of

sialic acids on the surface of the host cells is essential for infection. In the case of influenza virus, an hemagglutinin has been found to be responsible for attachment to cells. This hemagglutinin, exposed in the surface of the virion, recognizes and binds to cell surface sialic acids [Glick *et al* 1991].

1.3.3.5. Summary

Sialic acids are responsible for some cell to cell interactions regulating adhesion, agglutination and aggregation. They contribute to viscosity of macromolecular structures like mucus and cell surfaces, conferring, also, proteolytic resistance to glycoproteins and regulating enzyme conformation. Sialic acids play a role in the antirecognition effect of some cells by the immune system. They are receptors for some toxins, bacteria and viruses. Sialic acids also regulate interaction of hormones and immunoglobulins with their receptors.

1.3.4. DEGRADATION OF SIALIC ACIDS BY NEURAMINIDASE

Sialidase (neuraminidase, *N*-acetylneuraminosyl glycohydrolase, E.C. 3.2.1.18) catalyses the hydrolysis of sialic-acid-containing glycoproteins, glycolipids, and oligo- and polysaccharides [Cabezas 1991].

Sialidases can only cleave α-glycosidic linkaged glycosides. In general, $\alpha(2\rightarrow3)$ bonds are rapidly hydrolyzed, whereas $\alpha(2\rightarrow6)$ and $\alpha(2\rightarrow8)$ linkages are hydrolyzed much more slowly by most sialidases, with the exception of the neuraminidase from *Arthrobacter ureafaciens* [Schauer 1982].

Practically all neuraminidases have an acidic pH for hydrolysis (pH 3.5-7.0) and an acidic pI [Rosenber & Schengrund 1976].

1.3.5. NEURAMINIDASE PRODUCERS

The discovery of neuraminidases was made by using culture filtrates of *Clostridium perfringens* and *Vibrio cholerae* [McCrea 1947; Burnet & Stone 1947]. These workers observed that these filtrates could cause erythrocytes to be non agglutinable by influenza virus. They called the agent in *Vibrio* filtrates the "receptor-destroying enzyme" because it destroyed viral receptors

on erythrocytes, displaying an enzyme-like characteristics. Since then, many neuraminidase enzymes have been characterized.

Neuraminidases have a wider distribution in nature than have the sialic acids. They have been isolated from mammals through bacteria to viruses. Neuraminidases from mammals are the least characterized because of the difficulty of their purification. They have been isolated from different tissues, such us brain, liver kidney, mammary glands, muscle, erythrocytes and placenta [see for review Rosenberg and Schengrund 1976].

The viral neuraminidases have been studied in most detail. Neuraminidases have been found in orthomyxoviruses, paramyxoviruses and metamyxoviruses [Schauer 1982]. Molecular complexes having sialidase activity have been isolated and characterized, with sizes ranging from 100 to 200 KD. In the case of the monomeric enzymes 50 to 60 KD covers the range generally reported [Drzeniek 1972]. Of the viral neuraminidases, the best characterized is the one from influenza virus [Air & Laver 1989].

Sialidases have been found also in many bacteria, several strains of fungus and protozoa. The most important ones are shown in Table 1.1 . There is a large range of neuraminidases produced by these organisms. Some of them produce just one type of this enzyme, while there are others like *Arthrobacter ureafaciens* [Otha *et al* 1989], *Streptococcus viridans* [von Nicolai *et al* 1980] that produce more than one neuraminidase. Also, in terms of their structure, most neuraminidases are monomeric, although there are some exceptions: *Clostridium chauvoei* neuraminidase is composed of two subunits [Heuermann *et al* 1991].

NEURAMINIDASE PRODUCERS

ORGANISM	REFERENCE
Acanthamoeba castellanii	Pellegrin *et al* 1991
Acanthamoeba polyphaga	Pellegrin *et al* 1991
Actynomyces israelii	Moncla & Braham 1989
Actynomyces meyeri	Moncla & Braham 1989
Actynomyces naeslundii	Moncla & Braham 1989
Actynomyces odontolyticus	Moncla & Braham 1989
Actynomyces vicous	Moncla & Braham 1989
Arthrobacter sialophilus	Flashner *et al* 1977
Arthrobacter ureafaciens	Ohta *et al* 1989
Bacteroides fragilis	Guzman *et al* 1990
Bacteroides loeschii	Takeshita *et al* 1991
Clostridium chauvoei	Heuermann *et al* 1991
Clostridium perfringens	McCrea 1947
Clostridium septicum	Rothe *et al* 1991
Clostridium sordellii	Rothe *et al* 1989
Corynebacterium aquaticum	Sondag-Thull 1989
Corynebacterium diphteriae	Moriyama & Barksdale 1967
Klebsiella aerogenes	Pardoe 1970
Naegleria fowleri	Eisen & Franson 1987
Pasteurella multicocida	Drzeniek *et al* 1972
Propionibacterium acnes	Schauer 1982
Pseudomonas aeruginosa	Leprate & Michel-Briand 1980
Salmonella typhimurium	Hoyer *et al* 1991
Streptococci Group B	Milligan *et al* 1980
Streptococcus mitis	Reinholdt *et al* 1990
Streptococcus pneumoniae	Hughes & Jeanloz 1964
Streptococcus sanguis	Varki & Diaz 1983
Streptococcus viridans	von Nicolai *et al* 1980
Streptomyces albus	Myhill & Cook 1972
Trichomonas foetus	Romanouska & Watkins 1963
Trypanosoma cruzi	Pereira 1986
Vibrio cholerae	Burnet & Stone 1947

TABLE 1.1

In many of these organisms the production of neuraminidase can be induced by the presence of the products from the hydrolysis of sialic acids. Neuraminidase expression from *Streptococcus pneumoniae* was stimulated by N-acetylneuraminic acid (NANA) and *N*-acetylmannosamine [Kelly *et al* 1966]. The same type of induction has been also observed in the synthesis of neuraminidase by *Pasteurella multocida* [Drzeniek *et al* 1972] and *Bacteroides fragilis* [Russo *et al* 1990]. Inducible expression may be explained by feed-back regulation of neuraminidase production.

1.3.6. PURIFICATION OF BACTERIAL NEURAMINIDASES

Neuraminidases from many different bacteria have been purified. There is no agreement in terms of molecular weight and number of enzymes produced per organism. Table 1.2. shows a list of some bacterial neuraminidases that have been purified. Some of them yielded a single product after purification, while others gave more than one protein with neuraminidase activity. Also, for some neuraminidases, contradictory results have been reported from different purification procedures. For *Vibrio cholerae*, analytical ultracentrifugation gave a protein of 90 KD [Ada *et al* 1961], whereas gel filtration gave a size of 68 KD [Drzeniek 1972].

In *Arthrobacter ureafaciens*, a gel filtration method gave two isoenzymes of 39 KD and 51 KD [Uchida *et al* 1979]. Ten years later, the same group, using a similar purification technique, obtained four isoenzymes from the same organism, with molecular weights different from those reported previously [Ohta *et al* 1989][Table 1.2].

A wide range of molecular weights have been reported for neuraminidase from *Streptococcus pneumoniae*. Some reports indicate several neuraminidase enzymes with molecular weights from 35 KD [Tanenbaum *et al* 1970] to 70 KD [Tanembaum & Sun 1971].

To explain the diversity of molecular weights found when purifying pneumococcal neuraminidase several non-exclusive explanations could be given. The different techniques used to determine the molecular weight of

PURIFIED BACTERIAL NEURAMINIDASES

ORGANISM	MOLECULAR WEIGHT	REFERENCE
Actynomyces viscous	113000	Henningsen *et al* 1991
Arthrobacter sialophilus	87000	Flashner *et al* 1977
Arthrobacter ureafaciens	39000 & 51000	Uchida *et al* 1979
	88000,(66000)x2 & 52000	Ohta *et al* 1989
Bacteroides loescheii	87000	Tekeshita *et al* 1991
Clostridium chauvoei	300000 = 2 x 150	Heuermann *et al* 1991
Clostridium perfringens	56000	Cassidy *et al* 1965
Clostridium septicum	110000	Rothe *et al* 1991
Corynebacterium aquaticum	55600	Sondag-Thull 1989
Corynebacterium diphteriae	65000	Moriyama & Barksdale 1967
Propionibacterium acnes	33000	Schauer 1982
Streptococci Group B	106000	Milligan *et al* 1988
Streptococcus pneumoniae	>100000	Hughes & Jeanloz 1964
	35000	Tanenbaum *et al* 1970
	70600	Tanembaum *et al* 1971
	69800	Stahl & O'Toole 1972
	107000	Lock *et al* 1988b
	65000	Scalon *et al* 1989
Streptococcus viridans	45000 & 86000	von Nicolai *et al* 1988
Vibrio cholerae	90000	Ada *et al* 1961
	68000	Drzeniek 1972b

TABLE 1.2

this enzyme could have given dissimilar results for the same protein. Also, the variable size may have been caused by proteolytic degradation of the enzyme [Lock *et al* 1988]. Finally, the size difference could correspond to the presence of several neuraminidase isoenzymes in the pneumococcus.

Several studies have revealed no significant kinetic differences between the pneumococcal neuraminidase from several purifications [Hughes & Jeanloz 1964; Stahl & O'Toole 1972]. Interestingly pI determinations have given different values for these enzymes [Stahl & O'Toole 1972; Scalon *et al* 1989]. Nevertheless, as has been already mentioned, it is not clear that the purified enzymes are really different. Therefore, no conclusions should be taken from these kinetic studies.

1.3.7. ROLE OF NEURAMINIDASE IN BACTERIA AND VIRUSES

Since the discovery of the neuraminidases, many studies have been focused towards a better understanding of their role, if any, in virulence. Neuraminidases from bacteria, viruses and eukaryotes may share a common role since it has been postulated that they have a monophylogenic origin [Hoyer *et al* 1992]. Many hypothesis have been formulated trying to attribute to neuraminidase a role in pathogenicity although there are still no conclusive explanations.

In the particular case of viral sialidases, many roles have been proposed: viral penetration into the host cell, provision of low-molecular-weight metabolites for viral propagation, release of newly formed virus from the host cell, destruction of substances which protect the cell surface from virus binding and viral binding to the host cell surface by enzyme-substrate interaction. None of them have been widely adopted because of conflicting evidence for each hypothesis [Rosenberg & Schengrund 1976].

In the case of pathogenic bacteria, it is still difficult to establish a correlation between the production of neuraminidase and pathogenicity. It is clear that these enzymes are capable of altering the characteristics of important sialic-acid-containing macromolecules on mammalian cell surfaces and in

intercellular matrices. They can also attack circulatory macroglobulins, enzymes, and hormones.

For some bacteria it has been shown that neuraminidase could be involved in attachment. *Bacteroides fragilis* possesses an adhesin that mediates a neuraminidase-dependent attachment of the organism to mammalian cells. This ligand interacts with a mammalian cell receptor that contains a galactoside residue, exposed after neuraminidase treatment [Guzman *et al* 1990].

The pneumococcus might have a similar mechanism of attachment, as suggested by the fact that it specifically attaches to glycoconjugates that could act as receptors for this organism, on the surface of pharyngeal cells, after exposure by neuraminidase treatment [Andersson *et al* 1983]. In contrast, recent experimental data suggests that the pneumococcus does not interact with epithelial cells of the respiratory mucosa *in vitro*. Interaction has only been observed with the epithelial-derived secretions [Feldman *et al* 1992].

Neuraminidase can enhance toxin binding to cells, as has been demonstrated in the case of *Vibrio cholerae* by Galen *et al* (1992). By incubating mouse fibroblasts with a *V. cholerae* filtrate containing neuraminidase, they demonstrated that cholera toxin binding to these cells increases five to eight fold. Therefore, neuraminidase plays a subtle but significant role in the binding and uptake of cholera toxin by susceptible cells. Since under certain environmental conditions the production of the toxin is reduced, they postulated that neuraminidase could enhance toxin binding contributing to the pathogenic process.

The case of cholera toxin is not the only one in which neuraminidase acts to potentiate a second agent. Studies performed by Reinholdt *et al* (1990) showed that degradation of the carbohydrate moiety of IgA1 by neuraminidase accompanied IgA1 protease activity in *Streptococcus oralis* and protease producing strains of *Streptococcus mitis* biovar 1. Previously, Male (1979) noticed that pneumococcal IgA1 protease appeared to cleave

substrates more rapidly than *Haemophilus infuenza* protease. This was due to the presence of neuraminidase activity in the crude enzyme preparation from the pneumococcus. Therefore, neuraminidase is able to remove sialic acids from the cleavage region of IgA1 protease in the immunoglobulin, facilitating access of the protease to this region.

1.3.8. ROLE OF PNEUMOCOCCAL NEURAMINIDASE

There is still very little known about the role in disease of pneumococcal neuraminidase. Although it has not yet been shown as a virulence factor, there are some data suggesting that this enzyme might be involved in virulence. Firstly all fresh clinical isolates produce the enzyme [Kelly *et al* 1967]. Also, abnormally high levels of free NANA (1.42 mg/100ml), when compared to control patients (0.37 mg/100ml), have been detected in CSF from patients with acute pneumococcal meningitis. This has been correlated with the presentation of coma and adverse outcome. It was suggested that neuraminidase may be a factor involved in the pathogenesis of pneumococcal meningitis [O'Toole *et al* 1971]. These results would explain the findings of Kelly & Greiff's (1970) who observed that inoculating mice intracerebrally with neuraminidase, a variety of neurological signs such as paralysis, convulsions, hyperactivity and tremor of extremities were displayed prior to death. When inoculation of neuraminidase was via the intraperitoneal route, mice became inactive and had ruffled fur. In contrast, O'Toole & Stahl (1975) found that intrathecal inoculation of dogs with crude neuraminidase did not show any clinical morbidity. Therefore, a role for neuraminidase in meningitis is still uncertain.

Immunization experiments [Lock *et al* 1988] have given more support to the idea of a role in virulence for the pneumococcal neuraminidase. They observed a small but significant increase in survival time ($p < 0.05$) in groups of mice that had been immunized with purified pneumococcal neuraminidase, pretreated in 3.4% v/v formaldehyde, and challenged with virulent pneumococci. This treatment caused a reduction in enzyme activity of 60 %. No protection was observed when formaldehyde treatment was omitted. It was thought that when an enzyme such as neuraminidase is used as an

immunogen while retaining all or some of its activity, its protective effect may be partially countered or overcome by its toxicity.

The existence of contradictory and unclear results, when analyzing the data on pneumococcal neuraminidase in the experiments mentioned above, makes more critical the need for a rigorous genetic analysis of this enzyme in pathogenicity. To achieve this, the molecular form of Koch's postulates [Falkow 1988] should be applied. They can be summarized as follows:

1. "The virulence property should always be associated with pathogenic strains". As has been mentioned above, neuraminidase activity has been found in all the clinical isolates tested.

2. "Specific inactivation of the gene that specifies the virulence factor should cause a measurable loss in virulence". This could be achieved by cloning the neuraminidase gene from the pneumococcus, followed by construction of an isogenic neuraminidase negative mutant of this bacterium. This mutant could be used in animal experiments to compare its virulence with the wild type. Also, the cloning of the neuraminidase gene from the pneumococcus would allow studies of the structure and function of this enzyme which could be used for the understanding of the mechanism of action of neuraminidase during pneumococcal infection.

3. "Reversion of the mutated gene or allelic replacement by recombination or by complementation should restore pathogenicity". There would be two ways to achieve this. Firstly, the mutated neuraminidase gene could be replaced with the wild type gene. Secondly, the mutated gene could be complemented with the neuraminidase expressed from a plasmid carrying this activity. Re-establishment of virulence in these organisms could be analyzed. As complementation of neuraminidase activity from a plasmid could result in overexpression of this enzyme, giving disparate results compared to the wild type, replacement of the mutated gene seems more appropriate.

In conclusion, to apply the molecular form of the Koch's postulates for the pneumococcal neuraminidase, the cloning of the neuraminidase gene would be required as a starting point. This cloning plus further characterization of the neuraminidase gene is one of the main subjects of this thesis.

1.3.9. CLONING OF BACTERIAL NEURAMINIDASE GENES

During the last few years interest in neuraminidases from bacterial pathogens has increased enormously with respect to their possible role in virulence. In order to carry out further studies on these neuraminidases some of them have already been cloned and sequenced. Vimr *et al* (1988) isolated a neuraminidase gene (*nanH*) from a cosmid library of *Vibrio cholerae* 395. The *nanH* gene was subcloned and located within a 4.8 Kb *Bgl*II restriction endonuclease fragment. This neuraminidase gene has been sequenced by Galen *et al* (1990).

Several neuraminidases from *Clostridium spp.* have been cloned. The first is from *C. perfringens* that was cloned in *E. coli* using a pUC18 vector. The neuraminidase gene was encoded by a 1.4 Kb DNA fragment. DNA sequence gave a predicted amino acid sequence corresponding to a 42.7 KD protein [Roggentin *et al* 1988]. A neuraminidase gene was cloned from *C. sordellii* G12 by using as a DNA probe a synthetic oligonucleotide predicted from the N-terminal sequence of the purified protein. A 44.7 KD size protein was estimated from the DNA sequence of this neuraminidase gene [Rothe *et al* 1989]. Finally, a neuramindase gene has been cloned from *C. septicum* using degenerate oligonucleotides designed from conserved amino acid sequences of other bacterial sialidases. The complete nucleotide sequence of this sialidase gene was determined giving a predicted size protein of 110.000 KD [Rothe *et al* 1991].

Russo *et al* (1990) mobilized a plasmid-based library of *Bacteroides fragilis* DNA into a *B. fragilis nanH*⁻ mutant to isolate a neuraminidase gene from this organism. The neuraminidase structural gene was localized within a 1.5 to 2.5 Kb fragment.

Yeung & Fernández (1991), screened a genomic library of *Actinomyces viscous* for neuraminidase activity and cloned a gene coding for this activity. The isolated recombinant contained 3.4 Kb insert that, in maxicells, directed the synthesis of a protein with an apparent molecular mass of 100 KD. Hybridization studies performed with this clone showed that *A. viscous* contains a single copy of the neuraminidase gene. Simultaneously, Henningsen *et al* (1991) cloned a neuraminidase from the same organism by screening pUC18 and pUC19 genomic libraries for the enzyme activity. DNA sequence analysis of the isolated gene gave a predicted protein of 113 KD. It is not known whether the two neuraminidase genes and their products are the same.

Hoyer *et al* (1992) cloned a sialidase gene from *Salmonella typhimurium* LT2 and its nucleotide sequence gave a predicted protein of 41.300 Daltons.

Berry *et al* (1988) isolated a clone expressing neuraminidase activity from a pneumococcal genomic library. This clone suffered a spontaneous deletion during its isolation. The resulting clone after the deletion, called pJCP301, had a 3 Kb insert of contiguous pneumococcal DNA, as shown by Southern Blot analysis. This recombinant expressed low levels of neuraminidase activity. It was suspected that the insert coded for a truncated neuraminidase which retained activity. An attempt to isolate a stable clone of the whole neuraminidase gene was made. Chromosomal DNA from the pneumococcus was digested with *Hind*III, and 3.3 Kb *Hind*III fragment, containing the neuraminidase coding sequence, was cloned into pJDC9. This vector was chosen since it possesses transcription terminators that stabilize high promoter activity of cloned pneumococcal DNA in *E. coli*. The recombinant, called pJCP302, did not express detectable neuraminidase activity. It was speculated that this DNA fragment also suffered a deletion similar to that which occurred in pJCP301 when into *E. coli*. The lack of neuraminidase activity from this recombinant may be due to the absence of the neuraminidase gene from its insert after the deletion occurred.

The cloning of bacterial neuraminidase genes mentioned has not given any

direct proof for the presence of neuraminidase isoenzymes in those organisms studied in this thesis. Data presented indicates the presence of at least two distinct neuraminidase genes in the pneumococcus.

1.3.10. RELATIONSHIP BETWEEN DIFFERENT NEURAMINIDASES

Although several neuraminidase genes have been cloned and sequenced, very little known about their products. One of the starting points in the analysis of these enzymes would be to establish sequence comparisons between them. This could reveal some useful information about their possible structure-function and evolutionary relationships.

Roggentin *et al* (1989) were the first to perform these comparisons. By aligning sialidase sequences from *Clostridium sordellii* G12, *Clostridium perfringens* A99, *Salmonella typhymurium* LT-2, *Vibrio cholerae* 395, they detected the following conserved sequence motif: Ser-X-Asp-X-Gly-X-Thr-Trp. This motif, commonly called the "aspartic motif", was repeated four times in each bacterial neuraminidase [Figure 1.4]. The distance between the motifs was also highly conserved. Homology between these bacterial sequences and some influenza A neuraminidases has been found.

After this analysis, the neuraminidase gene from *Clostridium septicum* has been sequenced and the product analyzed for the presence of the above motifs [Figure 1.4] [Rothe *et al* 1991].

Other neuraminidase enzymes with the aspartic motifs have been described in *Bacteroides fragilis* [Russo *et al* 1990] and *Actinomyces viscous* [Henningsen *et al* 1991], but in these cases the aspartic motif is present five times [Figure 1.4].

The presence of the aspartic motif is not restricted to bacteria and viruses. It has also been found in the neuraminidase from the protozoan *Trypanosoma cruzi*. This neuraminidase possesses four of these motifs although, the number of residues between them is not as conserved as in the bacterial neuraminidases [Pereira *et al* 1991].

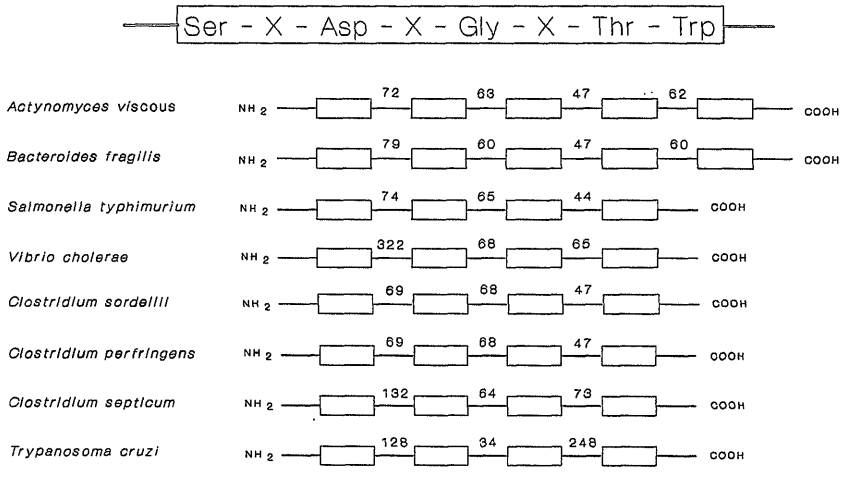


FIGURE 1.4 Representation of the "Aspartic motif" by boxes. Numbers indicate amino acid between motifs.

Although much effort has been expended trying to understand the function of these motifs, their role still remains intriguing. Studies on sialidases carried out by Hoyer *et al* (1992), using not only the aspartic motif, but, in some cases, whole sequences, have revealed the presence of amino acid sequence homologies between bacterial, viral and eukaryotic sialidases, suggesting the possibility of a common evolutionary origin for these enzymes.

1.3.11. LOCATION OF NEURAMINIDASE WITHIN DIFFERENT ORGANISMS

In bacteria, different subcellular locations for neuraminidase have been described. Some bacteria express a cell associated enzyme. This has been shown for *Corynebacterium diphteriae* which possesses an intracellular neuraminidase, which may be located in the cytoplasm [Moriyama & Barksdale 1967]. A similar location has been deduced for the enzyme from

Salmonella typhimurium. The lack of a signal peptide in the predicted protein sequence of the *S. typhimurium* enzyme is consistent with this conclusion [Hoyer *et al* 1992]. The neuraminidase from *Bacteroides fragilis* is thought to be surface exposed [Guzman *et al* 1990].

Other neuraminidases, being mainly cell associated, are also released into the growth media. For example in *Actynomyces viscous*, more than 80 % of the neuraminidase is cell associated and around 10% is secreted [Costello *et al* 1979]. A similar situation exists in *Acanthamoeba castellanii* [Pellegrin *et al* 1991], *Klebsiella aerogenes* [Pardoe 1970] and *Pasteurella multocida* [Drzeniek *et al* 1972].

In *Streptococcus pneumoniae* it has been difficult to localize neuraminidase. Lee & Howe (1966) showed that pneumococcal neuraminidase is secreted during the logarithmic phase and only a small fraction of the total neuraminidase is cell associated at any given time. However, it remains unclear whether it is released into the media as a consequence of bacterial autolysis, or is actively secreted.

Many bacterial neuraminidases are actively secreted. This is the case for *Pseudomonas aeruginosa* [Leprate & Michel-Briand 1980], *Clostridium perfringens* [McCrea 1947], *Vibrio cholerae* [Vimr *et al* 1988], *Clostridium sordellii* [Rothe *et al* 1989], and *Clostridium septicum* [Rothe *et al* 1991]. In the latter three cases, protein sequence information has shown the presence of a signal peptide.

Among the eukaryotes, the neuraminidase from *Trypanosoma cruzi* has also been described in detail. This neuraminidase is only expressed in trypomastigotes. Electron microscopy studies revealed the presence of this enzyme on the cell surface [Prioli *et al* 1991].

Finally, some viral neuraminidases have been well characterized. The one from influenza virus has been shown to be present on the surface of the

virions, being constituted of monomers organized into a tetrameric structure [Air & Laver 1989].

The variability of neuraminidase location, together with its presence in different organisms with very diverse pathogenesis, adds another complication on the understanding of the putative role this enzyme could play in pathogenesis.

1.3.12. PROJECT AIMS

Although there are some indications for a putative role in virulence for IgA1 protease and neuraminidase in many organisms, in the case of the pneumococcus very little is known about these enzymes.

IgA1 protease activity has been detected in practically all pneumococcal clinical isolates. Cloning of the IgA1 protease gene from the pneumococcus would allow a detailed study of the structure and function of this enzyme. This information would be required, in the long term, once a suitable animal model has been found, to understand the mode of action of this enzyme *in vivo*.

Pneumococcal neuraminidase has been studied for many years. There are several indications suggesting a possible role in virulence for this enzyme. The cloning of the neuraminidase gene from the pneumococcus would allow the analysis of the structure and function of this enzyme. This information will be used to understand the mechanism of action of this enzyme during pneumococcal infection in animal models.

The requirements for the study of IgA1 protease and neuraminidase in the pneumococcus set the scene for the work described in this thesis which had the following aims:

a) Construction of a pneumococcal genomic library using a Lambda vector. Lambda was chosen since pneumococcal DNA in other vectors has proved unstable [see chapter 3].
b) Development of a screening method for IgA1 protease suitable for library

screening. Use of neuraminidase screening methods already described.

c) Isolation of the genes encoding IgA1 protease and neuraminidase from the pneumococcal genomic library.
d) The DNA sequencing of IgA1 protease and neuraminidase genes.
e) Analysis of the IgA1 protease and neuraminidase DNA sequences to allow design of further studies on these enzymes.

All these studies will set the basis for the understanding of the possible role of these enzymes in the pneumococcal disease.

PART II

MATERIALS AND METHODS

2.1. Bacterial Strains and Plasmids

The bacterial strains and plasmids used in this study are listed TABLE 2.1. and TABLE 2.2. The *E. coli* strains LE392 and Q359 were grown on Luria agar medium whereas the strain JM101 was normally grown B-Agar and stored in M9 minimal agar media. All these strains were grown in Luria broth. The growth temperature was 37°C. *Streptococcus pneumoniae* strains were propagated on Brain heart infusion broth (BHI, Oxoid), blood agar base or columbia agar (Oxoid) plus 5% defibrinated horse blood. All bacterial strains were stored on glass beads in the appropriate media with 10% glycerol as the cryoprotectant. Strains were kept at -70°C or -20°C.

2.1.1. Chemicals

All chemical were obtained from BDH Limited, Fisons or Sigma Biochemicals unless specified

2.1.2. Growth Media

All media were prepared with distilled water and sterilized at 15 psi (pounds per square inch) for twenty minutes.

Luria broth (1L)

10g Peptone
5g Yeast extract
5g NaCl

For Luria agar (L-agar) 1 litre of L-broth was solidified with Bacto Agar (Difco) to a final concentration of 2%.

L-Soft top agar (200ml)

1.30g Bactoagar (Difco)
2g Peptone (BBL)
0.5g NaCl

L-Soft top agar was used to plate Lambda phage as indicated in Section 2.5.1.1.

M9 minimal agar (400ml)

100ml 4x M9
5.7g Bacto agar (Difco)
4ml 20% Glucose
400µl 20% $MgSO_4$
200µl Thiamine
H_2O up to 400ml

4x M9 (1L)

24g Na_2HPO_4
12g KH_2PO_4
4g NH_4Cl
2g NaCl

B-agar (400ml)

4g Bactopeptone (Oxoid, Merck)
3.2g NaCl
6g Bacto agar

When using *E. coli* JM101 carrying pJDC9, 1mg/ml of erythromycin from a 50mg/ml stock (in 95% ethanol), 2µl/ml of X-Gal(2% in dimethylformamide) and 1µl/ml of IPTG (100mM in H_2O) were added when agar still molten (held at 42°C). When using *E.coli* JM101 carrying pJCP301, 100µg/ml of Ampicillin from a 25µg/ml stock (in H_2O) were added when agar was still molten.

B-Soft top agar (200ml)

2g Bactopeptone (Oxoid, Merck)
1.6g NaCl
1.2g Bactoagar

B-Soft top agar was used to plate M13mp18/mp19 bacteriophages as indicated in Section 2.3.2.3.

Columbia agar & Blood agar base

These were made according to the manufacturers instructions, sterilized and allowed to cool to 45°C. Horse blood (Difco) was added to a final concentration of 5%.

Brain Heart Infusion

BHI was made according to the manufacturers instructions.

2.2. Transformation of Bacterial Cells

2.2.1. Production of competent cells

L-Broth (20 mls) was inoculated with 100μl of an overnight culture and bacteria grown to mid-log phase (OD_{600} 0.5). The cells were washed firstly in 20ml in ice-cold 100mM $CaCl_2$ (Sigma) then in 10ml and finally resuspended in 1ml of the same solution. Cells were left on ice for at least one hour before being used for transformation.

2.2.2. Transformation with plasmid DNA

Competent cell (200μl) and up to 50μl of the DNA (in water or TE) to be transformed were mixed and placed on ice for 1 hour. The cells were heat-shocked at 42°C for 2 min. L-broth (500 μl) was added to the cells which were then incubated for a further one hour at 37°C and plated onto L-agar plates (1 to 200μl per plate) which contained the appropriate antibiotics, substrates or inducers. The plates were incubated overnight at 37°C.

2.2.3. Transformation with M13 Bacteriophage DNA

Competent *E. coli* JM101 cells (prepared as described in Section 2.2.1) were incubated on ice for 1 hour with the bacteriophage DNA and heat-shocked at 42°C for 2 min. The transformed cells together with 200μl mid log phase JM101 were added to 3ml molten B-agar (held at 42°C) containing 4μl of 200mg/ml IPTG, and 40μl of 20mg/ml X-Gal in dimethylformamide. The suspension was mixed and poured immediately onto a B-agar plate, rocked to

disperse and once set incubated at 37°C overnight.

2.3. Procedures for DNA Extraction

2.3.1. Extraction of Chromosomal DNA

Different protocols for chromosomal DNA extraction were used according to the DNA purity required. For procedures requiring high purity DNA such us cloning or Southern blot DNA hybridization, protocols I and III were used, whereas when low purity of DNA was required, like dot blot DNA hybridization, protocol II was chosen.

PROTOCOL I

S. pneumoniae chromosomal DNA was prepared by a method based on that of Saito & Miura (1963). Bacteria were streaked from a glycerol stock onto a blood agar plate, and an optochin disk was placed on the streak. The plate was incubated under low oxygen conditions for up to 24 hours. If the colony morphology and sensitivity to optochin confirmed the growth as *S. pneumoniae*, a single colony was picked into 10ml of BHI broth (starter culture) and incubated for 8 hours at 37°C. The starter culture was then used to inoculate 500ml of BHI broth which was again incubated at 37°C. All *S. pneumoniae* isolates were grown without shaking. The 8 hour starter culture and the overnight cultures were streaked onto blood agar plates with optochin discs to confirm purity of the pneumococcal cultures, prior to harvesting the cells. Cells (200ml) were harvested by centrifugation in a Sorvall GSA rotor at 7000 rpm for 15 min at 4°C and washed in TES (25 mM NaCl, 2.5 mM EDTA, 15 mM Tris-HCl pH 8). The cells were finally resuspended in 4ml TE-Sucrose (50 mM Tris-HCl pH 8, 1mM EDTA, 25% sucrose). This was followed by the addition of 0.8ml 500 mM EDTA and 0.5ml 10% SDS and 50μl Proteinase K (10mg/ml). After 15 min on ice a further incubation at 65°C was carried out overnight. This was followed by several phenol:chloroform extractions until no white material was seen at the interface of the two phases after centrifugation at 5000 rpm in a Heraeus-Christ centrifuge for 20 min at 4°C. The top, aqueous layer was removed

into a corex tube. To precipitate the DNA, 2 volumes of ethanol and 0.1 volumes of 3M Sodium acetate pH 5.3 were added followed by storage at -70°C for 2 hours. DNA was collected by centrifugation at 4500 g for 30 min at 20°C. DNA was finally resuspended in 500 μl of TE.

PROTOCOL II

S. pneumoniae was grown in 20 mls of BHI broth as the starter culture in Protocol I. The pneumococcal culture was also checked for purity as before. Cells were harvested by centrifugation at 5000 rpm in a Heraeus-Christ centrifuge for 10 min at 5°C. The pellet was resuspended in 0.1 ml lysis buffer (0.1% sodium deoxycolate , 0.01% sodium dodecyl sulphate, 0.15M NaCl), and incubated at 37°C for 10 min. This was followed by the addition of 0.9 ml SSC (0.15M NaCl, 0.015 M sodium citrate) and incubation at 65°C for 15 min. The crude chromosomal DNA preparation obtained was suitable for dot blot hybridization.

PROTOCOL III

S. pneumoniae was grown in 200 ml of BHI broth as described in protocol I using half of the starter culture. The purity of the culture was also confirmed as described. Cells were harvested by centrifugation in a Sorvall GSA rotor at 7000 rpm for 15 min at 4°C followed by two washes in 10% sucrose. The cell pellet was finally resuspended in 10 ml of lysozyme solution (2mg/ml lysozyme, 10.3% sucrose, 25mM Tris-HCl pH 7.4, 25mM EDTA) and incubated at 37°C for 5 min. Then, 2.4 ml of 0.5M EDTA were added with a further incubation at 37°C for 5 min. This was followed by the addition of 1.4ml of 10 % sodium dodecyl sulphate and proteinase K to 100μg/ml. This preparation was incubated for 1 h at 50°C. This was followed by phenol:chloroform extraction and thoroughly mixed by gentle inversion of the tube. The two phases were separated by centrifugation (Heraeus-Christ, 5000 rpm 20 min at 4°C). The top, aqueous layer was removed into a fresh tube and subjected to further cycles of phenol-chloroform extractions until there was no white material seen at the interface of the two phases after centrifugation. At all times the aqueous phase was treated gently to minimize the mechanical shearing of DNA. Finally, an equal volume of chloroform

was added, mixed for ten minutes and centrifuged in Heraeus-Christ centrifuge at 3500 rpm for 5 min at 4°C. To the aqueous phase 2.5 volumes of ethanol were added, and stirred with a glass rod, DNA became attached to the rod. The rod was carefully washed in 70% ethanol and then 100% ethanol. DNA was redissolved by immersing the rod in TE and incubating overnight at 5 °C.

2.3.2. Extraction of Plasmid DNA

Plasmid DNA extraction were done as described by Maniatis (1982) using the following solutions:

Solution I
50mM glucose
25mM Tris-HCl
10mM EDTA

Solution II
0.2M NaOH
1% SDS

Solution III
5M acetate (11.5 ml glacial acetic acid)
3M potassium ions (60ml 5M potassium acetate)
H_2O to 100 ml

2.3.2.1. Small Scale Extraction of Plasmid DNA

Small scale preparation of plasmid DNA used 1.5ml stationary phase bacterial cultures. Cells collected by centrifugation were resuspended in 100μl of Solution I and incubated for 30 min on ice. Solution II (200μl) was added and the tube was carefully inverted and placed on ice for a further 5 minutes. Solution III (150μl) was added, mixed and placed on ice for between 10 and 60 min. The preparation was centrifuged in a bench top

microfuge for 5 min and the supernatant retained. Protein was removed by phenol extraction (Section 2.3.3.) and the plasmid DNA subsequently precipitated at room temperature by adding 2 volumes of ethanol. DNA was collected by centrifugation in a microfuge for 5 min and resuspended in 50μl TE containing 20 μg DNAse free RNAse.

2.3.2.2. Large Scale Extraction of Plasmid DNA

Stationary phase bacterial cultures (400ml) were used for larger scale preparation of plasmid DNA. The cells were collected by centrifugation, resuspended in 10 ml of solution I and left to stand on ice for 10 min. Freshly made solution II (20ml) was mixed in and left on ice for another 10 min. After the addition of 15ml of ice-cold solution III, the preparation was placed on ice for a further 10 min. Cell debris was removed from the plasmid preparation by centrifugation at 4°C for 20 min at 35000g. Isopropyl alcohol (0.6 volumes) was added to the supernatant, mixed and left to stand at room temperature for a minimum of 15 mins. DNA was collected by centrifugation at 4000g for 30min at 20°C. The DNA pellet was dried under vacuum and resuspended in sterile distilled water to a final volume of 17ml. Caesium chloride was added to a final concentration of 1mg/ml and ethidium bromide to 50μg/ml. Chromosomal and plasmid DNA were separated at 40000 rpm using a Sorvall TV850 vertical rotor in a Sorval OTD 60 centrifuge for 20 hours at 20°C. DNA was visualized under UV light, and the lower band of plasmid DNA extracted. Ethidium bromide was removed by equilibration with caesium chloride-saturated isopropanol. Caesium chloride was removed by exhaustive dialysis against TE at 5°C. Plasmid DNA was stored dissolved in TE at -20°C.

2.3.2.3. Extraction of M13mp18/19 Template DNA

The recombinant bacteriophage M13mp18/19 was transformed into JM101 (Section 2.2.3) and white plaques were picked into 5 ml of L-broth containing 100μl of an overnight culture of JM101 and incubated at 37°C for 5 hours with vigorous aeration. An aliquot of 1.5 ml was decanted into a 1.5ml eppendorf and cells were collected by centrifugation in a microfuge for

5 min. The cell pellet was used to extract replicative form DNA to check the identity of the insert as described in Section 2.3.2.1. The supernatant was used to extract single stranded DNA as follows: 1.2 mls of the supernantant were mixed with 300μl of feshly prepared 20%PEG6000/2.5M NaCl solution and vortexed briefly. The mixture was incubated for 15 min at room temperature and centrifuged in a microfuge for 10 minutes to pellet the virus. The supernatant was discarded, the sample was centrifuged for a further 2 min and all traces of supernatant removed. The viral pellet was resuspended in 100 μl of TES buffer (20mM Tris-HCl pH 7.7, 10mM NaCl and 0.1 mM EDTA) and extracted with 50μl of phenol saturated with TE buffer by vortexing vigorously for 20 sec. followed by centrifugation in the microfuge for 3 min to separate the phases. The organic phase with phenol was removed with a pipettor. A further phenol extraction was done as before. The aqueous phase was extracted with 50 μl of chloroform by vortexing for 20 sec. followed by spinning in a microfuge for 3 min to separate the phases. From the aqueous phase, 80μl were transferred to a 1.5ml eppendorf by using a pipettor. DNA was precipitated by the addition of 9μl of 3M Sodium acetate pH 5.0, and 200μl of ethanol followed by storage at -20°C overnight. DNA was then collected by centrifugation in a microfuge for 10 min and resuspended in 10μl of TE. Template DNA in 1μl was visualized by agarose gel electrophoresis (Section 2.4.4) and 1μg was used in a sequencing reaction (Section 2.7.).

2.3.3. Phenol Extraction and Ethanol Precipitation of DNA

One volume of phenol:chloroform (1:1 w/v) containing hydroxyquinoline and equilibrated in Tris-HCl pH 7.5 was added to the DNA sample and mixed carefully (for chloroform extraction one volume of chloroform was added to the DNA sample). The aqueous phase was separated either at 20°C for 20 min at 18000 g, or in a bench top microfuge for 5 min, and then retained. For ethanol precipitation sodium acetate to a final concentration of 300mM, and 2 volumes of ethanol (stored at -20°C) were added. The sample was placed at -20°C or -70°C for a minimum of 30 min and the DNA collected by centrifugation in a bench top microfuge for 5 min.

2.4. Techniques Used in Routine DNA Manipulations

2.4.1. DNA Restriction digests

Restriction endonuclease cleavage of DNA was performed according to the manufactures recomendations typically in 10 μl reactions with one unit of enzyme at 37°C. Restriction endonucleases were purchased from Pharmacia Biochemicals Inc or Life Technologies Ltd (GIBCO/BRL).

2.4.2. DNA Dephosphorylation

DNA dephosphorylation was based in the method described by Maniatis *et al* (1982). After DNA digestion, a phenol/chloroform extraction and DNA ethanol precipitation was done. DNA was treated with 0.01 units of calf intestinal alkaline phosphatase (CIP) from Boehringer, using the reaction buffer supplied by the manufacturers, in a 50 μl final volume. To dephosphorylate protruding 5' termini DNA was incubated at 37°C for 30 min, then another aliquot of CIP was added with a further 30 min incubation at 37°C. To dephosphorylate blunt ends or recessed 5' termini an incubation for 15 min at 37°C and 15 min at 56°C was required. Then a second aliquot of CIP was added with another two incubations at the same temperatures. The reaction was stopped in both cases by the addition of 40μl H_2O, 5μl 10%SDS and 10 μl of a solution containing 100mM Tris-HCl pH8.0, 1M NaCl and 10mM EDTA. DNA was twice treated with phenol/chloroform followed by ethanol precipitation after which it was ready for ligation.

2.4.3. DNA Ligation

T4 ligase, purchased from Pharmacia Biochemicals Inc, was used in ligase buffer at 14°C overnight. Ratios vector to insert DNA were calculated empirically.

LIGASE BUFFER 10 X

1M $MgCl_2$	2μl
1M DDT	2μl
100mM ATP*	2μl
1M Tris-HCl pH 7.5	10μl
H_2O	4μl

* Freshly made or from frozen stock

2.4.4. Agarose Gel Electrophoresis

DNA fragments were separated by agarose gel electrophoresis using 0.7% Seakem agarose for large DNA fragments (> 3Kb) and 1% for small ones (< 3Kb) in TAE buffer pH 7.7 (40mM Tris-acetate, 1mM EDTA) with 0.5μg/ml ethidium bromide and visualized using a longwave UV transilluminator. Electrophoresis was performed in TAE containing 0.5μg/ml ethidium bromide at 100V for 1 to 4 hours according to the size of the gel, or at 10V-20V overnight. DNA samples were mixed with an equal volume of DNA loading buffer (50% glycerol, 5X TAE, 0.25% bromophenol blue, 0.25% xylene cyanol) prior to loading. The DNA size markers used were bacteriophage lambda DNA cut with *Hind*III or *Bgl*II , or 1Kb ladder (BRL/GIBCO). Lambda markers were incubated at 65°C in DNA loading buffer for 10 min before use.

2.4.5. DNA Extraction From Agarose Gels

For subcloning, DNA fragments were purified from agarose gels by two different techniques:

2.4.5.1. Low Melting Point Agarose DNA Extraction

This method was based in the one described by Langridge *et al* (1980)

SOLUTIONS

CETAB is Hexadecyltrimethylammonium bromide (Sigma)

Butanol-CETAB and H_2O-CETAB were prepared by dissolving 1g of CETAB (Sigma) in 100ml of Butanol saturated in water. When dissolved

100ml of water saturated in butanol were added. After mixing, the suspension was left overnight at room temperature. The top phase was butanol-CETAB and the bottom phase was H_2O-CETAB.

PROCEDURE

The DNA fragment in question was firstly separated by agarose gel electrophoresis using 1% low melting point agarose (BRL) in the following buffer: 40mM Tris-HCl pH 7.9, 5mM NaAc, 1mM EDTA, 0.5μg/ml EtBr. The electrophoresis was also run in the same buffer. Agarose containing the fragment was excised from the gel and melted at 65°C-70°C for 30 min. The following steps were performed at 37°C. DNA was extracted by adding 1 vol of each H_2O-CETAB and Butanol-CETAB to the melted agarose. The mixture was vortexed for 1 min and centrifuged in a microfuge for 1 min. The top phase (butanol) was transferred to a fresh tube. The rest of the operations were done at room temperature. Butanol was extracted by adding 1/4 vol of 0.3M NaAc pH 7.0 followed by vortexing and 1min centrifugation in a microfuge. The lower phase was collected and mixed with 1 vol of chloroform. After centrifugation for 1 min in a microfuge, the top phase, containing the DNA, was transferred to a fresh tube. DNA was precipitated by the addition of 1/25 vol of 3M NaAc pH 5.3 and 3 vol of EtOH.

2.4.5.2. Electroelution of DNA

The DNA fragment in question was first separated by standard agarose gel electrophoresis using TAE buffer and Seakem agarose. Agarose containing the DNA was excised from the gel and the DNA was recovered by electrolution as follows: The gel slice was placed in dialysis tubing containing 500μl of 1M Tris-HCl pH 7.5 and subject to 120V for 1 hour in TAE buffer, polarity was reversed for the last 30 sec. DNA was collected by ethanol precipitation (Section 2.3.3) from the solution surrounding the slice.

2.5. Construction of a LambdaEMBL301 Genomic Library

2.5.1. Preparation of LambdaEMBL301 Arms

2.5.1.1. LambdaEMBL301 plating

E. coli (LE392 or Q359) were used as plating cells. They were grown in L-broth at 37°C to mid logarithmic phase. The prepared cells (200μl) were mixed with phage suspension in a 3ml capped test tube. This was followed by the addition of 20% maltose (25μl) and 1M $MgCl_2$ (25μl). The phage was allowed to absorb to the cells at 37°C for 20 min. As a control an identical mixture without phage was prepared and treated in the same way. To these preparations 2.5ml of L-Soft Top agar (Section 2.1.2) was added and mixed by inverting the tube twice. The content of each tube was immediately poured onto L-agar plates. Once the agar was set plates were incubated at 37°C overnight.

2.5.1.2. LambdaEMBL301 Large Scale Preparation

The method used for the purification of large scale LambdaEMBL301 was based on the one described by Hadfield (1987). Vector DNA stock should be initiated from a single pure phage plaque, rather than by subcloning from an existing phage stock. For this purpose phage was plated from a stock as described in section 2.5.1.1. From the resulting plates containing phage, a single plaque was inoculated in 10ml of L-broth containing 10mM $MgCl_2$. This was followed by incubation at 37°C with vigorous shaking for 7 hours.

For the large scale preparation, 200μl from the supernatant of the above culture were used to inoculate 500ml L-broth containing 10mM $MgCl_2$ in a 2 litre flask. This culture was grown at 37°C overnight. The resultant culture was dense due to surviving Lambda-resistant cells and the debris of lysed sensitive cells. To kill the surviving cells 5ml of chloroform were added to the culture followed by shaking for 15 min at 37°C. The culture was allowed

to stand on the bench for 10 min, to allow chloroform to settle. 400mls were decanted into centrifuge pots, leaving the chloroform behind, and centrifuged at 8000 rpm in a Sorvall GSA rotor for 10 min at 4 °C. Supernatants were carefully collected, free of cell debris. The number of phage was detected by plating 10-fold dilutions of these supernatants (Section 2.5.1.1.). A total titer $>10^{12}$ phage was taken as sufficient for the large scale phage preparation.

Pancreatic DNAse (Sigma) and RNAse (Sigma) were added, each to 10μg/ml (40μl of 10mg/ml stocks; made up in water, and stored at -20°C). This was followed by incubation at 37°C for at least 1 hour. To precipitate the phage, solid NaCl to a final concentration of 1M was added, followed by PEG 6000 to a final concentration of 10% w/v. The phage preparation was cooled in ice-water for at least 1 h or overnight at 4°C. Phage was pelleted by spinning at 9000 rpm for 10 min in a sorvall GSA rotor. The pellet was resuspended in 4 ml of Lambda Tris buffer (10mM Tris pH7.4, 10mM $MgSO_4$, 0.01% Gelatin) with a minimun of shearing force. The phage suspension was transferred to a 30 ml plastic universal tube. This suspension contained both phage and cell debris. Most of the debris was removed by pelleting as follows: (a) the suspension was centrifuged at 5000 rpm for 5 min and the supernatant was collected; (b) the pellet was washed in 1ml of Lambda Tris buffer and the new supernatant was pooled with the above; (c) step (b) was repeated. Phage suspension was mixed with solid CsCl (0.71 times the weight of the phage suspension) and mixed until dissolved. The solution was transferred to a 35 ml ultracentrifuge tube and the tube filled with a CsCl solution of the same density (1.45 - 1.50g/ml). Phage was separated by centrifugation at 40,000 rpm using a Sorvall TV850 OTD 60 centrifuge for 20 hours at 4°C. The phage, present in a bluish-white band at the middle of the tube was removed in less than 1ml solution, using a 2ml syringe and a 21 gauge needle introduced through the side of the tube. Caesium chloride was removed by exhaustive dialysis against lambda Tris buffer at room temperature. Viral protein coats were removed by several phenol:chloroform extractions until no proteins appeared in the interphase of phage suspension and phenol:chloroform. This was followed by one chloroform extraction after which Lambda Tris buffer was removed by exhaustive dialysis against TE at

5°C. Phage DNA was precipitated with ethanol at -20°C (Section 2.3.3), collected by centrifugation in a bench top microfuge for 10 minutes at room temperature and washed in 70% ethanol. After vacuum drying the DNA was resuspended in 300μl of TE final volume and stored at -20°C.

2.5.1.3. Separation of Lambda Arms from Central Replaceable Fragment

LambdaEMBL301 (5μg) were digested with 10 units of each *BamH*I and *EcoR*I with the appropriate enzyme buffers in a final volume of 50 μl. Once digestion was completed, TE was added up to 300μl followed by phenol:chloroform extraction. DNA was precipitated with ethanol (Section 2.3.3) in a dry-ice/ethanol bath for 5 min followed by a 15 min centrifugation in a microfuge. The DNA pellet was washed in 70% ethanol and resuspended in 400μl of TE. These DNA precipitation steps were repeated twice more. The final pellet was dried under vacuum and resuspended in 5μl of TE.

To confirm the impossibility of LambdaEMBL301 to religate to central replaceable fragment, 2μg of LambdaEMBL301 arms were ligated in a 5μl ligation mixture. This was followed by an *in vitro* packaging of this mixture (Section 2.5.3) followed by plating (Section 2.5.1.1.). The lack of plaques in the plating indicated that no religation had occurred.

2.5.2. Partial Digestion of Chromosomal DNA

Chromosomal DNA was prepared as Section 2.3.1, PROTOCOL I. Chromosomal DNA was partially digested with *Sau3A*I as described by Maniatis *et al* (1982), by adding 2-fold decreasing units of enzyme to 1μg aliquots of chromosomal DNA followed by incubation at 37°C for 1 h. Enzyme activity was stopped by incubation at 65°C for 10 min. From each digestion 100ng were analyzed in a 0.7% agarose gel electrophoresis (Section 2.4.4.). The digests which gave the production of fragments between 15 and 20 Kb, were selected and the remaining 900 ng used for the final cloning to the arms without any further treatment.

2.5.3. Final Ligation and *In Vitro* Packaging

For the final ligation, 2μg of LambdaEMBL301 arms were ligated to 400ng of partially digested chromosomal DNA in a final volume of 20μl. An aliquot of the ligation mixture (1μl) was analyzed by agarose gel electrophoresis, and 5μl of the ligated DNA was packaged into bacteriophage Lambda heads by the method of Hohn (1979) using the Lambda DNA *in vitro* packaging kit produced by Amersham International PLC as described in the manufacture's instructions.

2.5.4. LambdaEMBL301 genomic library amplification

The packaged library was plated onto non-dried L-agar plates using *E. coli* strain LE392 at a high density of recombinants per plate (10^3-10^4). After overnight growth at 37°C, 2 ml of Lambda tris buffer was added onto the soft top containing the plaques which were rocked at 5°C overnight. Lambda tris buffer was then collected from the plates and mixed with the soft top agar that was scraped off. A 1/100 volume of chloroform was added and mix thoroughly to kill the cells. This mixture was also vortexed to release as many phage as possible from the agar. The lambda tris buffer containing phage was collected from the agar by centrifugation at 5,000 rpm 5°C for 20 min in a Heraeus-Christ centrifuge. The collected buffer containing the amplified genomic library was stored at 5°C.

2.6. DNA Hybridization

2.6.1. Preparation of Nylon Filters for DNA Hybridization

2.6.1.1. Transfer of DNA to Nylon Filters by Southern Blotting

DNA was transferred to filters as described by Southern (1975). DNA samples were separated by agarose gel electrophoresis as described in Section 2.4.4. and the gel photographed along side a linear ruler. The DNA was de-purinated by soaking the gel in 0.25M HCl for 10 min. The gel was then rinsed briefly in distilled water and placed in denaturing solution (0.5M

NaOH, 1.5M NaCl) for 15 min with occassional shaking. The gel was again rinsed in distilled water and placed in neutralizing solution (0.5M Tris-HCl pH 7.5, 3M NaCl) for 30 min with occasional shaking as before. The gel was rinsed again and placed on six sheets of pre-wetted (20xSSC) whatman paper (3mm) without trapping air bubbles. A pre-wetted (10xSSC) sheet of nylon membrane (hybond-N, Amersham International PLC) was placed on the gel with a pre-wetted sheet (10xSSC) Whatman paper on top, again taking care to avoid air bubbles. Four sheets of dry Whatman paper were placed above this with a stack of paper towels. Finally a glass plate and a 500g weight were placed on the top. The lower sheets of Whatman paper were regularly soaked with 20xSSC. The apparatus was left for at least 3 hours or overnight for the gel to transfer and the dismantled. The nylon filter was aired dried, wrapped in Saran wrap and exposed to UV light from a long wave transilluminator for 5 min to fix the DNA to filter. Filters were stored at room temperature in the dark until required for DNA hybridization (Section 2.6.2.).

20 X SSC
3M NaCl
0.3M Trisodium citrate

2.6.1.2. Transfer of DNA to Nylon Filters by Dot Blotting

Before applying to nylon filters DNA was mixed with an identical volume of 2M ammonium acetate followed by heating at 100°C in a boiling water bath for 5 min and then incubated on ice for 5-10 min. DNA samples were then applied onto nylon filters soaked in 1M ammonium acetate by either using a Bio-Dot microfiltration apparatus (Bio-Rad) as manufactures instructions, or in the case of non-radioactive hybridization (Section 2.6.2.2.) by hand on untreated filters. Filters with blotted DNA were wrapped in Saran wrap and exposed to UV radiation, as in the case of Southern blotting, and also stored at room temperature in the dark.

2.6.1.3. Preparation of Filters for Plaque Hybridization

For plaque hybridization Colony/Plaque Hybridization Transfer Membranes from Dupond Biotechnology System (NEF-978/978A) were used as manufacturer's instructions. Essentially, plates were stored at 4°C follow by plaque lifting. Membranes containing phage were treated once with 0.5M NaOH for 2min and twice with 1M Tris-HCl pH7.5. After drying at room temperature, membranes were prehybridized in (1%SDS, 1M NaCl) for 6 hours at 65°C, followed by addition of both denatured 100μg/ml salmon sperm DNA and radioactive probe (Section 2.6.2.1.1.). Incubation of the membranes for 6-24 hours at 65°C was followed by washes as described in Section 2.6.2.1.2.

2.6.2. DNA Hybridization Procedures

2.6.2.1. Use of Radioactive Probes

2.6.2.1.1. Production of a radiolabelled DNA probe.

Plasmid DNA was cleaved with the appropriate restriction endonucleases and the fragments separated by agarose gel electrophoresis on a 1% low melting point agarose (BRL). The required DNA fragment was excised from the gel and added to sterile water (1.5ml water per gram agarose). The sample was placed in a boiling water bath for 7 min and then stored at -20°C. Prior to use the sample was boiled for an additional 5 min and then incubated on ice for 10-60 min.

DNA was radiolabelled using random hexanucleotide primers as described by Feinberg and Volgelstein (1983). The DNA (10-25ng) to be labelled was denatured by placing in boiling water for 5 min then quickly cooled to 37°C prior to use. The labelling reaction was initiated by the addition of reagents in the following order:

xμl H_2O (to a total volume of 25μl)
5μl Oligo-labelling buffer
1μl ^{32}P α-dCTP (specific activity 100mCi/ml)
0.5μl Klenow (large fragment) DNA polymerase I (1000 units/16.7μl)

The mixture was incubated for at least 4 hours or overnight. Reaction was stopped by the addition of 75μl of H_2O. Labelled probed was stored at -20°C until used.

Solution O 1.25M Tris-HCl pH 8.0
0.125M Magnesium chloride

dNTP mix 100mM dATP in TE buffer
100mM dGTP in TE buffer
100mM dTTP in TE buffer

Solution A 1ml Solution O
18μl 2-ß-mercaptoethanol (14.2M)
5μl of dATP, dGTP, dTTP

Solution B 2M Hepes (adjusted to pH 6.6 with 4M NaOH)

Solution C Hexadeoxynucleotides resuspended in TE buffer to give a final concentration of 90_A units/ml.

Oligo-labelling buffer
10μl Solution A
25μl Solution B
15μl Solution C

Nucleotides and hexanucleotides were obtained from Pharmacia and [α-^{32}P]dCTP from Amersham International PLC.

2.6.2.1.2. Hybridization of the DNA immobilized on filters with the probe.

Souther blot, dot blot or plaque hybridization filters were incubated at 65°C with shaking in 100 ml of hybridization solution (see below) for 2 hours. This solution was the discarded and replaced by 20 ml hybridization solution (see below) containing the radiolabelled probe DNA which has been boiled for 5 min before adding. The filter was incubated overnight at 65°C with shaking. Hybridization solution is 3xSSC (Section 2.6.1.1.), 2x Denhards (50x is 1% each Ficoll, BSA, polyvinolpyrollindine), 200μg/ml salmon sperm DNA, 0.1% SDS, 6%PEG6000. Prehybridization solution was the same except with 5x Denhards. Both solutions were stored at -20°C without the salmon sperm DNA. Salmon sperm DNA was sheared by forcing through a narrow gauge syringe needle and denatured by boiling prior to use.

After the hybridization period the filters were washed three times by shaking in 250ml 2xSSC, 0.1%SDS at 65°C for 15 min and three times in the final wash solution of choice (2xSSC for approximately 70% homology sequences, 0.5xSSC for approximately 85 % homology and 0.1%xSSC for approximately 95% DNA homology) under the same conditions. SDS concentration (0.1%) and temperature (65°C) were not varied. The filters were air dried.

The filters were wrapped in Saran wrap for autoradiography and placed in a cassette carrying intensifying screens. Kodak X-Omat AR film was exposed to the filters at -70°C. Films were developed in Agfa-Geveart automatic film processor.

2.6.2.2. Use of Non-Radioactive Probes

The Enhanced Chemiluminescent (ECL) DNA hybridization Kit (Amersham) was used following manufacturer's instructions. In this method horseradish-peroxidase (HRP) was covalently linked to the denatured probe using glutaraldehyde (one molecule of HRP is attached approximately every 30 bases of DNA). DNA probes labelled in this way were hybridized to target

sequences following manufacturer's instructions. The blots were then soaked in a solution containing luminol which being excited by oxidation with HRP, in the presence of H_2O_2, emits light when decaying. This emission was detected using X-ray film.

2.7. DNA Sequencing

DNA was sequenced by the chain termination method described by Sanger *et al* (1977) in which DNA synthesis from deoxynucleotide triphosphates is terminated by the addition of dideoxynucleotides triphosphates. The M13 cloning vectors, M13mp18 and M13mp19 were used to generate single stranded DNA templates. For double stranded DNA sequencing pJDC9 was used.

Sequencing reactions were performed using the Sequenase Version 2.0 kit supplied by United States Biochemical Corporation. The protocol recommended by the manufactures was followed in the case of single stranded DNA. In the case of double stranded DNA, 3μg of plasmid DNA in 18μl of TE were mixed with 2μl of freshly prepared 2M NaOH and allowed to stand for 5 min. Denatured DNA was precipitated overnight at -20°C by the addition of 50 μl of ethanol and 2μl of 3M sodium acetate pH 5.3. DNA was collected by centrifugation for 10 min in a microfuge and resuspended in 7μl of TE for immediate sequencing reaction as described for single stranded DNA.

Both single and double stranded DNA fragments were sequenced by using either a primer annealing 40 bp before the vector's multiple cloning site (-40 universal primer) or oligonucleotide primers synthesized for this purpose. DNA was radiolabelled by incorporating [α-^{32}S]dATP in the extension reactions. The radiolabelled DNA fragments were separated by gradient gel electrophoresis (Biggin *et al* 1983). Preparation of the gel used the following solutions:

Gel solution 1

7ml	2.5x TBE acrylamide/urea mix
90μl	10% ammonium persulphate
5μl	TEMED

Gel solution 2

40ml	0.5x TBE acrylamide/urea mix
360μl	10% ammonium persulphate
15μl	TEMED

0.5x TBE acrylamide/urea mix (1L)

430g	urea
50ml	10x TBE
150ml	40% acrylamide

2.5x TBE acrylamide/urea mix (1L)

430g	urea
250ml	10x TBE
150ml	40% acrylamide
50g	sucrose
50mg	bromophenol blue

40% acrylamide (1L)

380g	acrylamide
20g	bisacrylamide
	deionized with 50g amberlite per litre

Electrophoresis grade ammonium persulphate was purchased from BIO-RAD, TEMED from Sigma Chemical company Ltd, acrylamide and bisacrylamide (NN'-methylenbisacrylamide) from BDH. 10x TBE is 0.089 Tris-borate, 0.089M boric acid, 0.002M EDTA.

To prepare the gel, clean gel plates (20cm x 50cm) were taped together separated by 0.4mm spacers. 10ml gel solution 2 followed by 7ml of gel

solution 1 were drawn up into a 25ml pipette. Four air bubbles were introduced to form a rough gradient. The liquid was introduced between the gel plates and the cavity filled with the remainder of gel solution 2. The comb was positioned and the plates clamped along each side. Gels were routinely left to set for one hour or overnight.

A vertical gel electrophoresis system was used. Running buffer in the top tank was 0.5x TBE and in the lower 1x TBE according with the gradient of the gel. The gel was pre-run at a constant power of 40W and the wells rinsed with running buffer prior to loading. Electrophoresis was performed at 40W for 2 hours 30 min to visualize the smallest DNA fragments and for increasingly longer periods (3 to 12 hours) to separate larger fragments.

After electrophoresis the gel plates were prised apart and the gel was soaked in fixing solution (10% methanol, 10% acetic acid) for 10 mins and then rinsed briefly in distilled water. It was transferred to pre-wetted filter paper, covered with Saran wrap and dried under vacuum at 80°C. Autoradiography used Dupond Cronex film and took place at room temperature.

Nucleotide sequences were analyzed using the University of Wisconsin Genetic Computer Group sequence analysis software [Devereux *et al* 1984] on the Leicester University VAX/VMS cluster. The program FASTA, which uses the algorithm of Pearson and Lipman (1988) was used for sequence comparisons.

2.8. Polyacrylamide Gel Electrophoresis

Proteins were separated and visualized in denaturing gels using the method of Laemmli (1970) for discontinuous polyacrylamide gel electrophoresis (PAGE). A Mini protein II cell modular mini electrophoresis system from Bio-Rad was used following manufacturer's instructions. The solutions used for the separating and the staking gel are described below. After addition of TEMED to the separating solution, it was poured into the gel plate spaces and overlaid with n-butanol. After polymerization the n-butanol was removed

and the gel surface washed with water and blotted dry. The gel solution was poured onto the separating gel, and a comb inserted at the top of the assembly. Protein samples to be electrophoresed were prepared by the addition of SDS/PAGE loading buffer to a ratio 5:1 sample to buffer, and placed in a boiling water bath for 5 min prior to loading onto the gel. Samples (20μl) were loaded and run through the gel at 200 V for approximately 1 hour, until the blue dye font reached the bottom of the gel. The separated proteins were visualized by staining with Coomasie blue stain for 15 min following by destaining, or the polyacrylamide gels were western blotted (Section 2.9.) or treated for autoradiography (Section 2.10.)

The solutions required for polyacrylamide gel electrophoresis are as follows:

30 % ACRYLAMIDE SOLUTION (Protogel, New Jersey)

SEPARATING GEL (10%)

1.25 ml 30 % acrylamide solution
5 ml 0.75M Tris-HCl pH8.8, 0.2% SDS
0.5 ml 15mg/ml ammonium persulphate
3.25 ml H_2O
16 μl TEMED

STACKING GEL (5%)

1.25 ml 30% acrylamide solution
5 ml 0.25M Tris-HCl pH6.8, 0.2% SDS
0.5 ml 15 mg/ml ammonium persulphate
3.25 ml H_2O
8 μl TEMED

SDS PAGE RUNNING BUFFER (1L)

3.03 g Tris base
1.0 g SDS
14.4 g Glycine
pH should be 8.3-8.6 without adjustment

SDS PAGE loading buffer

1.25 ml 0.5M Tris-HCl pH 6.8
0.2 g SDS
2 ml glycerol
0.5 ml 2-mercaptoethanol
10 mg bromophenol blue
6.25 ml H_2O

COOMASIE BLUE STAIN

0.25% (w/v) Coomasie blue R250 made up in SDS PAGE destain.

SDS PAGE destain (1L)

300 ml methanol(AR)
100 ml glacial acetic acid
600 ml H_2O

2.9. Western Blotting

This procedure was adapted from the method of Towbin *et al* (1979). A Minitrans-blot electrophoretic transfer cell (Bio-Rad) was used following manufacturer's instructions using nitrocellulose membranes. Gels were blotted either 1 hour at 100 V or overnight at 30 V.

After transfer, the membranes were blocked with 3% skimmed milk in 100mM Tris pH 8.0 for 1 hour or overnight at room temperature, followed by three washes in the same buffer. Membranes were then incubated in a 1:500 primary antibody dilution in 100mM Tris pH 8.0 for 2 hours at room temperature after which three further washes in the same buffer were done. A dilution of 1:1000 rabbit anti-IgG serum conjugated to horseradish-peroxidase (Sigma) in 100mM Tris pH 8.0 was added to the membranes followed by 2 hours incubation at room temperature. After three washes in the same buffer, blots were developed in a solution containing 100ml 100mM Tris pH 8.0, 60mg chloronaphtol(in 3ml of methanol) and 50μl 30% H_2O_2.

2.10. Treatment of Polyacrylamide Gels for Autoradiography

After PAGE electrophoresis gels were placed in a bath containing 10% glacial acetic acid and allowed to fix for 15 min. The acid was poured off, the gel rinsed in distilled water and then soaked in Amersham Amplify (a fluorographic enhancer) for 20 min. The Amplify was drained off and the gel transferred to a Whatman 3MM paper, covered with saranwrap and dried at 80°C under vacuum on a gel drier for about 30 min. Dried gels were exposed to Kodak X-Omat AR film overnight at -70°C overnight.

2.11. IgA1 Protease Screening

2.11.1. Human IgA1 ^{35}S Labelling

Human IgA1 from human plasma was from Protogen AG. This immunoglobulin was labelled with ^{35}S-labelling reagent (SLR)(code SJ.440) from Amersham following manufacture's instructions.

2.11.2. Screening for IgA1 Protease

Fresh plaques resulting from a LambdaEMBL301 recombinant plating (Section 2.5.1.1.) were inoculated with a sterile toothpick in tubes containing 10ml L-broth and 10mM $MgCl_2$. After overnight growing in continuous shaking at 37°C, supernatants were assayed for IgA1 protease activity as follows: 100μl of each phage supernatant was mixed with 5μl of human IgA1 labelled with ^{35}S and incubated overnight at 37°C. A 15μl sample of each mixture was resolved in a 10% polyacrylamide gel electrophoresis (Section 2.8.), followed by amplifying treatment and autoradiography (Section 2.10.). IgA1 protease activity was manifested by the disappearance of the band corresponding to the heavy chain of IgA1 after cleavage by the protease.

2.12. Neuraminidase Assays

For all these assays the fluorogenic substrate 2'-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid (MUAN) (Sigma) was used.

2.12.1. Plate Assay for Neuraminidase Expression

This method is based in the technique developed by Vimr *et al* (1988). Plates containing recombinant phage were overlayed with 2ml of molten 0.75% agar containing MUAN (1mg/ml), 50mM sodium acetate pH 5.5, 150mM NaCl, and 4mM $CaCl_2$. Plates were incubated at 37°C for one hour. Neuraminidase positive clones were detected by their intense green-blue fluorescent halos when excited with 366 nm light from a UV transiluminator.

2.12.2. Neuraminidase Assay from *E. coli* Liquid Cultures

This method was based in the technique described by Berry *et al* (1988). Clones were grown in 200μl of L-broth, supplemented with the appropriate antibiotic in wells of microtiter plates. The plates were placed on a rocker at 37°C overnight. the plates were centrifuged at 2000 x g for 10 min, and the supernatant was discarded. The cells from each well were then resuspended in 70μl of a solution containing 10mM Tris-HCl pH 8.0, 10% sucrose, 50mM EDTA and 1mg/ml lysozyme. This was followed by an incubation at 37°C for one hour. Cell debris was precipitated by centrifugation at 2000 x g for 10 min, and 20 μl of each supernatant were transferred to a new microtiter plate where another 20 μl of a solution containing 20mM NaCl, 0.5M Sodium phosphate pH 7.0 were added , followed by the fluorogenic substrate MUAN to a final concentration of 7μg/ml. After incubation at 37°C for 20 min, neuraminidase activity was detected by fluorescence when placed on a UV transilluminator (366 nm).

2.12.3. Quantitative Assay for Neuraminidase Activity

This method was based on the neuraminidase assay used by Lock *et al* (1988). 10μl of enzyme sample was mixed with 10μl of 0.35% w/v MUAN as described by Lock *et al* (1988), and incubated for 10 min at 37°C. The reaction was stopped by the addition of 2ml of 0.5M sodium carbonate buffer pH 9.6. Neuraminidase activity was determined by measuring the amount of 4-methylumbelliferone (MU) which has been released. MU was quantitatively measured using a Perkin Elmer spectrofluorimeter (excitation wavelength 366nm, emission wavelength 446nm). Neuraminidase activity was given in

fluorimeter units (direct reading from spectrofluorimeter)

2.13. Pneumococcal Growth Curve

For the growth curve, 400ml of BHI were inoculated with 1.16 x 10^8 pneumococci from a frozen stock. From the 16 ml aliquots taken every hour, 1ml was centrifuged in a microfuge for 5min and the supernatant was assayed for neuraminidase activity (Section 2.12.3.). The 15ml left were disrupted at 12 psi in a cell disrupter (Stanstead Fluid Power Ltd) after which neuraminidase activity was measured (Section 2.12.3.). The activity obtained from the cell disruption minus the activity from the supernatants was considered as cell associated neuraminidase activity. Colony forming units (CFU) were measured by plating in blood agar base 10-fold dilutions of the aliquots taken every hour.

2.14. Pneumococcal Cell Fractionation

A 400 ml BHI broth was inoculated with 1.16 x 10^8 from a frozen stock and grown for 9 hours at 37°C. From this culture, 120 ml were centrifuged at 5,000 rpm for 10 min 4°C in a Heraeus-Christ centrifuge. The resulting supernatant was assayed for neuraminidase activity. The pellet was washed twice in PBS. The supernatants of these washes were also assayed for neuraminidase activity. The final pellet was resuspended in 14ml of PBS and subject to cell disruption at 12 psi in a cell disrupter (Stanstead Fluid Power Ltd). The disrupted cell suspension was centrifuged at 5,000 rpm at 5°C in a microfuge in order to sediment unbroken cells. The resulting supernatant was subject to two further centrifugations until no pellet of unbroken cells was present. From each of these centrifugations supernatants and pellets were assayed for neuraminidase activity. The final supernatant was centrifuged at 45,000 rpm in a Sorvall Ti75 rotor at 4°C and the resulting pellet was washed three times in PBS using the same speed and temperature. The final pellet containing the cell membranes was resuspended in 500μl of PBS and assayed for neuraminidase activity together with the supernatants and pellets of the previous 45K spins and PBS washes.

2.15. Electron Microscopy of Immunogold Stained Bacteria

This technique was carried out by the Electron Microscopy Unit (EMU), School of Biological Sciences, University of Leicester. An outline of the immunolabelling technique used for electron microscopy is summarized here.

Bacteria were fixed and embedded in resin by standard techniques used by EMU. Sections of this resin were blocked with phosphate buffered saline (PBS) pH 7.2 containing 0.1% Tween 20 (PBST) and 1% bovine serum albumin (BSA) for 10 min. This was followed by another blocking step with PBST containing 0.5% Gelatin for 20 min. Sections were then incubated at 4°C with primary antibody diluted, 1:100 for anti-neuraminidase antibodies, or 1:3000 for control anti-capsule antibodies, in PBST containing 1% BSA for 90 mins, followed by 5 washes with PBST. Section labelling was continued by incubation with secondary antibody (rabbit anti-IgG Au(10nm), Amersham) diluted 1:100 in PBST for 90 min. This was followed by 3 washes in PBST containing 1% BSA, and then 4 washes in PBST. Sections were fixed with 1% glutaraldehyde in PBS and treated with a standard techniques used by EMU for subsequent electron microscopy.

TABLE 2.1.BACTERIAL STRAINS

BACTERIAL STRAIN	GENOTYPE	REFERENCE
Escherichia coli strains:		
LE392	F^-, *hsd*R514 (rk^-,mk^-), *sup*E44, *sup*F58, *lac*Y1, *gal*K2, *gal*T22, *met*B1, *trp*R55, $Lamda^-$	Maniatis *etal*(1982)
Q359	rK^- mk^- *sup*E, *ton*A (P2)	Hadfield (1987)
JM101	F^-, *sup*E, *thi*, (*lac-pro*AB), $traD_{361}$, *pro*AB, *lac*IqZM15	Yanish-Perron *et al* (1982)
Streptococcus pneumoniae strains:		
3551	type I	Clinical isolate
R36A	(rough) type II derived	NCTC10319
11733	(rough) type II derived	ATCC11733
12213	(rough) type II derived	ATCC12213
DP1601	(rough) type II derived	Shoemaker *et al* (1974)
GBO5	type III	Clinical isolate
GB05A	type III	Clinical isolate
GBO5B	type III	Clinical isolate
SP01	untyped	Clinical isolate
SP02	untyped	Clinical isolate

TABLE 2.2. PLASMIDS

VECTOR	CHARACTERISTICS	REFERENCE
LambdaEMBL301	Phage cloning vector	Lathe *et al* (1987)
pJDC9	Pneumococcal cloning vector	Chen & Morrison (1987)
pJCP301	Neuraminidase recombinant	Berry *et al* (1988)
M13mp18/19	Sequencing vector	Yanisch-Perron *et al* (1985)
pUC18	Cloning vector	Yanisch-Perron *et al* (1985)

PART III

CLONING OF AN IgA1 PROTEASE GENE FROM *Streptococcus pneumoniae*

3.1. INTRODUCTION

Prevention of pneumococcal disease remains a serious matter, as the use of the twenty three valent polysaccharide vaccine does not provide complete protection for the population [Lee *et al* 1991]. For this reason, a search for pneumococcal factors that are involved in virulence is required to improve the current vaccine. IgA1 protease could be one of the candidates as it is expressed by many pathogens. This enzyme activity have been detected in most pneumococcal clinical isolates tested [Pratt & Boulnois 1987]. Isolation of the gene coding for IgA1 protease activity (*iga*) from the pneumococcus would allow a detailed study of this enzyme. The starting point would be to determine the DNA sequence of the cloned gene, followed by analysis of the predicted IgA1 protease amino acid sequence. This analysis would provide some hints in the understanding of the structure and function of this enzyme. The information obtained from these studies could be used, in the long term, to study the mechanism of action of pneumococcal IgA1 protease *in vivo*, once a suitable animal model has been found.

The cloning for this protease from pneumococcus has already been attempted with little success [Pratt & Boulnois 1987]. In the study of Pratt & Boulnois (1987) problems of instability of cloned pneumococcal DNA in the cosmid vector used were encountered. In this thesis, the isolation of a stable IgA1 protease clone in a lambda vector is described.

3.2. CONSTRUCTION OF A LambdaEMBL301 PNEUMOCOCCAL GENOMIC LIBRARY.

To overcome the instability problems found when cloning pneumococcal DNA, the construction of the genomic library was carried out in a LamdaEMBL vector [Lathe *et al* 1987][FIGURE 3.1]. This vector allows the cloning of 9 to 22 Kb DNA fragments. Also, by using an *in vitro* packaging technique for the lambda recombinant DNA, the cloning efficiency is many fold higher than plasmid cloning using standard DNA transformation protocols.

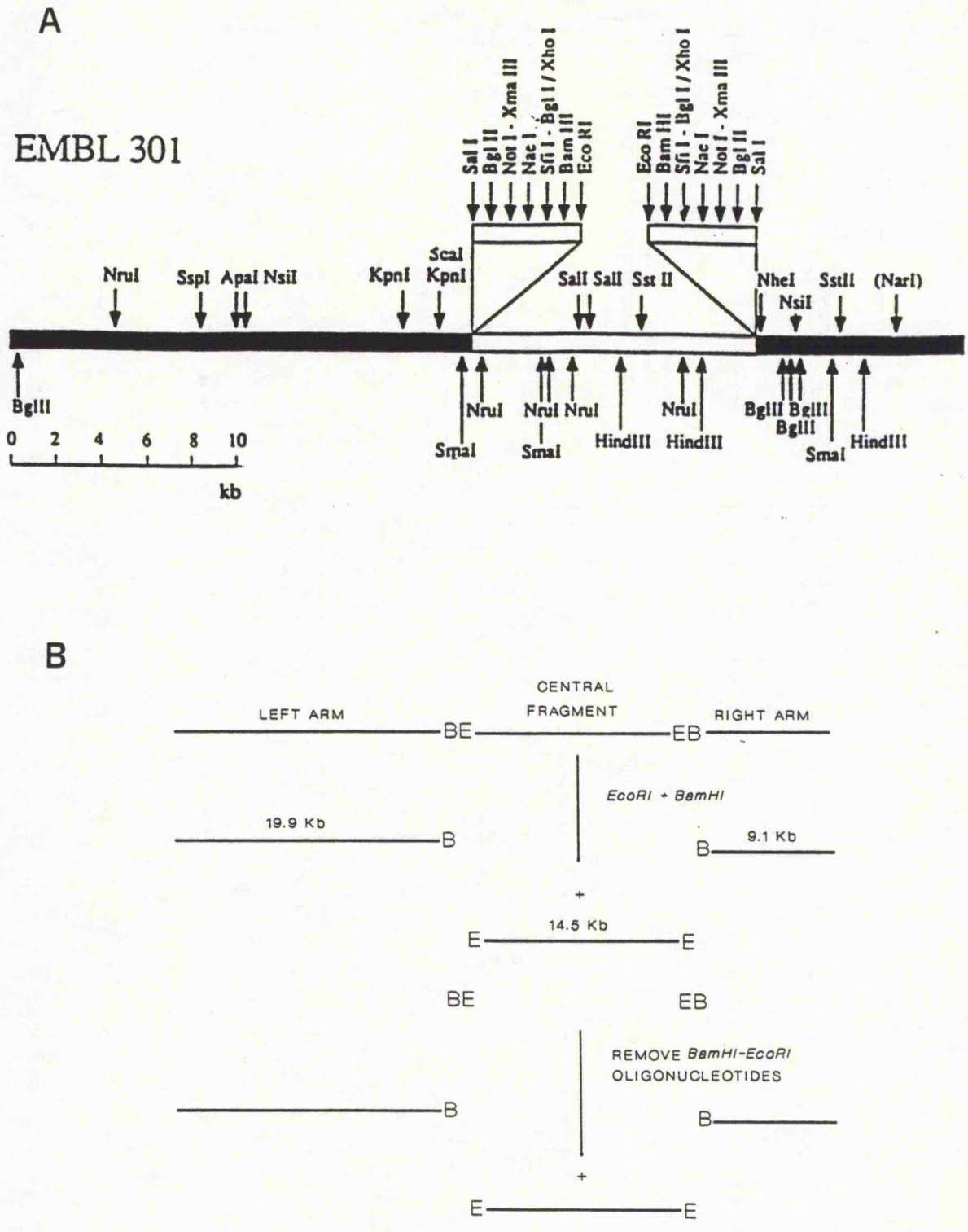


FIGURE 3.1. **Bacteriophage LambdaEMBL301.**

A. Restriction map of the bacteriophage LambdaEMBL301. The replaceable fragment is filled in white.

B. Preparation of Lambda arms. The central replaceable fragment is removed by digestion with *EcoR*I and *BamH*I. The *BamH*I-*EcoR*I oligonucleotides are then removed by ethanol precipitation of the arms.

S. pneumoniae R36A was the strain used for the construction of the genomic library. Several reasons led to the selection of this strain for the construction of the gene library. R36A is a rough nonpathogenic strain and the lack of virulence is probably due to the absence of capsule. R36A is derived from D39, a highly pathogenic typeII strain. This makes the manipulation of the strain less hazardous. Finally, this strain produces IgA1 protease as well as neuraminidase, which are the two main targets of the cloning described in this thesis.

Two pneumococcal genomic libraries were made. Although the first one was lost during the amplification process, the construction of both libraries will be reported, since a neuraminidase gene was cloned from the first library before the amplification process.

Construction of the first pneumococcal genomic library.

A stock of 16 μg of LamdaEMBL301 arms was prepared. After 3 ethanol precipitations, arms were tested for the absence of polylinker sequences that could interfere with the final ligation to the chromosomal DNA. This was accomplished by the addition of 1 μg of arms to 5μl final ligation reaction mixture. As a control 500ng of pUC18 digested with *Sst*I was used. Following an overnight incubation at 15^{0}C, the reaction products were analyzed by electrophoresis of the 500 ng of the control (pUC18/*Sst*I with ligase), alongside the control (pUC18/*Sst*I without ligase)[FIGURE 3.2]. Disappearance from the gel of the band corresponding to the pUC18 vector with ligase, indicated that ligation was successful.

A sample of the arms ligation reaction was packaged into Lambda phage heads and 5μl and 10 μl aliquots were plated on either *E. coli* LE392 or Q359. No plaques were found on either *E. coli* strain used, indicating that no religation of the arms with the 14.5 Kb middle fragment had occurred, and hence that polylinker sequences had been efficiently removed. Thus, the digested LambdaEMBL301 with *BamH*I and *Eco*RI was ready for the ligation to chromosomal DNA.

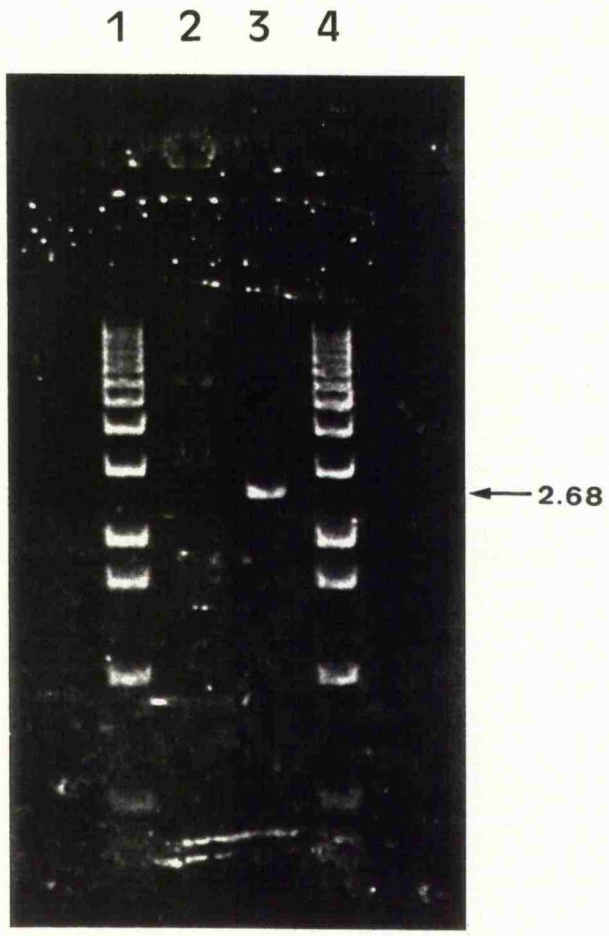


FIGURE 3.2. **Ligation to test purity of LambdaEMBL301 arms.**
To test the efficiency of the T4 ligase used for the ligation of LambdaEMBL301 arms, pUC18 digested with *Sst*I was ligated (track 2). Control (no T4 ligase) is shown in track (3). Tracks (1) and (4) contain 1Kb ladder DNA size marker.

Chromosomal DNA from the pneumococcal strain R36A was purified and aliquots of 1 μg were prepared for *Sau3A*I partial digestion and two fold dilutions of *Sau3A*I were added. After 1 h incubation at 37°C, it was found that the ideal size range of DNA for LambdaEMBL301 cloning (15-20 Kb) was achieved by using 1/64 units (digestion 6) and 1/128 units (digestion 7) of *Sau3A*I per μg of DNA [FIGURE 3.3].

It should be mentioned at this point that several attempts to construct the genomic library were made, following the method described by Maniatis *et al* (1982). In this method, the chromosomal DNA is subject to *Sau3A*I partial digestion, followed by a size fractionation step in a sucrose gradient. Several problems were found when using this technique. Only a small amount of digested DNA was recovered after precipitation from the gradient. Also, the recovered DNA did not ligate efficiently to the vector. Due to these difficulties, it was decided to omit the size fractionation step and ligate the DNA immediately after the partial digestion, without any further treatment.

Taking into account the compatibility of *BamH*I and *Sau3A*I ends from the lambdaEMBL301 arms and the partially digested chromosomal DNA respectively, the final ligation was set as follows: 500 ng of each selected digest (digestions 6 and 7) were mixed with 2 μg of LambdaEMBL301 arms in two separate ligation reactions of 20 μl final volume. The products of the reactions were analyzed by electrophoresis of 1 μl of each. The appearance of a major single band in the gel, compared to the three bands of the digested arms [FIGURE 3.4], indicated that the reaction was successful and the products were, therefore, packaged into Lambda phage heads *in vitro*. LambdaEMBL301 DNA was used as a control. Small aliquots of the packaging mixtures were plated on the *E. coli* LE392, which is nonselective for recombinants, and *E. coli* Q359 which selects recombinants. The amounts of packaging mixtures plated as well as the number of recombinants from each packaging reaction is shown in TABLE 3.1. Although the arms showed no religation ,indicating that no non-recombinants would be expected in the ligation with chromosomal DNA, surprisingly, larger number of plaques were obtained on LE392 than Q359.

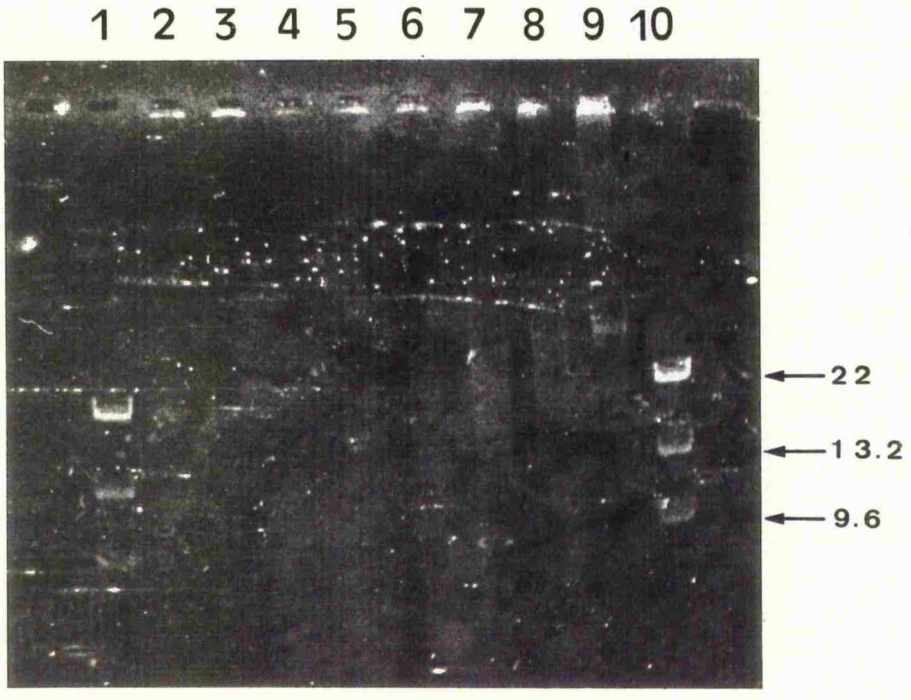


FIGURE 3.3. ***Sau3A*I partial digestion of chromosomal DNA.**

Aliquots of 1μg of chromosomal DNA were digested for an hour with the following ammunts of *Sau3A*I: (2) 1/2 units; (3) 1/4 units; (4) 1/8 units; (5) 1/16 units; (6) 1/32; (7) 1/64 units; (8) 1/128 units; (9) 1/256 units. Tracks (1) and (10) correspond to the DNA size marker which was wild type Lambda DNA digested with *Bgl*II.

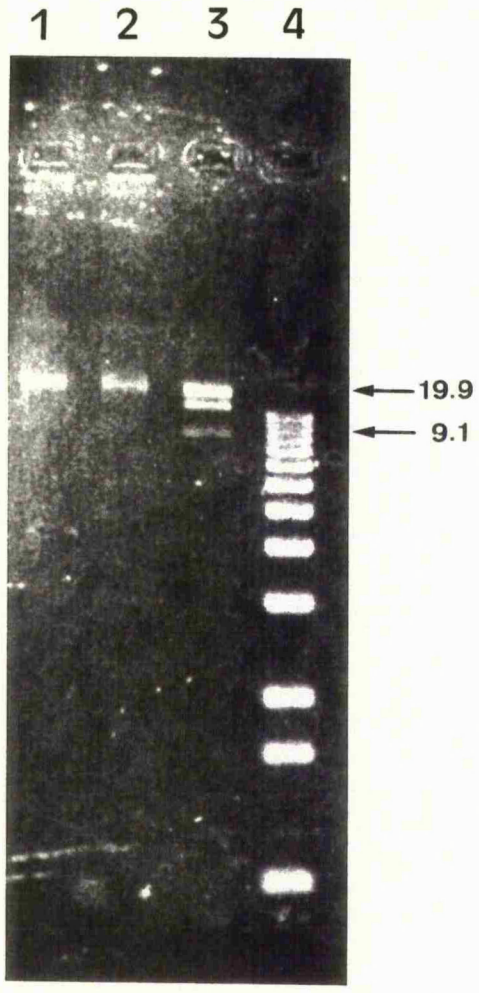


FIGURE 3.4. **Ligation of LambdaEMBL301 arms to chromosomal DNA.**

(1) Ligation of 500ng of digestion 6 to 2μg of LambdaEMBL301 arms.

(2) Ligation of 500ng of digestion 7 to 2μg of LambdaEMBL301 arms.

(3) LambdaEMBL301 arms without T4 ligase.

(4) 1Kb ladder DNA size marker.

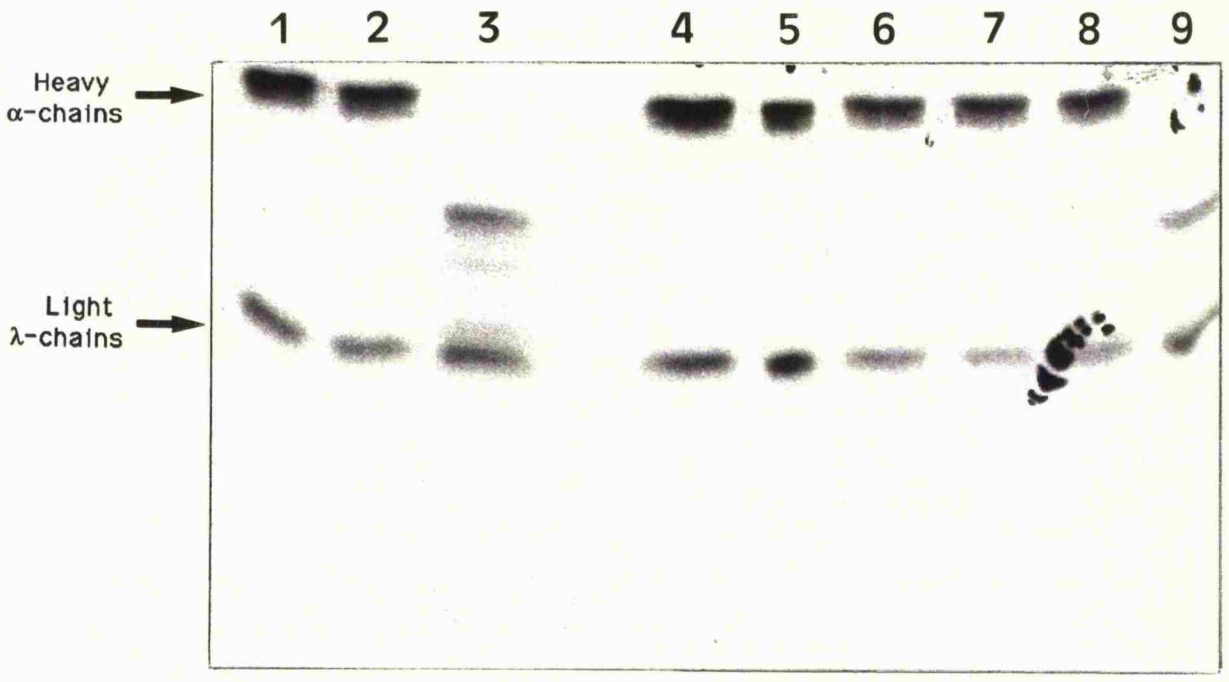


FIGURE 3.5. **Screening for IgA1 protease in pools of phage plaques.**

Human IgA1 was labelled with ^{35}S, and incubated with phage supernatant at 37°C overnight. These samples were subjected to SDS PAGE. The gel was dried and exposed to X-ray film overnight. Lane (3) shows the presence of IgA1 protease activity in pool 43. Lane (5) shows human IgA1 alone. Lane (9) shows human IgA1 incubated with pneumococcal R36A culture supernatant. Heavy and light chains of IgA1 are indicated.

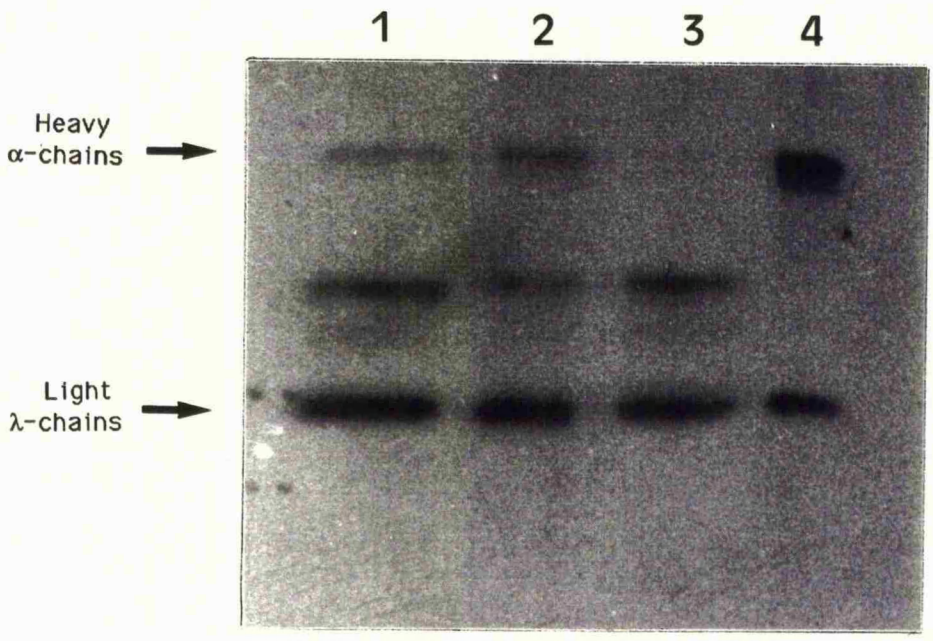


FIGURE 3.6. **Comparison of the IgA1 protease activity from phage pools 32 and 43.**
Screening was done as described in FIGURE 3.5. As a positive control pneumococcal R36A strain was incubated with ^{35}S-IgA1 this shown in track (1). As a negative control ^{35}S-IgA1 alone was used (4). Radiolabeled IgA1 incubated with supernatant from pools 32 (2) and 43 (3).Immunoglobulin heavy and light chains are indicated.

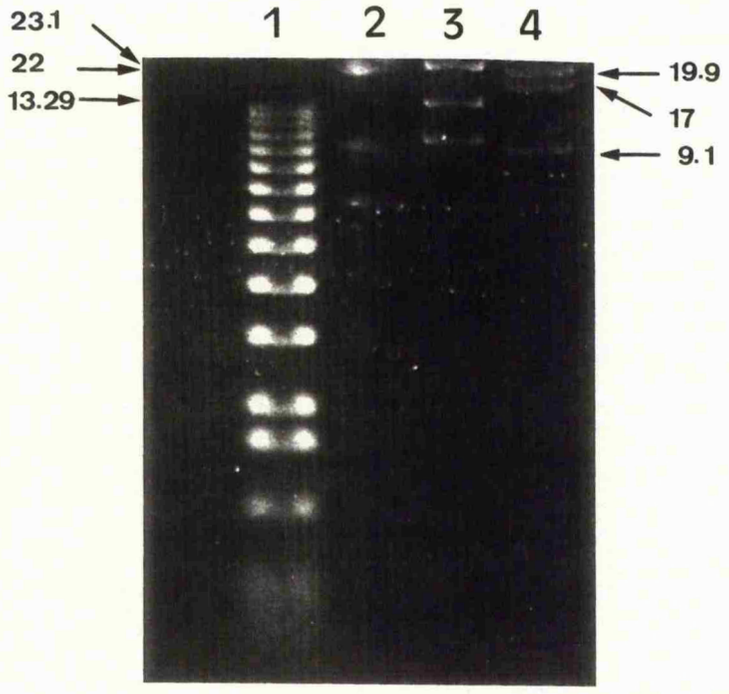


FIGURE 3.7. **Determination of the insert size in LambdaEMBL301-iga1.**
LambdaEMBL301-iga1 was digested with *Xho*I (4). The digestion was resolved in a 0.7% agarose gel electrophoresis. As markers 1Kb ladder (1), and wild type Lambda vector digested with *Hind*III (2) and *Bgl*II (3) were used. Fragment sizes are indicated.

PACKAGED SAMPLE	SAMPLE VOLUME	LE392	Q359
LIGATION 6	5 µl	79	22
	10 µl	183	109
	20 µl	388	204
LIGATION 7	5 µl	80	76
	10 µl	236	112
	20 µl	415	221
LambdaEMBL301	5 µl	>1000	0

TABLE 3.1. Number of plaques resulting from the plating of the packaging mixtures on LE392 and Q359.

The remainder of each ligation was packaged *in vitro* and the number of plaques obtained, is shown in [TABLE 3.2]. It should be noticed that a large aliquot of the packaging reaction corresponding to TABLE 3.1 was included here to obtain a more accurate final count of recombinants.

Several points should be noticed from [TABLE 3.2]. Firstly, the controls gave the expected results; no plaques were present from plating the arms whereas from LambdaEMBL301, a large number of plaques appeared when plated in LE392, but none were seen when plated on Q359. Secondly, smaller number of plaques was obtained when plating was done on Q359 rather than LE392. This was difficult to explain as the control gave no plaques. Finally, more than 25,000 recombinants was the estimated total number of plaques obtained. It was assumed that most of these were recombinants.

Unfortunately, when the library was being amplified, a technical problem caused its loss. It should now be pointed that some of the plates used for checking the package efficiency were screened for neuraminidase activity. One of them gave a plaque expressing this activity. This recombinant is discussed later in this thesis.

PACKAGING MIXTURES

	6	6A	6B	6C	7	7A	7B	7C	ARMS	λEMBL301
LE392	40 / 340	5 / 74	5 / 80	5 / 40	40 / 268	5 / 41	5 / 26	5 / 43	5 / 0	10^{-1} / $>10^{3}$
	40 / 452	10 / 113	10 / 61	10 / 47	40 / 235	10 / 126	10 / 67	10 / 46	10 / 0	10^{-2} / 890
	20 / 177	20 / 188	20 / 199	20 / 112	20 / 108	20 / 283	20 / 130	20 / 62	20 / 0	10^{-3} / 101
Q359	40 / 409	5 / 45	5 / 125	5 / 28	40 / ?	5 / 22	5 / 10	5 / 25	5 / 0	10^{-1} / 0
	40 / 410	10 / 89	10 / 39	10 / 24	40 / 152	10 / 92	10 / 3	10 / 44	10 / 0	10^{-2} / 0
	20 / 157	20 / 106	20 / 101	20 / 39	20 / 106	20 / 138	20 / 39	20 / 78	20 / 0	10^{-3} / 0
$\bar{X}$/µl	9.725	7.328	6.500	4.140	>4.340	10.020	3.920	4.250		
X/500µl	4862.5	3,664	3,250	2,070	2,170	5,010	1,960	2,125	TOTAL = > 25,111.5	

TABLE 3.2. **Results from the plating of the first R36A genomic library on the *E. coli* LE392 and Q359 strains. Numbers above the diagonals represent number of µl plated whilst those under the diagonal represent number of plaques obtained. X/µl: average number of plaques per µl of packaging mixture considering both *E. coli* strains together. X/500µl: number of plaques estimated in each packaging mixture.**

Construction of the second pneumococcal genomic library

The same lambdaEMBL301 arms and *Sau3AI* partial digested chromosomal DNA used for the first genomic library, were also used in the construction of the second library. In contrast to the previous library, only 10 µl of each ligation containing chromosomal DNA was packaged. Since a maximum of 5µl was allowed per packaging reaction, two reactions for each ligation mixture had to be set up. The estimated number of recombinants obtained from this library is shown in [TABLE 3.3]. Although in some platings larger number of plaques were obtained on LE392 than Q359, in others the opposite occurred. Since all the phages were expected to be recombinants and, in some platings, more recombinants were obtained when using the recombinant-selective strain it was assumed that all the plaques obtained were

recombinants. On this basis, a library of 21,950 recombinants had been constructed [TABLE 3.3].

	PACKAGING MIXTURES				
	6A	6B	7A	7B	LambdaEMBL301
LE392	8	112	33	62	380
Q359	2	82	54	86	0
PLAQUES/5μl	5	97	43.5	74	TOTAL = 21,950 PLAQ.
PLAQUES/500 μl	500	9,700	4,350	7,400	

TABLE 3.3. Number of plaques obtained from the plating of ligation mixtures "6" and "7". 5 μl of each packaging reaction was plated. As a control 10^{-2} μl of LambdaEMBL301 were used. The estimated number of recombinants from all the packagings is given.

For the amplification of the library *E. coli* LE392 was used. There were two reasons for this. Firstly, as already mentioned, all the phages from the packaging were considered to be recombinants. Secondly, the plaques were much larger when phage was plated in LE392, indicating that a higher yield in phage could be obtained by using LE392. This could be due to different phage infection efficiencies between LE392 and Q359. The amplification gave 3.35 x 10^4 recombinants/μl in a final volume of 16 ml.

3.3. CLONING OF PNEUMOCOCCAL IgA1 PROTEASE GENE.

For the cloning of the IgA1 protease a non-amplified pneumococcal genomic library was used to reduce the potential loss of unstable recombinants during the amplification process. The packaged mixture was plated on *E. coli* LE392. Low number of plaques per plate were used (an average of 160 plaques/plate) to avoid cross contamination when plaques were being picked. Two main screening techniques were tested for the screening of IgA1 protease production. They were based on plaque overlays with agar containing human ^{35}S-labelled IgA1 either free or bound to immunobeads. Nitrocellulose filters were placed on top of the agar, to collect the cleaved

radioactive IgA1 fragments. These filters were subsequently exposed to a film. A large background corresponding to a considerable diffusion of the labelled IgA1 was found in all cases, therefore, an alternative screening method was designed. For this screening, 735 plaques were picked, inoculated in pools of five in LB media and, following overnight growth at 37°C, culture supernatants were incubated with ^{35}S-labelled IgA1. These samples were electrophoresed in 10% polyacrylamide gels which, after drying, were used for autoradiography. If IgA1 protease activity was present, the heavy chain of the immunoglobulin would be cleaved and readily seen by autoradiography.

Of 56 pools IgA1 protease activity was seen in two [FIGURE 3.5]. Pool number 32 together with a control supernatant from R36A gave less cleavage than pool number 43. This was shown by the presence of some of the heavy chain that had not been hydrolyzed [FIGURE 3.6]. Since pool 43 gave complete cleavage, it was chosen for further study. 10-fold dilutions of the phage-containing supernatant from pool 43 was plated and single plaques were grown and screened for protease activity. A single recombinant expressing IgA1 protease activity was isolated out of 18 plaques examined. This recombinant, called LamdaEMBL301-Iga1, was subject to a large scale DNA purification, as described in materials and methods.

To analyze the insert size present in LambdaEMBL301-iga1, restriction digests using *Xho*I and *Not*I were used since these enzymes, known to have few sites in most DNAs, cut in the polylinker and are unlikely to cut in the insert. The products of these digests were resolved by electrophoresis in a 0.7% agarose gel. As presumed, none of these enzymes cut within the insert. The estimated insert size in LambdaEMBL301-iga1 was 17 Kb [FIGURE 3.7].

3.4. DNA HYBRIDIZATION STUDIES OF IgA1 PROTEASE PRODUCING CLONE.

To demonstrate the pneumococcal origin of the insert, a southern blot experiment was used. Chromosomal DNA from *S. pneumoniae* R36A and

LambdaEMBL301, were digested with *BamHI*, *EcoRI* and *HindIII*. These digests were resolved by agarose gel electrophoresis followed by blotting onto nitrocellulose filters of fragments larger than 1Kb. To generate a probe, LambdaEMBL301-iga1 was digested with *XhoI* and resolved by electrophoresis on low melting point agarose gel. An attempt to isolate the 17 Kb band corresponding to the insert DNA was done. Unfortunately, the 19.9Kb band corresponding to the left arm of the vector [FIG. 3.1.] was difficult to resolve from that of the insert, therefore, contamination of the insert could not be avoided, and the probe contained a mixture of both fragments. This probe was used in DNA hybridization studies after which, filters were washed in a 0.2 X SSC. The results from this southern blot are shown in FIGURE 3.8. Any bands over 19.9Kb in the hybridization of digested lambda recombinants would correspond to vector sequences plus insert DNA. From each set of digestions of chromosomal and recombinant DNAs with a particular enzyme, the presence in the blot of identical size bands, smaller that 19.9 Kb, would indicate the existence of the same DNA fragments in the chromosome of the pneumococcus and in LambdaEMBL301-iga. This is not conclusive proof of the pneumococcal origin of the gene, as nonspecific hybridization to pneumococcal DNA could have occurred due to the presence of lambda DNA in the probe. However, the presence of some bands of identical size in all digestions (FIG. 3.8: indicated by arrows) supports the idea that the hybridization is specific and shows the clone to contain pneumococcal DNA. In the case of the digestion with *EcoRI*, a partial digestion of LambdaEMBL301-iga1 may have caused the presence of many multiple bands.

3.5. CONCLUSION.

Two LambdaEMBL301 genomic libraries from *Streptococcus pneumoniae* R36A have been made. One of the libraries was lost during amplification. An IgA1 protease expressing phage was isolated from a non-amplified library. Two phages expressing IgA1 protease activity have been found. Only the phage expressing the highest protease activity has been purified. Preliminary southern blot analysis suggests that the insert carried by this phage was of pneumococcal origin.

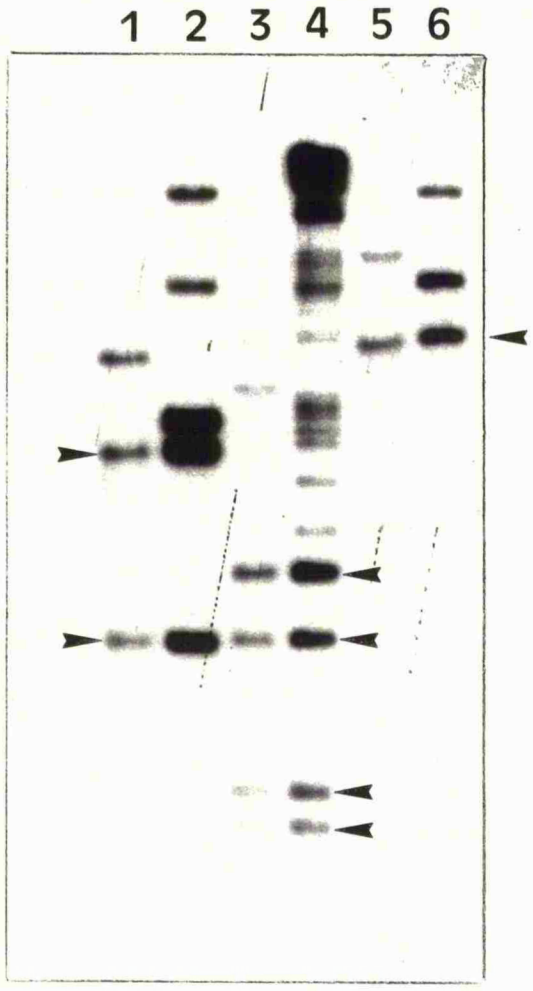


FIGURE 3.8. **Southern blot analysis to prove the pneumococcal origin of the LambdaEMBL301-iga1 insert.**

Chromosomal DNA from the pneumococcal strain R36A (1)(3)(5) together with DNA from LambdaEMBL301-iga1 (2)(4)(6) were digested with *BamH*I (1)(2), *Eco*RI (3)(4) or *Hind*III (5)(6). These DNAs were subject to Southern blot analysis. The filter was probed with the full length 17Kb insert from LambdaEMBL301-iga1 containing a small amount of the 19.9Kb LambdaEMBL301 left arm. The filter was washed in 0.2 X SSC. Arrows indicate DNA fragments from LambdaEMBL301-iga of a possible pneumococcal origin.

PART IV

CLONING OF A NEURAMINIDASE GENE FROM *Streptococcus pneumoniae*

4.1. INTRODUCTION.

For many years, a role for neuraminidase in virulence has been suggested. However, no conclusive results to support this contention have been reported. In the case of the neuraminidase from *Streptococcus pneumoniae*, there is some circumstantial evidence suggesting a possible role in virulence for this enzyme. One of the approaches to study this role in more detail would be to clone the neuraminidase gene from the pneumococcus. The DNA sequence determination of this gene would allow the analysis of the neuraminidase predicted amino acid sequence. This analysis would reveal useful information about the structure and function of this enzyme. This information could be used for the understanding of the its mechanism of action *in vivo*, especially during the establishment of the molecular version of Kock's postulates for this enzyme (see introduction).

A gene for pneumococcal neuraminidase was cloned in a cosmid vector [Berry *et al* 1988]. The purified recombinant, although expressing neuraminidase activity, suffered a spontaneous deletion, and lost part of the neuraminidase gene. An attempt to clone the intact gene into a plasmid vector yielded a recombinant carrying DNA sequences from the neuraminidase gene that failed to express the enzyme activity [Berry *et al* 1988]. It was postulated that this DNA fragment was very susceptible to deletions when cloned in the plasmid vectors in *E. coli*.

The instability problems found with the vectors used by Berry *et al* (1988) led us to attempt the cloning of the pneumococcal neuraminidase gene using the LambdaEMBL301 vector. Two different screening methods were used. One of them consisted of a test for neuraminidase activity using a fluorogenic substrate. The second one involved DNA hybridization using a probe containing an internal fragment of the neuraminidase gene cloned by Berry *et al* (1988).

4.2. CLONING OF A PNEUMOCOCCAL NEURAMINIDASE GENE.

The nonamplified pneumococcal DNA library, described in Section 3.2, was screened for the neuraminidase gene by examining clones for expression of

functional enzyme.

After plating, a total of 800 recombinant plaques were screened for neuraminidase activity by using an agar overlay containing the fluorogenic neuraminidase substrate MUAN. One recombinant produced fluorescence when exposed to U.V. light. This phage was subjected to plaque purification. A plug of agar containing the recombinant was washed in buffer to release the phage, which was plated on *E. coli* LE392. Only two plaques were obtained, grown in liquid and the culture supernatants were tested for neuraminidase activity. One gave a neuraminidase activity of 2.5 x 10^5 fluorimeter units per ml, while the other gave no neuraminidase activity. Phage DNA from the recombinant expressing neuraminidase activity, designated LambdaEMBL301-Neu1, was purified.

The insert size in LambdaEMBL301-Neu1 was determined by a *SalI* restriction digest since this enzyme cuts in the vector's polylinker and only cut once in the insert. In addition to lambda arms this digestion gave two bands of 15 Kb and 3.5 Kb of insert DNA [FIGURE 4.1.]. Therefore, LambdaEMBL301-Neu1 contained 18.5 kb of insert DNA.

4.3. SUBCLONING OF THE NEURAMINIDASE GENE OF LambdaEMBL301-Neu1 INTO pJDC9.

Unstable recombinant plasmids are frequently encountered when cloning pneumococcal DNA into conventional *E. coli* plasmid vectors. It was shown by Chen & Morrison (1987) that random pneumococcal DNA fragments cloned in *E. coli* exhibited a higher frequency of strong promoter activity than did similarly cloned *E. coli* fragments. The same workers designed a vector called pJDC9 [FIGURE 4.2]. This vector contains a *lacZα'* multiple cloning site surrounded by transcriptional terminators which protect plasmid control elements from excessive transcription originated at strong pneumococcal promoters. This vector also has an erythromycin-resistant marker expressed both in *E. coli* and *S. pneumoniae*. In an attempt to reduce potential instability problems pJDC9 was the vector chosen for use in subcloning the insert from LambdaEMBL301-Neu1.

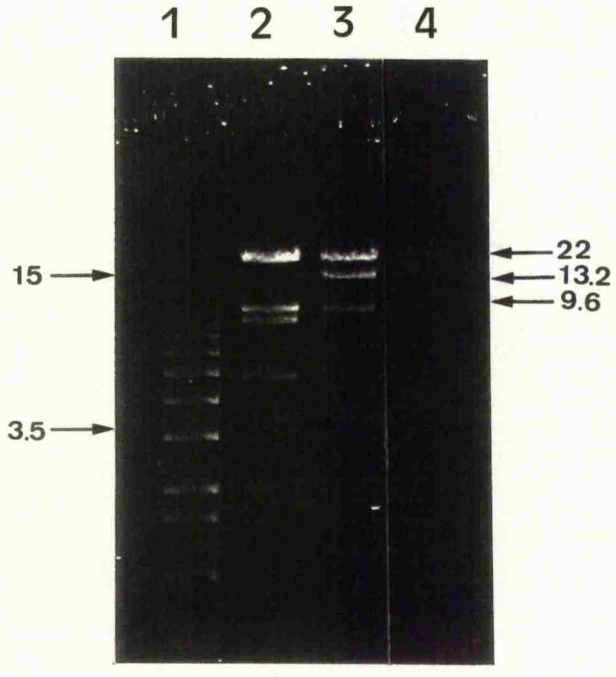


FIGURE 4.1. **Determination of the insert size in LambdaEMBL301-Neu1.**
LambdaEMBL301-Neu1 was digested with *Sal*I (3). As a reference LambdaEMBL301 vector was digested with the same enzyme (2). DNA size markers were 1Kb ladder (1) and wild type Lambda DNA digested with *Bgl*II (top fragment sizes 22, 13.2 and 9.6 Kb)(4).

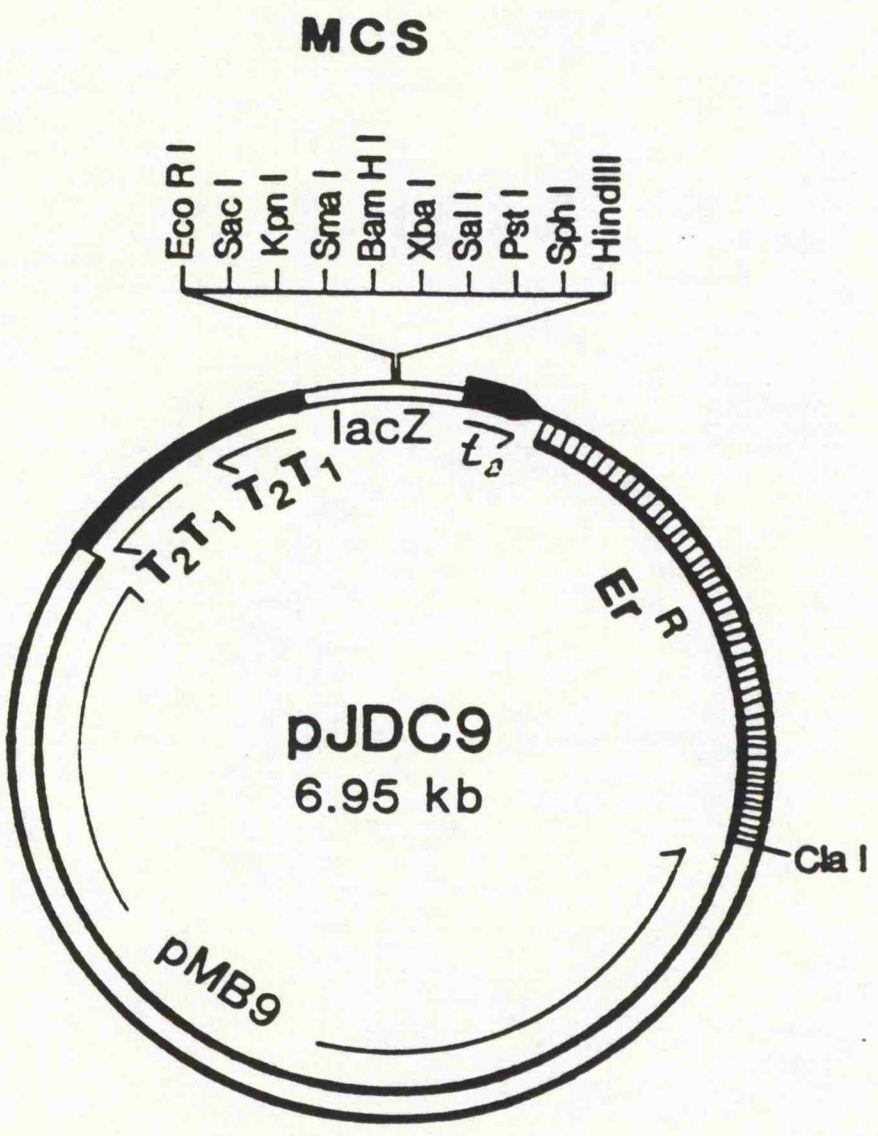


FIGURE 4.2. **Map of the plasmid pJDC9.**

This vector carries an erythromycin-resistant marker (Er). T_1, T_2 and t_E represent transcriptional terminators from pDS1 and pKK232-8. MCS represents the multiple cloning site. The pMB9 fragment is derived from pBR322. Figure from Chen & Morrison (1987).

The *E. coli* strain selected as host for subcloning was JM101 since it allows colour selection of recombinant plasmids carrying the *lacZα'* by the addition of the inducer IPTG and the *β*-galactosidase substrate X-Gal. In addition, this strain is also the most easily transformable of the JM series of strains.

The subcloning of fragments of insert DNA from LambdaEMBL301-neu1 into pJDC9 was achieved without prior restriction mapping. A set of Sau3AI partial digestions of the whole lambda recombinant, which yielded DNA fragments ranging from 3 to 12 Kb, was prepared [FIGURE 4.3]. Size fractionation of the digested products was done. Fragments selected were: N1=3-4 Kb; N2=4-5 Kb; N3=5-7 Kb; N4=7-10Kb; N5=10-12 Kb. The purified DNA fragments are shown in FIGURE 4.4.

To clone these *Sau3A*I-generated DNA fragments, pJDC9 was digested with *BamH*I and treated with alkaline phosphatase. The success of the alkaline phosphatase reaction was checked by religation reaction of the dephosphorylated vector followed by transformation into a suitably competent *E. coli* strain.

A 1:4 ratio of insert to vector was used in ligation reaction, and the products transformed into *E. coli* JM101 and selected on minimal media containing 1mg/ml erythromycin, with IPTG and X-Gal. The white colonies from each transformation were streaked onto fresh selective plates. The resulting colonies were grown in microtiter plates. Cells were lysed and assayed for neuraminidase production by the addition of the neuraminidase fluorogenic substrate MUAN. Positives were detected by the production of fluorescence in the wells of the microtiter plate, when placed onto a U.V. transilluminator [FIGURE 4.5]. The following table shows the number of recombinant colonies obtained from each transformation and the number expressing neuraminidase:

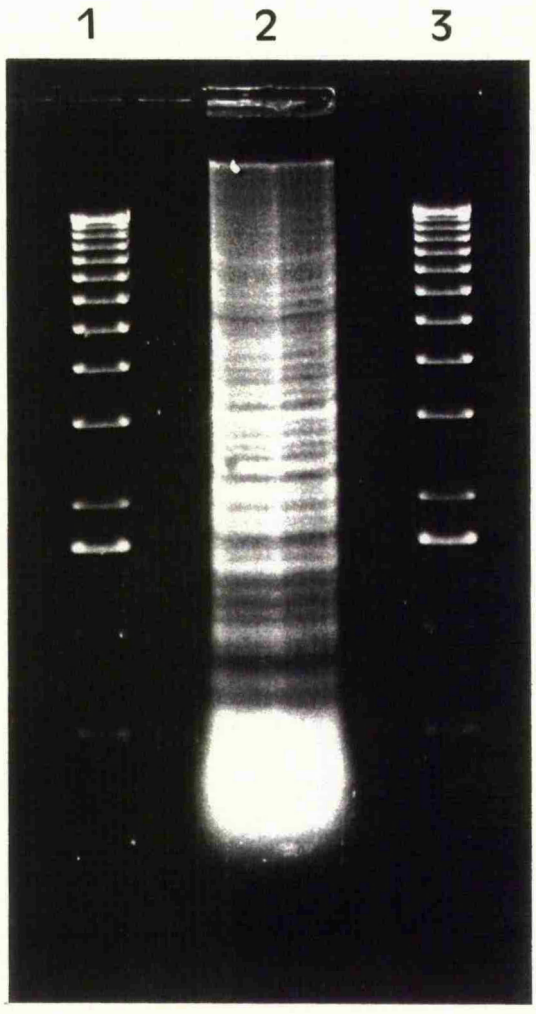


FIGURE 4.3. ***Sau3A*I partial digestion of LamdaEMBL301-Neu1.**

For this digestion 1/32 units/μg of *Sau3A*I were added to 40μg of Lambda neuraminidase recombinant followed by incubation at 37°C for 1 hour. Digested products were resolved in 1% low melting point agarose electrophoresis (2). 1Kb ladder was used as DNA size marker (1)(3).

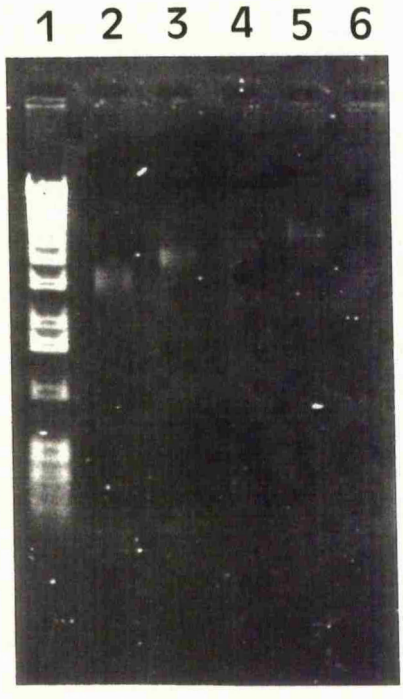


FIGURE 4.4. **Purified DNA fragments from *Sau3A*I-partial digested LambdaEMBL301-Neu1.** Partial digested DNA was size fractionated from a low melting point agarose gel. Purified DNA range fragments were: N1=3-4Kb (2); N2=4-5Kb (3); N3=5-7Kb (4); N4=7-10Kb (5); N5=10-12Kb (6). 1 Kb ladder was used as DNA size marker (1).

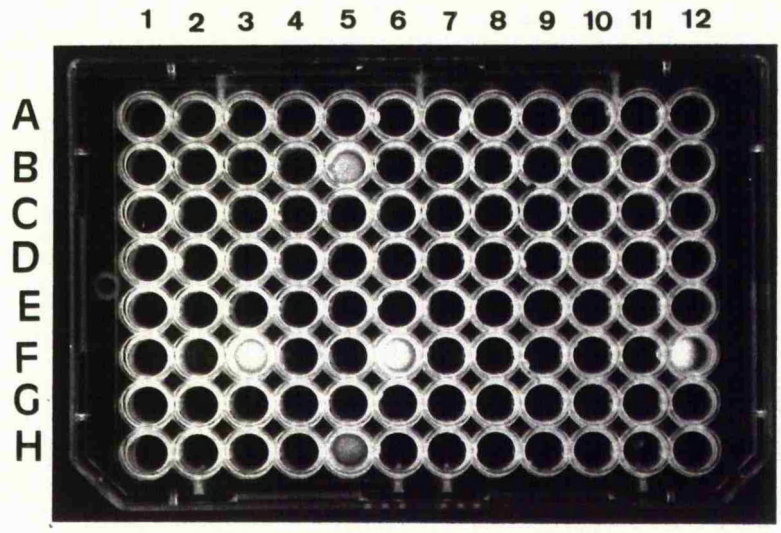


FIGURE 4.5. **Screening for neuraminidase recombinants.**

Size fractionated DNA from LambdaEMBL301-Neu1 was ligated to pJDC9 and transformed into *E. coli* JM101. Bacteria carrying the recombinants were grown in microtiter plates. After lysis, supernatants from these cultures were incubated in fresh microtiter plates with the neuraminidase fluorogenic substrate MUAN. Positive recombinants were detected by the production of fluorescence when the microtiter plate was placed on a UV transilluminator (366 nm). Positives are present in wells B5, F3, F6, F24 and H5. As a negative control a lysate of *E. coli* JM101 was incubated with MUAN (H11 & H12).

PREPARATION	No. COLONIES	No. POSITIVES
N1	38	1
N2	34	3
N3	50	0
N4	21	4
N5	5	1

The neuraminidase producing colonies from N1 and N2 were used for large scale plasmid preparation, since they were expected to carry the smallest insert of pneumococcal DNA capable of coding for neuraminidase activity. Cells from each plasmid preparation were tested for neuraminidase activity, to ensure stability of expression. The active clones detected were as follows:

PREPARATION	PLASMID	FLUORIMETER UNITS/ml
N1	pMC1250	82,540
N2	pMC2150	99,300
N2	pMC2180	98,700
N2	pMC2240	99,000

The four plasmids were digested with *Hind*III, *Pst*I and *Sal*I in order to determine their size [FIGURE 4.6]. The smallest plasmid was pMC1250 which had an insert of 3.55 Kb, while the other three gave the same restriction pattern, with a 4.3Kb insert. It was concluded that these three recombinants were identical. The one chosen for further studies was pMC2150 since it gave the highest neuraminidase activity.

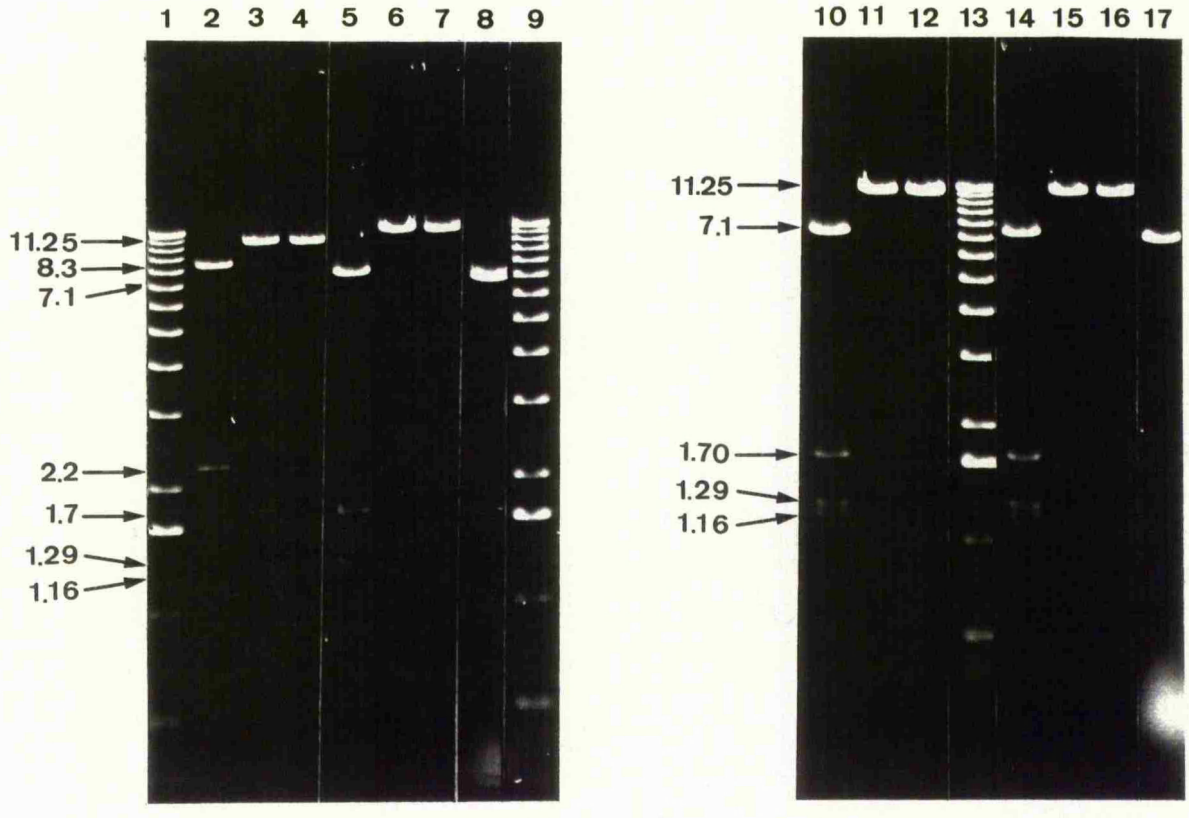


FIGURE 4.6. **Determination of the insert size of pMC1250 (2)(3)(4), pMC2150 (5)(6)(7), pMC2180 (10)(11)(12) and pMC2240 (14)(15)(16).**

These plasmids were digested with *Hind*III (2)(5)(10)(14), *Pst*I (3)(6)(11)(15) or *Sal*I (4)(7)(12)(16). As a reference pJDC9 was digested with *Eco*RI (8)(17). 1Kb ladder was used as a DNA size marker (1)(9)(13). DNA sizes are indicated in kilobases.

4.4. DNA ANALYSIS.

A detailed restriction map of pMC2150 was required to allow a primary comparison of this neuraminidase clone with the one isolated by Berry *et al* (1988) and to enable further subcloning into M13 vectors for DNA sequence determination.

4.4.1. DNA RESTRICTION MAPPING.

The restriction map of pMC2150 was constructed by performing single, double and triple restriction enzyme digests. The deduced DNA restriction map is shown in FIGURE 4.7. *EcoR*I cuts the plasmid several times but the cleavage sites were not mapped. On the other hand, restriction sites for *Pst*I, *Sal*I, *Sma*I and *Kpn*I were absent in the insert.

The restriction map of the insert was compared to that of pJCP301 [Berry *et al* 1988][FIGURE 4.7]. In pMC2150 no sites for *Pst*I nor *Kpn*I were found, while such sites were present in pJCP301. Also, restriction sites for *Cla*I, *EcoR*I, *Hind*III and *Xba*I mapped in pMC2150, were absent in pJCP301 [Berry *et al* 1988]. One possible explanation for this difference could be the spontaneous deletion that pJCP301 suffered when purified since, instead of a large insert expected of a cosmid clone, only a 3Kb insert was found. This might imply rearrangements within the original clone that could give rise to the different restriction map. This possibility cannot be considered as valid, since it has been shown that the insert carried by pJCP301 corresponds to a contiguous sequence in the chromosome of *Streptococcus pneumoniae* [Berry *et al* 1988]. Therefore this suggested the presence of two different neuraminidase genes in the pneumococcus, although the possibility of serotype specific neuraminidases could not be discarded, since the neuraminidase from pMC2150 and pJCP301 were cloned from a typeII and a typeI pneumococcal strain, respectively.

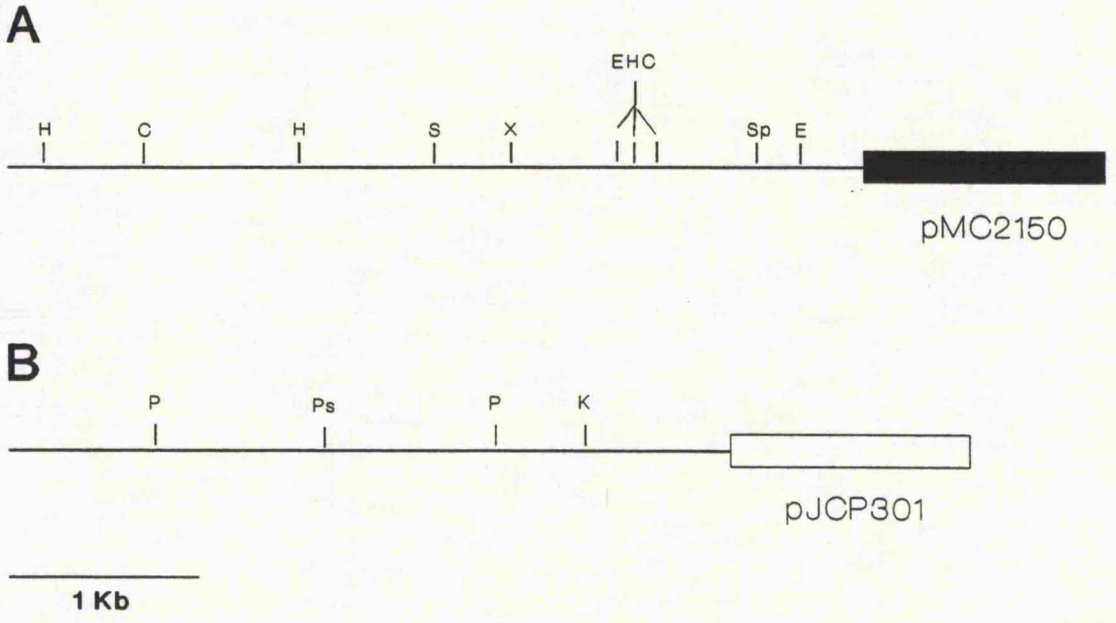


FIGURE 4.7. **Restriction map of cloned pneumococcal DNA in pMC2150 (A) and pJCP301 (B).** Restriction endonuclease sites: H, *Hind*III; C, *Cla*I;, S, *Sac*I; X, *Xba*I; E, *Eco*RV; Sp, *Sph*I;P, *Pvu*II; Ps, *Pst*I; K, *Kpn*I. No sites for *Sma*I, *Kpn*I, *Pst*I, or *Sal*I were present in the insert in pMC2150. Berry *et al* (1988) found no sites in pJCP301 for *Cla*I, *Eco*RI, *BamH*I, *Hind*III, or *Xba*I. Closed and open boxes indicate the vectors pJDC9 (A) and pHC79 (B), respectively.

4.4.2. DNA COMPARISON BY HYBRIDIZATION.

The inserts in plasmids pMC2150 and pJCP301 did not show any similarities in their restriction maps. For this reason DNA hybridization studies between the two clones were made to investigate whether the two neuraminidase genes were distinct.

The first approach used was dot blot hybridization. Chromosomal DNA from three different pneumococcal strains was used in these experiments. These three strains were the non-encapsulated laboratory strains R36A and DP1601 (both typeII derived), and GB05, a typeIII clinical isolate. All of these DNAs were purified using protocol I (Materials and Methods). The chromosomal DNA from R36A was different from that used for the construction of the genomic library. LambdaEMBL301-Neu1 DNA was also included. The

recombinant pJCP301 was used as a positive control and lambdaEMBL301 as a negative control. Three different amounts of each DNA (1.4 μg, 0.7 μg and 0.35 μg) were blotted onto nitrocellulose filters. These filters were probed with the 1.6 Kb *Pvu*II fragment from pJCP301 [FIGURE 4.7] and washed finally in 0.2 X SSC. The result of this dot blot is shown in FIGURE 4.8. The positive and negative controls worked as expected, but no hybridization was detected with DNA from LambdaEMBL301-Neu1 under the conditions used. On the other hand, hybridization was found for the three pneumococcal strains used. These results supported the idea of the presence in the same pneumococcal strain of the neuraminidases expressed from pMC2150 and pJCP301.

To ensure that the dot blot results were not a product of nonspecific hybridization of the probe with different regions in the chromosome, and that the lack of hybridization found in the case of the LambdaEMBL301-Neu1, was not a consequence of the high SSC concentration used, two sets of duplicate southern blot hybridizations were designed. For the first [FIGURE 4.9.A], chromosomal DNAs from *S. pneumoniae* R36A, DP1601 and GB05B, digested with *Cla*I, were probed with the 4.3 Kb insert from pMC2150. As a negative and positive controls, *Cla*I digested *E. coli* chromosomal DNA and the 4.3Kb purified insert from pMC2150 were used, respectively. The 1.6 Kb *Pvu*II fragment from pJCP301 was also included in this blot. For the second blot [FIGURE 4.9.B], chromosomal DNA from *S. pneumoniae* R36A, DP1601 and GB05B, was digested with *Pvu*II, and probed with the 1.6 Kb *Pvu*II insert from pJCP301 which was also used as a positive control. As a negative control, *E. coli* chromosomal DNA digested with *Pvu*II was used. The 4.3 Kb insert from pMC2150 was also included.

Each of the nitrocellulose filters from both sets of blots were washed in either 2 X SSC or 0.2 X SSC. In both, FIGURE 4.9. (A) and (B), only results from the 2 X SSC final wash are shown as the results from the two different stringencies were identical. No hybridization was detected in the negative control [FIGURE 4.9.A]. The positive control showed three bands of 0.75, 2.45 and 1.1 Kb, corresponding to the 4.3 Kb insert digested with

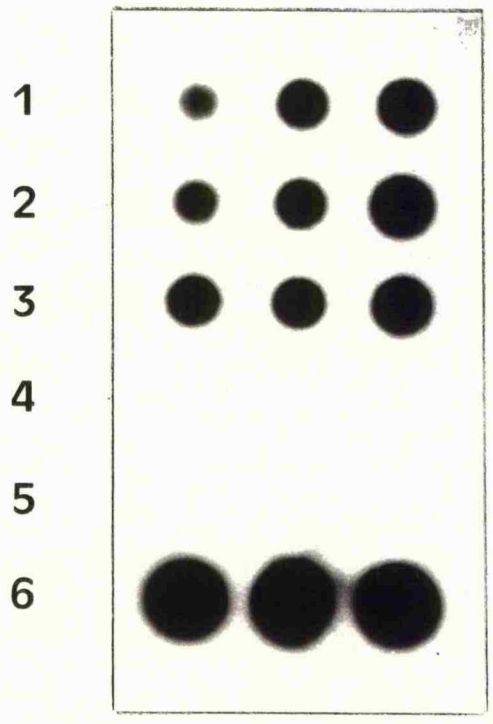


FIGURE 4.8. **Dot blot hybridization showing the presence of two distinct neuraminidase genes in pneumococci.**

Three different amounts of DNA from *S. pneumoniae* R36A (1), GB05B (2), DP1601 (3) were blotted together with DNA from LambdaEMBL301-Neu1 (4), LambdaEMBL301 (5) and the 1.6Kb *Pvu*II internal fragment from pJCP301 (6). This 1.6Kb *Pvu*II fragment was used as a probe. Filters were washed in 0.2 X SSC.

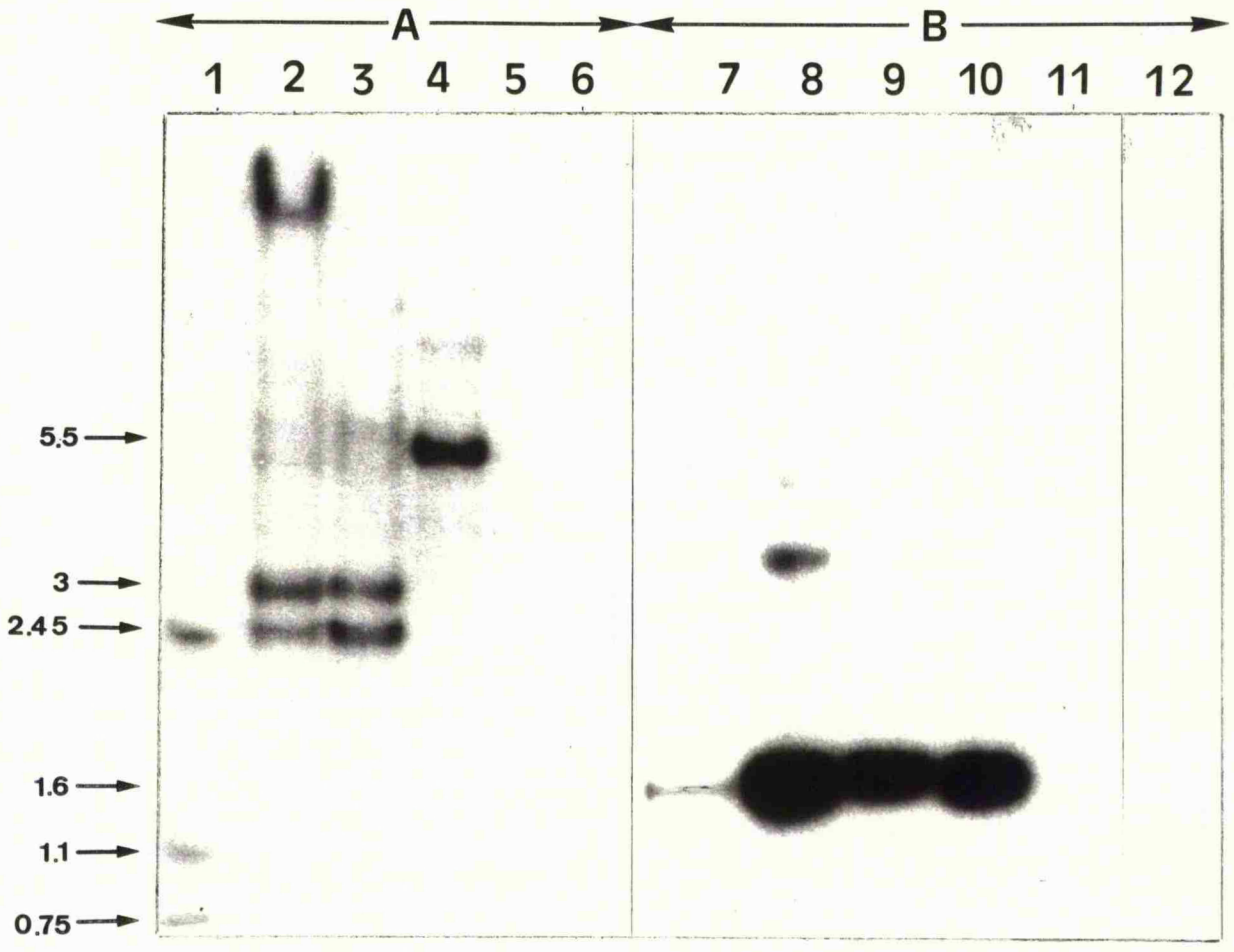


FIGURE 4.9. **Southern blot analysis of neuraminidase genes.**

A. Chromosomal DNAs from *S. pneumoniae* R36A (2), DP1601 (3) and GB05B (4) were digested with *Cla*I. The 1.6 Kb *Pvu*II internal fragment from pJCP301 was also included (5). As a positive control, the purified insert of pMC2150 (4.3Kb) was digested with *Cla*I (1) and, as a negative control, chromosomal DNA from *E. coli* digested with the same enzyme was used (6). These filters were probed with the 4.3 Kb insert from pMC2150.

B. Chromosomal DNAs from *S. pneumoniae* R36A (8), DP1601 (9) and GB05B (10) were digested with *Pvu*II. The purified insert from pMC2150 (4.3Kb), digested with *Cla*I, was also included (11). As a positive control the 1.6 Kb *Pvu*II internal fragment from pJCP301 was used (7), whereas as a negative control chromosomal DNA from *E. coli* digested with the same enzyme was included (12). These DNAs were probed with the 1.6Kb *Pvu*II internal fragment from pJCP301(6).

DNA sizes are indicated in Kilobases.

*Cla*I. Chromosomal DNA from the pneumococcal strains R36A and DP1601 showed hybridization of two 3.0 and 2.45 Kb *Cla*I fragments, the last corresponding to the internal *Cla*I fragment from pMC2150. In the case of GB05, a 5.5 Kb *Cla*I fragment hybridized with the pMC2150-derived probe. This suggests that there is a polymorphism in the neuraminidase gene from this clinical isolate, and one of the *Cla*I sites has disappeared. Given the restriction map of the insert present in pMC2150 [FIGURE 4.7.], which contains two sites for *Cla*I, it was expected that three *Cla*I fragments would hybridize in the R36A genome. Only two were apparent, although it is possible that one of the observed bands is a doublet. No hybridization was observed with pJCP301 at either of the two different washes used.

As expected, a single 1.6 Kb *Pvu*II fragment in the positive control, and R36A, DP1601 and GB05 hybridized to the pJCP301-derived probe [FIGURE 4.9.B]. In contrast, no hybridization was found with DNA from pMC2150 or the negative control. Thus the sequences coding for neuraminidase activity from pJCP301 and pMC2150 are distinct and non-serotype specific, since they are present in each pneumococcal stain tested.

4.5. RECOGNITION OF NEURAMINIDASE BY POLYCLONAL ANTIBODIES

Lock *et al* (1988) purified a neuraminidase from a type II clinical isolate of *Streptococcus pneumoniae* which was used subsequently to raise polyclonal antibodies in the mouse and rabbit. These antibodies reacted in western blots with the neuraminidase from pJCP301 and the pneumococcus [Berry *et al* 1988]. Although there was no homology in terms of DNA hybridization between the pneumococcal DNA from pJCP301 and LambdaEMBL301-Neu1, the possibility of protein similarities between the products with neuraminidase activity from the two clones could not be discounted.

To test for a serological relationship between these proteins, a western blot using rabbit antibodies to purified neuraminidase [Lock *et al* 1988] (kindly donated by Dr. James Paton) was done [FIGURE 4.10]. The anti-neuraminidase antibodies reacted with the LambdaEMBL301-Neu1 lysate and

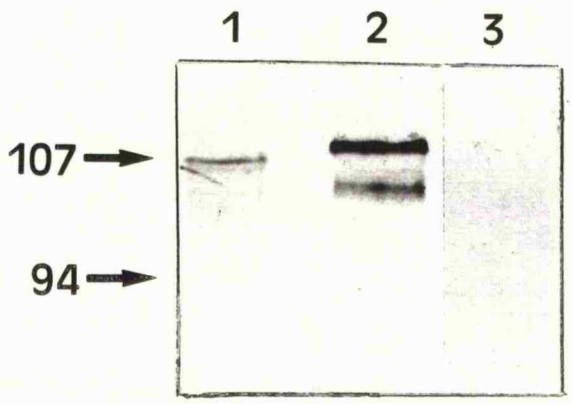


FIGURE 4.10. **Western blot analysis of the neuraminidase from LambdaEMBL301-neu1.** For the western blot purified neuraminidase from the pneumococcus (Lock *et al* 1988)(1) together with a lysate from LambdaEMBL301-Neu1 in *E.coli* JM101 (2) were used. As a negative control an *E. coli* lysate of vector LambdaEMBL301 was included (3). These samples were resolved by SDS PAGE electrophoresis. After blotting, membranes were incubated with rabbit anti-neuraminidase antibodies (Lock *et al* 1988). Numbers at the side indicate Kilodaltons.

also with the neuraminidase purified by Lock *et al* (1988). No antibody binding to protein in the LambdaEMBL301 lysate control was seen. This western blot showed several faint bands below the 107KD major band for the purified neuraminidase. Also a less intense band below the largest one is present in the lambda lysate. The appearance of multiple bands was probably due to the susceptibility to proteolysis of neuraminidase, as reported by Lock *et al* (1988). However, the neuraminidase from LambdaEMBL301-Neu1 seemed slightly larger than the one purified from the pneumococcus 107KD. Therefore, the antibodies that recognize the products from pJCP301 also react with the products from LambdaEMBL301-Neu1. This would indicate some similarities between the proteins encoded by these recombinants. Alternatively, as the pneumococcal neuraminidase was susceptible to proteolytic degradation during its purification, the purified protein could contain a mixture of the products from pJCP301 and LambdaEMBL301-Neu1.

4.6. SIMULTANEOUS PRESENCE OF THE TWO NEURAMINIDASE GENES IN SEVERAL PNEUMOCOCCAL STRAINS

As shown, the DNA sequences coding for neuraminidase activity from pJCP301 and pMC2150 are present in both the typeII-derived pneumococcal strains R36A and DP1601, and also in the typeIII strain GB05. To determine if the presence of these two neuraminidase genes is a general feature of *Streptococcus pneumoniae*, several pneumococcal strains were tested in dot blot experiments.

The following strains of *Streptococcus pneumoniae* were used: SPO1 and SPO2, both untyped clinical isolates; 3551, a type I clinical isolate used for the cloning of the neuraminidase present in pJCP301; the ATCC strains 11733 and 12213; and, the typeIII pneumococcal clinical isolates GB05-A and GB05-B.

Dotblot analysis was carried out using ^{32}P for probe labelling. For the ATCC strains an AMERSHAM ECL (Enhanced Chemiluminescent) detection kit, involving the use of a non-radioactive probe was used. All DNA samples

were probed with the 1.6 Kb *Pvu*II fragment from pJCP301, and with the 1.7 Kb *Hind*III fragment from pMC2150 [FIGURE 4.7]. As positive or negative controls pJCP301 and pMC2150 were used.

The results are shown in FIGURE 4.11. The intensity of the signals should be ignored since the preparations of chromosomal DNA did not allow analysis of equal amounts of DNA. Hybridization of both neuraminidase genes was detected in all the strains tested. These results support the conclusion that the neuraminidases from pJCP301 and pMC2150 are simultaneously present in all strains of pneumococci tested.

4.7. CLONING OF THE NEURAMINIDASE CONTAINED IN pJCP301 FROM THE LambdaEMBL301 PNEUMOCOCCAL GENOMIC LIBRARY

The neuraminidase cloned by Berry *et al* 1988 suffered a spontaneous deletion reflecting the likely instability of pneumococcal DNA cloned in cosmid vectors. This has been observed in some other studies [Pratt & Boulnois 1987]. However, no stability problems were encountered in the cloning of the pneumococcal IgA1 protease and neuraminidase when using a LambdaEMBL301 pneumococcal genomic library (this thesis). The availability of this genomic library led to attempts to clone the intact sequence responsible for the production of neuraminidase in pJCP301. An internal DNA fragment of this recombinant was used as a probe. This would allow a better comparison between this neuraminidase and the one encoded by pMC2150.

The amplified LambdaEMBL301 pneumococcal genomic library was plated on *E. coli* LE392 and 2,688 amplified recombinants were screened. Recombinant plaques were transferred by plaque lift onto Dupont nitrocellulose filters. As a negative control LambdaEMBL301 was used. The probe used was the 1.6 Kb *Pvu*II fragment from pJCP301. A final filter wash of 0.2 X SSC was used. Two hybridizing recombinants were found and termed LambdaEMBL301-Neu2 and LambdaEMBL301-Neu3. These recombinants were grown overnight in broth to amplify the number of

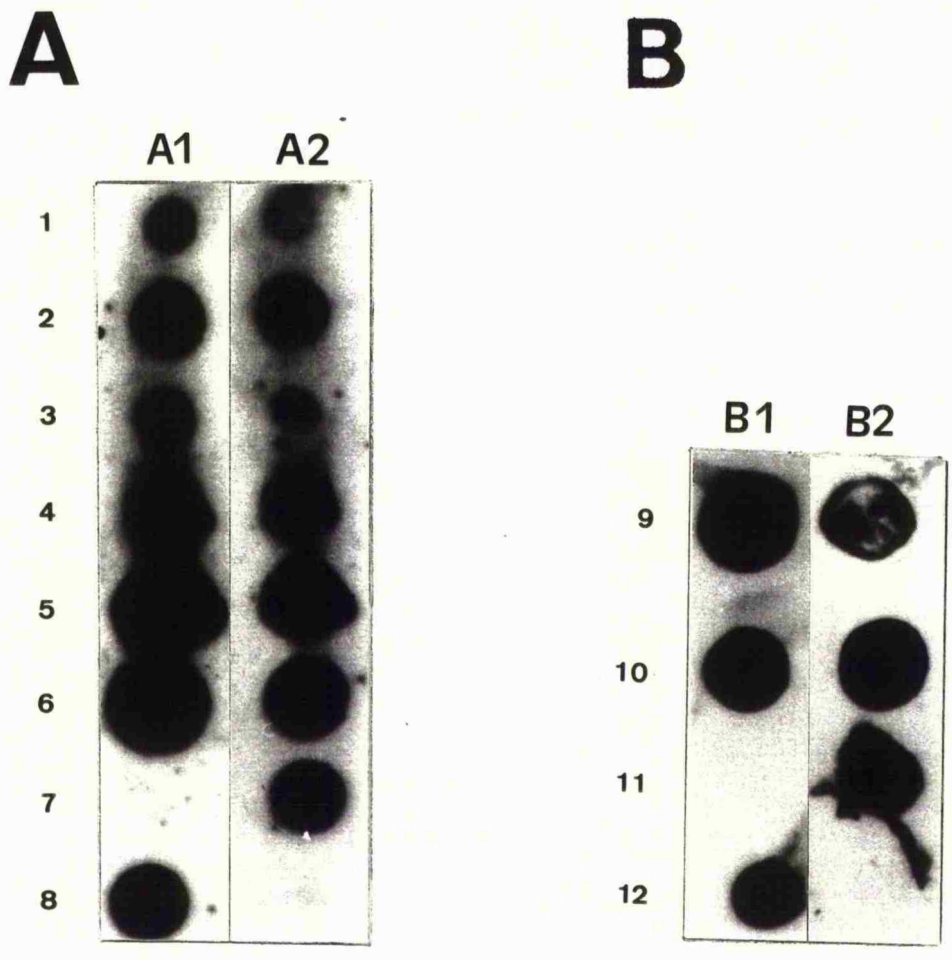


FIGURE 4.11. **Detection of neuraminidase genes in pneumococcal strains.**

A. Chromosomal DNA from the pneumococcal strains SP01 (1), SP02 (2), 3551 (3), GB05 (4), GB05A (5), GB05B (6) were hybridized together with the 1.6 Kb *Pvu*II fragment from pJCP301 (7) and the 1.7 Kb *Hind*III fragment from pMC2150 (8). The blots were probed with either the 1.7 Kb *Hind*III fragment from pMC2150 (A1) or the 1.6 Kb *Pvu*II fragment from pJCP301 (A2).

B. Chromosomal DNA from the pneumococcal strains 11733 (9) and 12213 (10) were blotted together with the 1.6 Kb *Pvu*II fragment from pJCP301 (11) and 1.7 Kb *Hind*III fragment from pMC2150 (12). The blots were probed with either the 1.7 Kb *Hind*III internal fragment from pMC2150 (B1) or the 1.6 Kb *Pvu*II fragment from pJCP301 (B2). Non-radioactive probes were used in this blot.

phages. The supernatants from phage cultures had no neuraminidase activity. Nevertheless the recombinants were subjected to further purification to investigate the possibility of phage contamination with other recombinats.

LambdaEMBL301-Neu2 and -Neu3 were re-plated and re-probed with the 1.6 Kb *Pvu*II fragment from pJCP301. Many positives were obtained and two positive clones of each recombinant were amplified. No detectable neuraminidase activity was found in the phage supernatants of this second amplification.

LambdaEMBL301-Neu2 and LambdaEMBL301-Neu3 were plated and the plaques obtained were analyzed for the presence of the 1.6 Kb *Pvu*II fragment from pJCP301. Also, to confirm that these recombinants contained inserts distinct from that in pMC2150, a 1.7 Kb *Hind*III fragment from that plasmid was used as a probe. This was achieved by making replica filters from each plate to be screened of the filters obtained from each plate, one was hybridized with the 1.6 Kb *Pvu*II fragment from pJCP301, and the other with the 1.7 Kb *Hind*III fragment from pMC2150. LambdaEMBL301 was used as a negative control. A stringency of 0.2 X SSC was used for the final wash of the nitrocellulose filters. No hybridization was found when using the probe from pMC2150 whereas hybridization was detected with the pJCP301 probe.

From this experiment 12 positively-hybridizing plaques of LambdaEMBL301-Neu2 and -Neu3 were amplified overnight, but no neuraminidase activity was found in any supernatant. At that stage it was unknown whether or not this was due to a lack, in the lambda recombinant clones, of either the whole neuraminidase gene sequence, or the flanking sequences maybe required for the expression of the enzyme. To check that the sequences contained in the two lambda clones corresponded to that carried by pJCP301, these recombinants were purified.

LambdaEMBL301-Neu2 and LambdaEMBL301-Neu3 DNA were digested with *Pvu*II, *Pst*I, *Kpn*I or *Not*I. In all of these digests identical bands were

obtained for both recombinants. It was concluded that they were identical. To assess if the 1.6 Kb *Pvu*II fragment from pJCP301 was present in LambdaEMBL301-Neu2, these two recombinants, together with LambdaEMBL301 vector, as a control, were subject to *Pvu*II digestion. The samples resulting from these digests were analyzed by electrophoresis through 1% agarose gel [FIGURE 4.12]. The 1.6 Kb *Pvu*II fragment, present in both LambdaEMBL301-Neu2 and pJCP301, was absent in LambdaEMBL301. This result, although not conclusive, suggests that LambdaEMBL301-Neu2 contains the DNA sequences present in pJCP301, and should encode for the neuraminidase activity detected by Berry *et al* (1988).

4.8. CONCLUSION

A recombinant coding for neuramindiase activity has been isolated from a LambdaEMBL301 pneumococcal genomic library. DNA restriction mapping and hybridization studies have revealed that the sequences responsible for the expression of this enzyme activity are distinct from the neuraminidase recombinant isolated by Berry *et al* (1988). Also the DNA sequences of both recombinants have been detected in all the pneumococcal strains tested.

Polyclonal antibodies that react with neuraminidase purified from the pneumococcus, reacts with the products of both pJCP301 and pMC2150, although the reason for the cross-reaction is still unclear.

Finally, the recombinant DNA sequences from the neuraminidase-producing clone pJCP301 have been isolated from a LambdaEMBL301 pneumococcal genomic library. No neuraminidase activity was detected in the new recombinant clone. The reason for this still remains unclear.

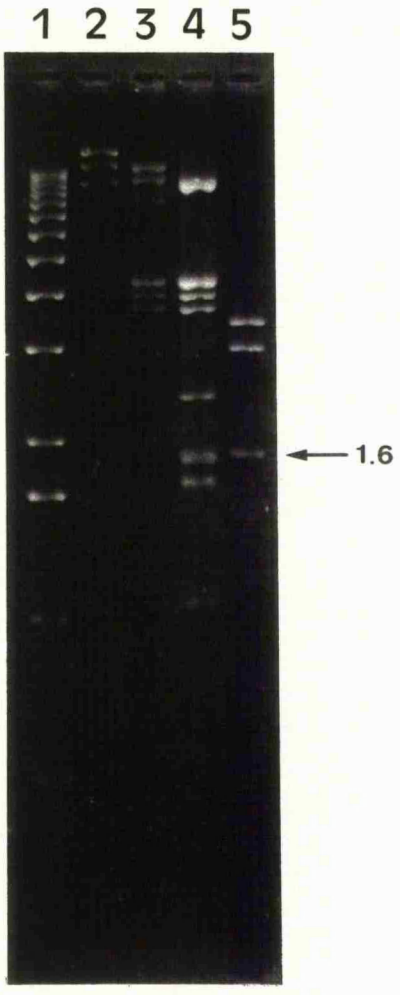


FIGURE 4.12. **Restriction enzyme digests of LambdaEMBL301-Neu2.**
LambdaEMBL301-Neu2 was digested with *PvuII* (4). To discern between fragments corresponding to vector and insert, LambdaEMBL301 vector was digested with the same enzyme (3). Also to locate the position of the 1.6 Kb *PvuII* fragment from pJCP301, this plasmid was digested with the same enzyme (5). The 1.6 Kb fragment is indicated. DNA size markers are 1Kb ladder (1) and wild type Lambda DNA digested with *Bgl*II (top fragment sizes 22, 13.2 and 9.6 Kb)(2).

PART V

DNA SEQUENCE DETERMINATION

OF THE

NEURAMINIDASE GENE

IN LambdaEMBL301-Neu1 CLONE

5.1. INTRODUCTION.

After the cloning of the pneumococcal neuraminidase in LambdaEMBL301, further studies on the gene and its product were carried out in an attempt to understand the structure and function of this enzyme. The first step in this study was the determination of the nucleotide sequence of the neuraminidase gene. This sequence information could be used to determine the predicted protein sequence which would be the starting point for the study of this enzyme. Hydrophobicity and secondary structure plots would reveal some information about the structure of this protein. The amino acid sequence can be used to search protein databases to determine any homologies to other proteins for which data are already available. By this comparison some hints on the neuraminidase structure could be found. Also the search for certain amino acid sequence motifs present in other proteins could provide more specific information about the structure and function of this enzyme.

Homology studies of this neuraminidase with other microbial neuraminidases could reveal the presence of common sequences which must be important in enzyme activity. These sequences could be used in the construction of pneumococcal insertion-duplication mutants, allowing the study of this enzyme in the pneumococcus. Also the use of these mutants in animal models, might ultimately provide useful information about the role of this enzyme in pneumococcal disease.

For the above reasons determination of the DNA sequence of the cloned pneumococcal neuraminidase gene in LambdaEMBL301 was the first aim of this study.

5.2. DNA SEQUENCE DETERMINATION.

For DNA sequence determination, DNA fragments from pMC2150 were subcloned into the phages M13mp18 and M13mp19. These phages, following infection of *E. coli*, go through two different phases during its life cycle. The first is the replicative phase that allows the phage to multiply. During this phase the phage DNA is in double stranded form.

Once enough DNA has been produced, the phage enters into the next phase in which single stranded DNA (sDNA) is produced and encapsidated and the phage released from the *E. coli* host cell [Wickner 1978]. DNA from the phage is easily purified.

To sequence both DNA strands, M13mp18 and M13mp19 phages were used [Yanish-Perron *et al* 1985]. These vectors possess the *lacZ* multiple cloning site from pUC18 and pUC19 vectors, respectively, which allows cloning of DNA fragments in two orientations. In this way, opposite strands of a DNA fragment can be sequenced.

For the sequencing of the neuraminidase gene contained in pMC2150 the three *Hind*III fragments of 1.29, 1.70 and 1.16 Kb length [FIGURE 5.1] were cloned into M13mp18 and mp19. For this purpose, M13mp18 and mp19 were digested with *Hind*III, followed by alkaline phosphatase treatment, to avoid self ligation. pMC2150 was also digested with the *Hind*III. Once the digestion was completed, without any further treatment, the DNA fragments were ligated into the M13mp18 and mp19 phosphatased vectors. The ligated DNA was transformed into *E. coli* JM101, since this strain carries the pili that are required for the phage infection. Recombinant phages were selected by lack of β-galactosidase production and replicative form (RF) DNA was prepared and analyzed by digestion with different restriction enzymes.

DNA templates for sequencing with the recombinant phages were prepared. DNA sequencing of the three *Hind*III fragments was carried out following the dideoxy method [Sanger *et al* 1977], in which ^{35}SdATP was incorporated. Synthetic oligonucleotides were used, as sequence information was obtained, to determine the complete DNA sequence. TABLE 5.1 shows the list of oligonucleotides used as primers. The products of sequencing reactions were resolved by electrophoresis in polyacrylamide gradient gels. The gels were fixed, dryed and exposed to X-ray film. DNA sequence information was taken by reading the developed films.

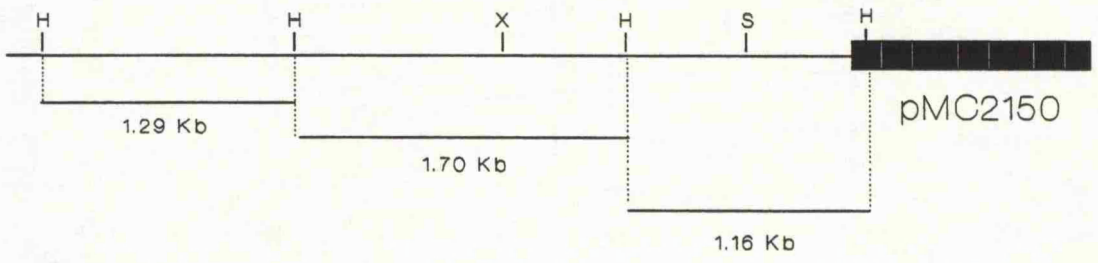


FIGURE 5.1. **Representation of the three *Hind*III fragments from pMC1250 that were cloned in M13mp18/19.**
Restriction sites are as follows: H, *Hind*III; X, *Xba*I and S, *Sph*I.

Once the sequence of the three *Hind*III fragments was determined, junctions between the three fragments were sequenced by double stranded sequencing of the plasmid pMC2150 with the primers number 8 and 24 [TABLE 5.1]. These primers annealed immediately before the *Hind*III sites so that the extension during the sequencing reactions runs through these restriction sites.

The sequence strategy for pMC2150 is shown in FIGURE 5.2. This figure shows the length of the DNA sequenced from each oligonucleotide. The numbers correspond to the primers shown in TABLE 5.1.

NUMBER		SEQUENCE (5'→3')
1	UNIVERSAL	AAACGACGGCCAGT
2	REVERSE	CAGGAAACAGCTATGAC
3		GCCATCCTACCATGAA
4		TCGTCAGGCTTGTCTGT
5		TATAAAGAGATTTTAG
6		GAGCTCAACCCTAACTG
7		ACAGCAGAACTACCTAA
8		AGGAACTTTGATCGCAG
9		GATATTACTCCGATG
10		TCATCCTCAGTAATGC
11		TACAGTGGATTCAGAGGA
12		CGCTTGTAACTCTTACTA
13		ACTCGAAACAAGTATTGTA
14		ATCTATAAGGAAGGTAGCT
15		CTTCACAAGGAGGTCTGA
16		TTCTAATTTGGAATTTAGT
17		AGCTAGGAGGTCACTCTC
18		CAGCAGCTTCACTTCCGT
19		CGTAGAAGAGTGGATC
20		CAGGATCTACAACAACGC
21		CTTGGTTGCTCCGATT
22		TCAACTCTCTGACCAT
23		ATGCACTTAGTTTGTC
24		TCAGAATCTTATCCAGCT
25		ATACCTGATATTTAC
26		TGACAAATCTAGCAAC

TABLE 5.1. List of synthetic oligonucleotides used in the sequencing of pMC2150 and pMC4170.

The DNA sequence from pMC2150 was analyzed for the presence of open reading frames (ORFs). This analysis was done allowing ATG and GUG as translation start codons. The results are shown in FIGURE 5.2. Several ORFs were found. The largest one (ORF1) was thought to be the one coding for neuraminidase activity. The main reason that led to this conclusion was the presence in the predicted amino acid sequence of the "aspartic box" motif found in other neuraminidases already sequenced (see introduction). These motifs will be described in more detail later in this section.

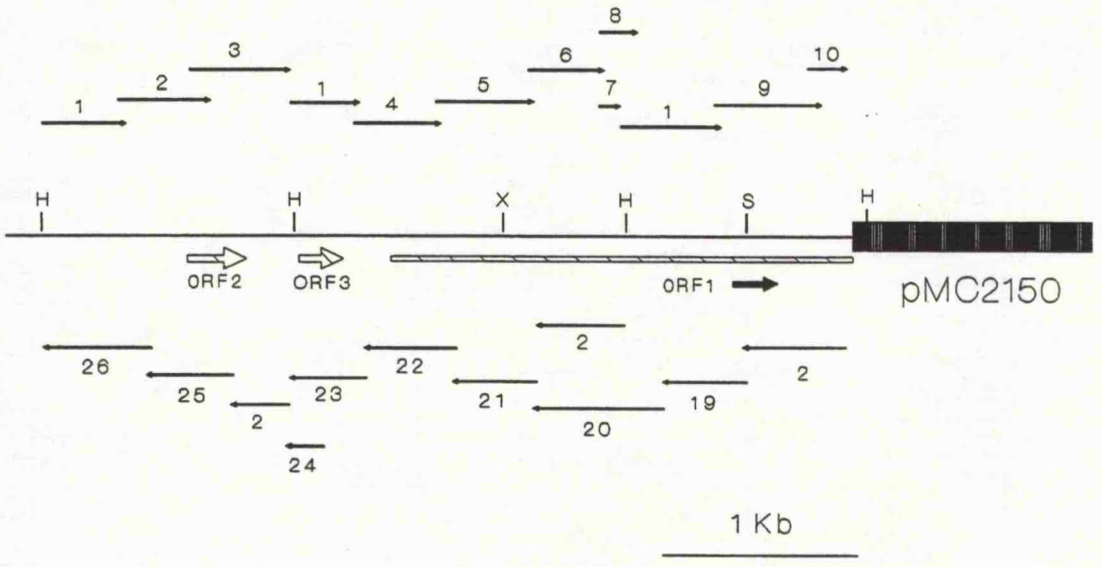


FIGURE 5.2. Sequence strategy of pMC2150.

Arrows indicates the fragment sequenced from every oligonucleotide as well as the direction of sequencing. Numbers over the arrows indicate different oligonucleotides used and correspond to the ones in TABLE 5.1. Restriction sites are as follows H, *Hind*III; X, *XbaI*; S, *SphI*. Large arrows indicate ORFs.

Since no translation stop codon was found in ORF1 and the coding sequence ran into the vector, a larger fragment containing the whole ORF1 from lambda EMBL301-Neu1 had to be cloned to obtain the C-terminal sequence of neuraminidase. Recombinants of the N4 (7-10 Kb insert) size fractionated library from the lambdaEMBL301-Neu1 subcloning in pJDC9, previously described in section 4.3., were analyzed. Four of these clones were grown in *E. coli* JM101 and produced large amounts of neuraminidase. DNAs from these four recombinants (pMC410, pMC470, pMC4170, pMC4210) were digested with *Eco*RI and *Hind*III. The agarose gels with the digested products are shown in FIGURE 5.3. In these digests pMC410, pMC470 and pMC4170 gave identical bands in both digestions, with estimated fragments of 9.9, 2.45 and 1.9 Kb for the *Eco*RI digests,

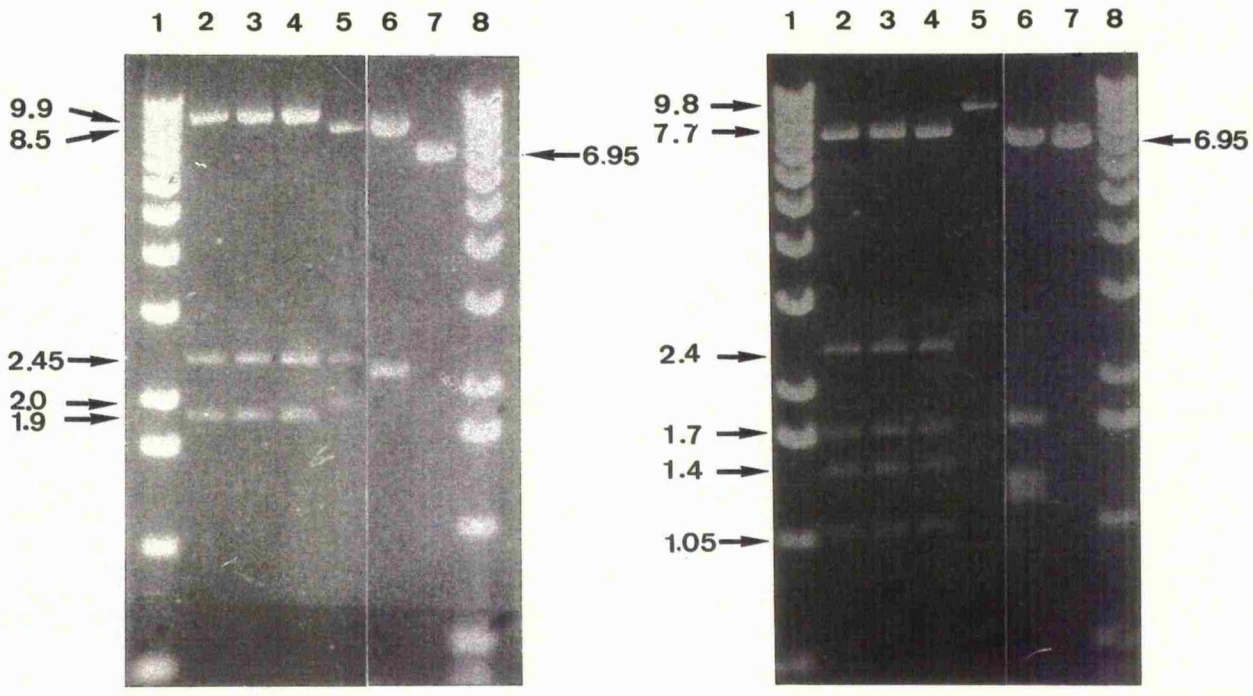


FIGURE 5.3. **Restriction digests of DNA from LambdaEMBL301-Neu1 subclones.**

A. Restriction digest using *EcoR*I of the following recombinants: (2) pMC410; (3) pMC470; (4) pMC4170; (5) pMC4210. As a control pMC2150 (6) and the vector pJDC9 (7) were digested with the same enzyme. (1) and (8) correspond to the 1Kb ladder DNA size marker.

B. Restriction digest using *Hind*III of the following recombinants: (2) pMC410; (3) pMC470; (4) pMC4170; (5) pMC4210. As a control pMC2150 (6) and the vector pJDC9 (7) were digested with the same enzyme. (1) and (8) correspond to the 1Kb ladder DNA size marker.

and 7.7, 2.4, 1.7, 1.4, and 1.05 Kb for the *Hind*III digests. Since the vector (pJDC9) is 6.95 Kb, the estimated insert size for these three recombinants was 7.3 Kb. In the case of the *Eco*RI digest from pMC4210 the resulting DNA fragments were 8.5, 2.45, and 2, and for the *Hind*III digest 9.8, 1.7, and 1.4 Kb, giving an estimated insert size of 6 Kb. pMC4210 was excluded as it was the only one showing a different DNA restriction pattern to the other three in agarose gel electrophoresis, and also had a smaller sized insert. Since the other three recombinants seemed to be identical, pMC4170 was selected for further analysis. This recombinant was subject to restriction enzyme mapping digests. No sites were found for *Kpn*I, *Sma*I and *Sal*I. The restriction map of this recombinant is shown in FIGURE 5.4.

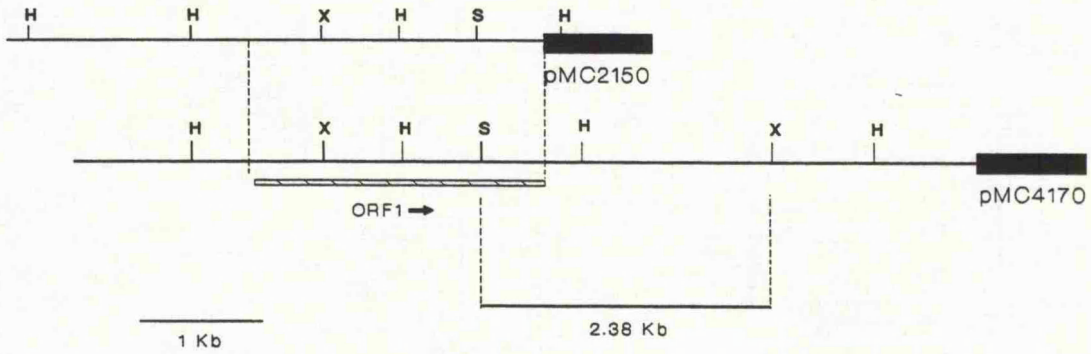


FIGURE 5.4. Restriction map of pMC4170 compared in scale to that for pMC2150.
The 2.38 Kb *Sph*I/*Xba*I fragment used for sequencing is indicated. The incomplete ORF1 sequence from pMC2150 is represented. Restriction sites are as follows: H, *Hind*III; S, *Sph*I and X, *Xba*I.

A 2.38 Kb fragment resulting from the digestion of pMC4170 with *Sph*I and *Xba*I was used to obtain the C-terminal end of the sequence from ORF1. The reason for choosing this fragment is that its first 0.5 Kb overlap with the end of the pMC2150 insert, which also corresponds to the

end of the incomplete ORF1 [FIGURE 5.4]. It was expected that the translation STOP codon of ORF1 would fall within the 1.88 Kb remaining of the 2.38 Kb fragment.

The subcloning of the fragments previously mentioned was achieved as follows. DNA from pMC4170 was digested with *Sph*I and *Xba*I. A 2.38 Kb fragment resulting from that digestion was electroluted from a 1 % agarose gel. This DNA was ligated to M13mp19 and M13mp18 previously digested with *Sph*I and *Xba*I. Recombinant clones were analyzed by digestion with the enzymes mentioned above. Once the recombinants carrying the 2.38 Kb fragment were obtained, DNA sequencing was carried out by using the appropriate oligonucleotides. The first primer used was number 10 [TABLE 5.1]. This oligonucleotide confirmed the overlap between ORF1 from pMC2150 and pMC4170. The rest of the oligonucleotides used for the sequence of the 2.38 Kb fragment are also shown TABLE 5.1. The sequence strategy adopted for this fragment is shown in FIGURE 5.5.

When the sequence was completed, as verified by the presence of a translation STOP codon (TAA) in ORF1, the next step was to verify that the cloned fragments were contiguous in the pneumococcal genome, since the genomic library used for the cloning of the neuraminidase gene was made with chromosomal DNA that had not been size fractionated after the *Sau3A*I partial digestion, and before its ligation to Lambda EMBL301. Therefore distant DNA fragments in the chromosome could have been ligated into the same recombinant. To exclude this possibility the sequenced DNA was examined by Southern blot analysis.

For the Southern blot, chromosomal DNA from *Streptococcus pneumoniae* R36A was prepared along with pMC2150 and pMC4170. These DNA samples were subject to single digests with *Eco*RI, *Hind*III or *Pvu*II, and double digests with *Hind*III and *Eco*RI. The products of these digests were subject to electrophoresis in a 0.7% agarose gel. DNA from this gel was blotted into a nitrocellulose filter followed by hybridization with a ^{32}P-

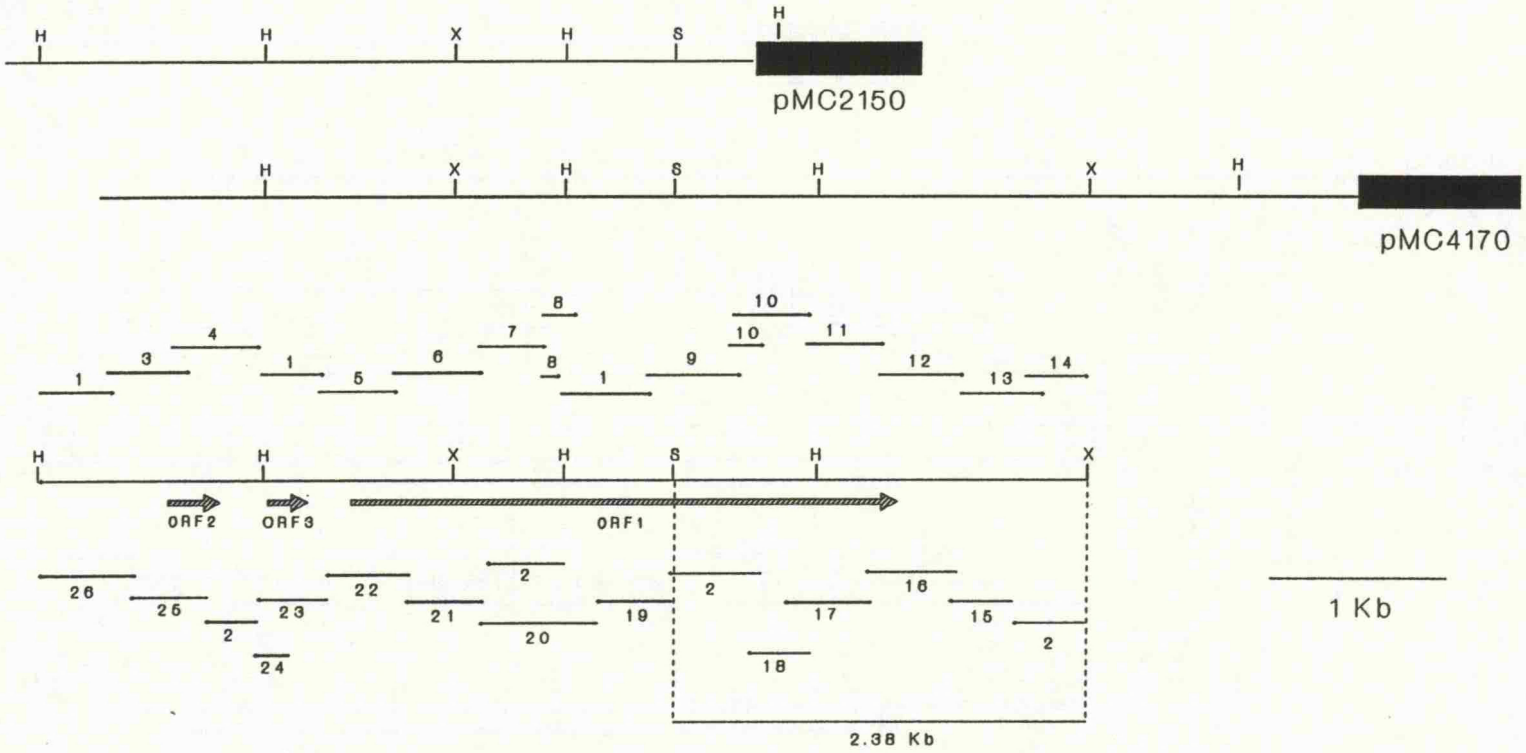


FIGURE 5.5. **Sequence strategy for pMC2150 and the 2.35 Kb *Sph*I/*Xba*I fragment from pMC4170.**

Small arrows indicate segments of DNA sequenced in both directions. Numbers above arrows corresponds to the oligonucleotide used for each segment and are specified in TABLE 5.1. Large arrows indicate ORFs. Restriction sites are as follows: H,*Hind*III; S, *Sph*I and X, *Xba*I.

labelled *Pvu*II fragment (3.54 Kb) from pMC4170 [FIGURE 5.6 A]. The filter was washed in 0.1 x SSC. The resulting blot is shown in FIGURE 5.7.

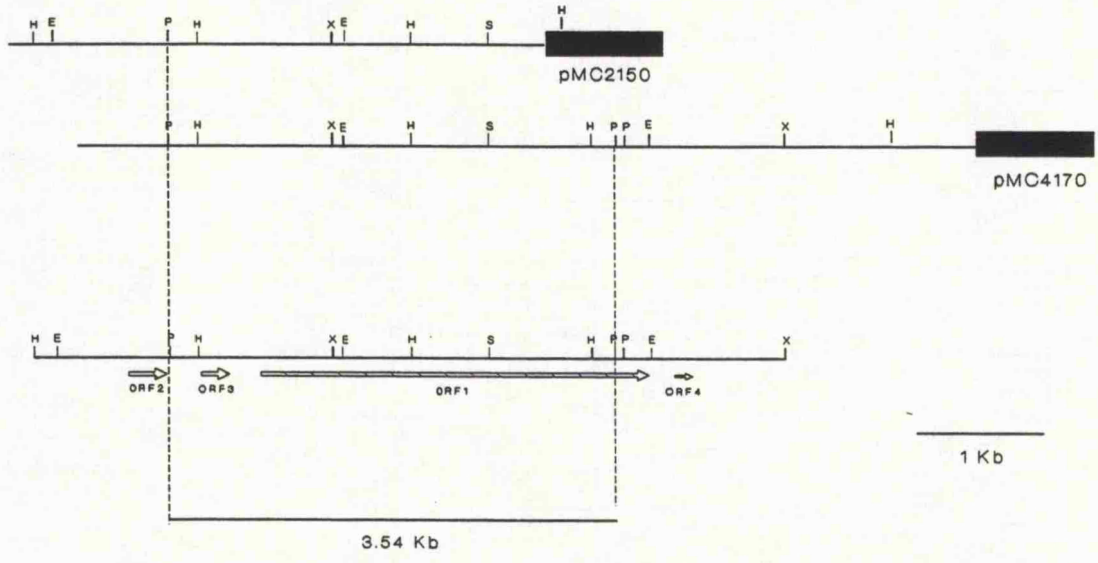


FIGURE 5.6.A. **Representation of the 3.54 Kb *Pvu*II fragment from pMC4170 indicating the region covered by it in both pMC4170 and pMC2150.**
The restriction map at the bottom of the figure corresponds to the total amount of sequenced DNA from pMC2150 and pMC4170. ORFs are indicated by arrows. Restriction sites are as follows: E, *Eco*RI; H, *Hind*III; P, *Pvu*II; S, *Sph*I and X, *Xba*I.

The first digestion shown was with *EcoR*I. The expected fragments from these digestion are indicated in FIGURE 5.6.B. Bands of 2.25 and 2.45 Kb were expected in the track corresponding to the chromosomal DNA [FIGURE 5.7] but only a single band of 2.45 Kb appeared. This band was coincident with the 2.45 Kb shown in the pMC4170 track. The presence in pMC2150 of the 2.25 Kb band absent from the chromosomal DNA digestion, suggested the possibility of a lack of contiguous DNA in this cloned fragment. Also, the appearance of a single band in the chromosomal DNA suggests that the 2.45 Kb band could be a doublet. Analysis of the DNA sequence revealed the presence of a single *Sau3A*I site in the *EcoR*I 2.25 Kb fragment [FIGURE 5.6.B] which indicates that the upstream sequence from this *Sau3A*I site arose from a non-contiguous fragment in the chromosome during the cloning.

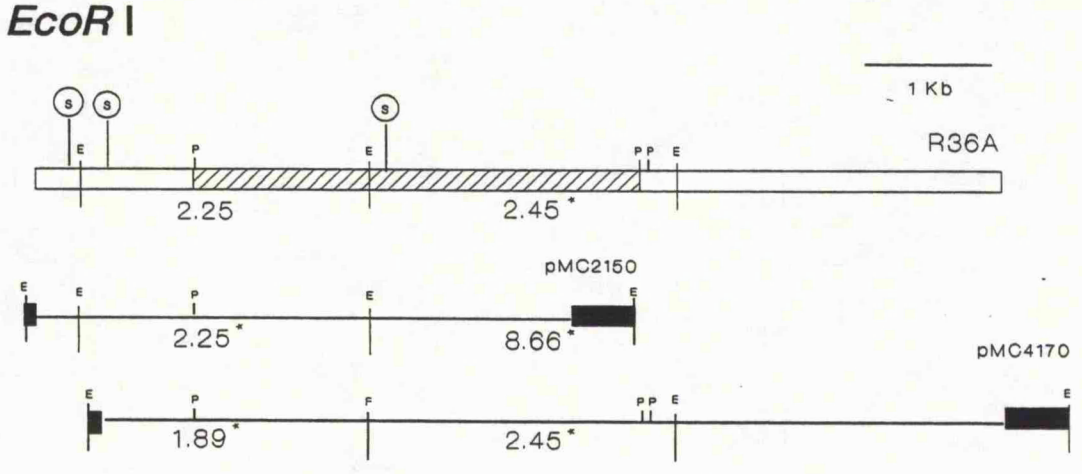


FIGURE 5.6.B. Expected fragments from the *EcoR*I digests.

The region covered by the 3.54 Kb *Pvu*II probe from pMC4170 is indicated in shadow in the chromosomal DNA. Bands present in the Southern blot are indicated by (*) over the corresponding size fragment. Restriction endonuclease sites are as follows: E, *EcoR*I; P, *Pvu*II; S, *Sau3A*I.

In the case of the *Hind*III digestion [FIGURE 5.6.C], the 1.7 Kb fragment from pMC2150 and pMC4170 was present in the chromosomal DNA. The 1.43 and 2.4 Kb fragments from pMC4170 were also present in the chromosome. In contrast, the 1.29 Kb fragment from pMC2150 was absent in the chromosomal DNA, supporting the idea of the lack of contiguity for this cloned DNA fragment.

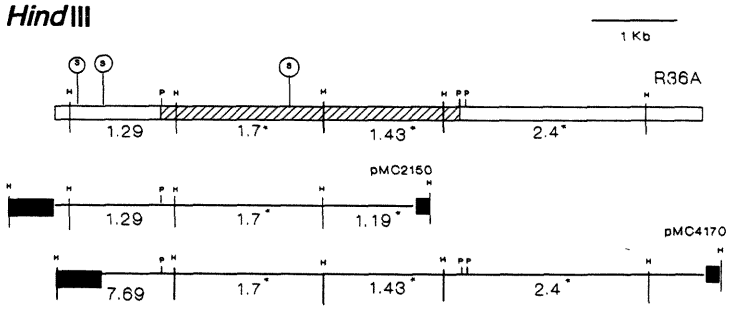


FIGURE 5.6.C. Expected fragments from the *Hind*III digests.

The region covered by the 3.54 Kb *Pvu*II probe from pMC4170 is indicated in shadow in the chromosomal DNA. Bands present in the Southern blot are indicated by (*) over the corresponding size fragment. Restriction endonuclease sites are as follows: H, *Hind*III; P, *Pvu*II; S, *Sau3A*I.

The *Hind*III and *Eco*RI double digest corroborated the contiguous nature of the DNA fragment between the second *Hind*III restriction site (upstream to ORF1 [FIGURE 5.6.A]) and the last *Eco*RI sites from the cloned DNA. This was shown by the presence of the 1.15 Kb, 0.55 Kb, 1.43 Kb and 0.47 Kb fragments, resulting from the *Hind*III and *Eco*RI double digest [FIGURE 5.6.D], in the Southern blot hybridization [FIGURE 5.7.]

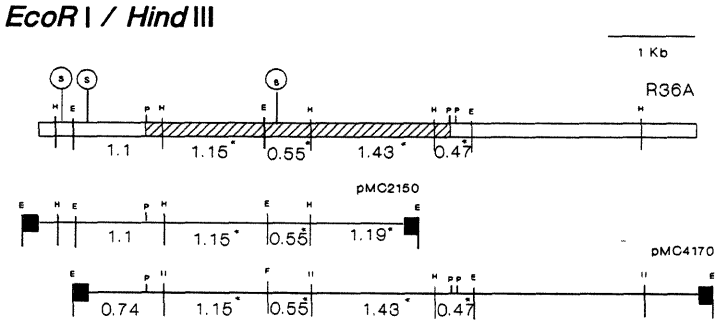


FIGURE 5.6.D. Expected fragments from the *EcoRI/HindIII* double digests.
The region covered by the 3.54 Kb *Pvu*II probe from pMC4170 is indicated in shadow in the chromosomal DNA. Bands present in the Southern blot are indicated by (*) over the corresponding fragment. Restriction endonuclease sites are as follows: E, *EcoR*I; H, *Hind*III; P, *Pvu*II; S, *Sau3A*I.

Finally, digestion with *Pvu*II showed the expected bands from FIGURE 5.6.E. in both chromosomal and plasmids DNAs [FIGURE 5.7], which means that the 3.54 Kb *Pvu*II fragment corresponds to a contiguous sequence in the chromosome.

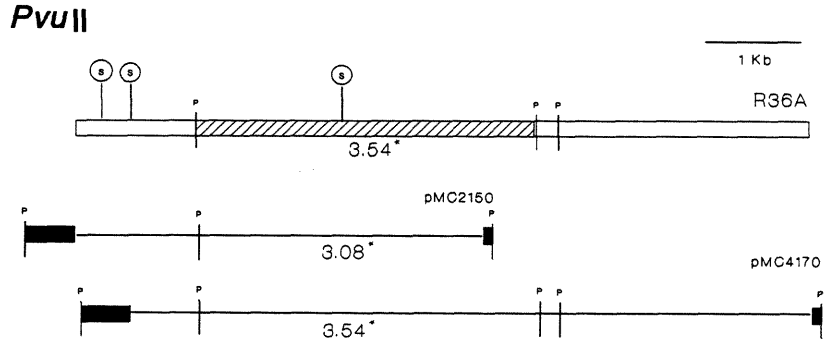


FIGURE 5.6.E. Expected fragments from the *Pvu*II digests.
The region covered by the 3.54 Kb *Pvu*II probe from pMC4170 is indicated in shadow in the chromosomal DNA. Bands present in the Southern blot are indicated by (*) over the corresponding size fragment. Restriction endonuclease sites are as follows: P, *Pvu*II; S, *Sau3A*I.

From this Southern blot analysis it can be concluded that the 5.43 Kb fragment, starting from the second *Sau3A*I site to the end of the sequence corresponds to a contiguous DNA fragment in the chromosome of *Streptococcus pneumoniae* R36A. This applies to the ORF1 corresponding to the neuraminidase gene, plus flanking sequences.

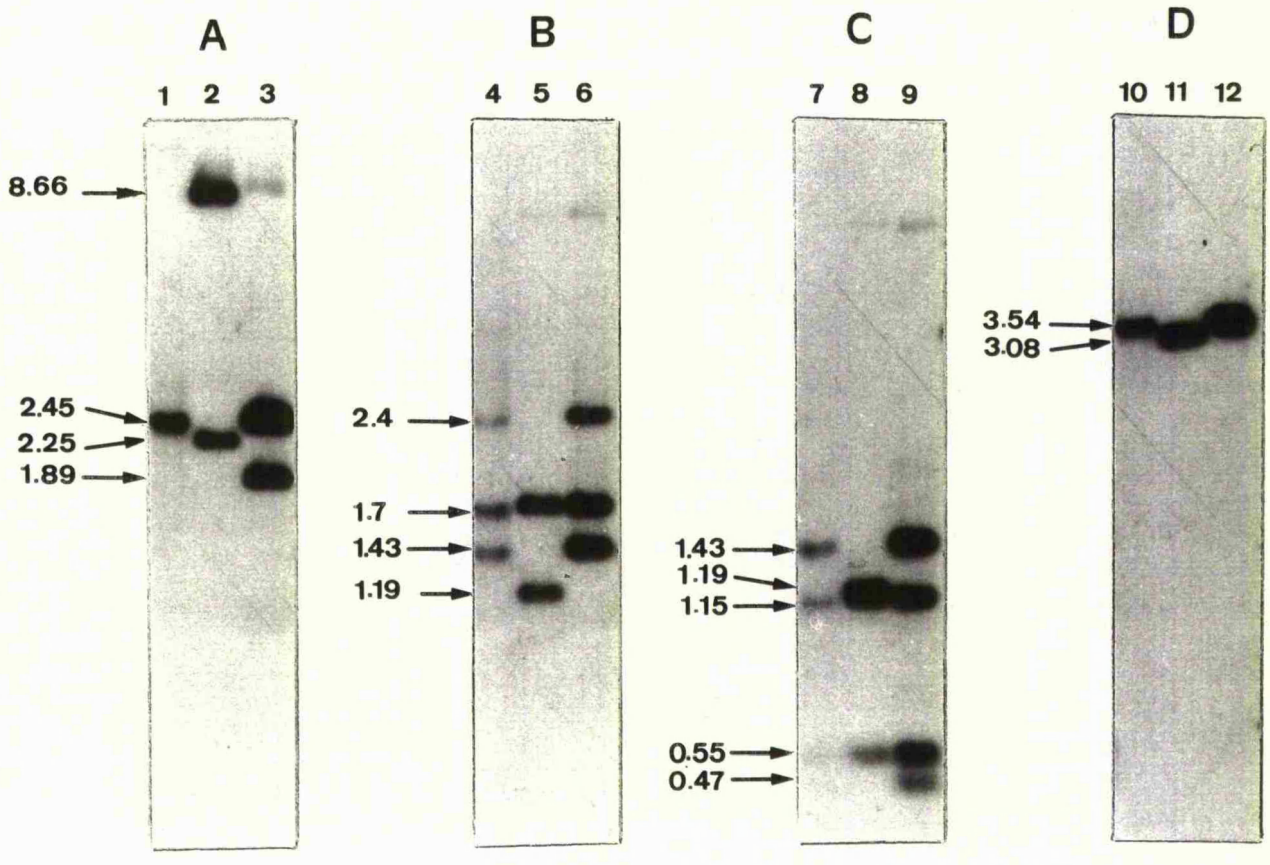


FIGURE 5.7. **Southern blot hybridization to determine contiguity of sequenced DNA.** Chromosomal DNA from the pneumococcal strain R36A (1)(4)(7)(10), together with the recombinants pMC2150 (2)(5)(8)(11) and pMC4170 (3)(6)(9)(12) were subject to single digestion with *EcoR*I (A), *Hind*III (B), *Pvu*II (D), and double digest with *Hind*III/*EcoR*I (C). Digested DNAs were resolved on a 0.7% agarose gel. DNA was blotted onto nitrocellulose membranes that were hybridized with ^{32}P-labelled 3.54Kb fragment resulting from digestion of pMC4170 with *Pvu*II. Membranes were washed in 0.1 X SSC. Numbers with arrows indicate kilobases.

5.3.DNA SEQUENCE ANALYSIS.

Once the DNA sequence had been determined, the next step was to analyze it for the presence of ORFs. These ORFs are regions of DNA that potentially encode the information for the synthesis of a protein. The following features were used to define an ORF:

a/ Presence of a translation start codon. The starting amino acid in the synthesis of virtually all bacterial polypeptides is N-formylmethionine. There are two codons (AUG and GUG) that code for this amino acid [Kozak 1983], to which $tRNA_F^{Met}$ binds for the initiation of protein synthesis.

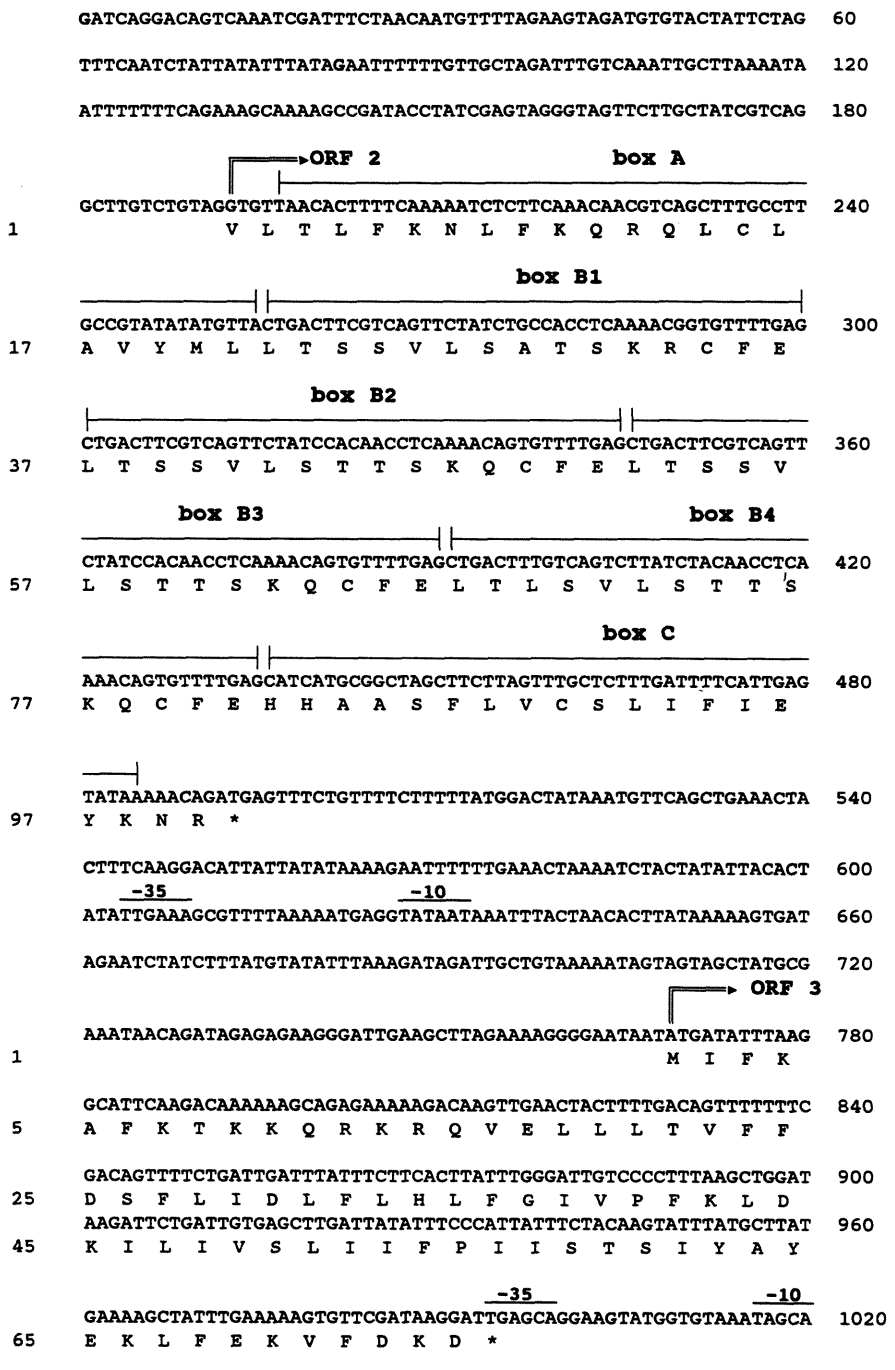
b/ Presence of translation stop codons. The three codons UAA, UAG, and UGA do not correspond to any amino acid. Is at this point where translation terminates [Kohli & Grosjean 1981].

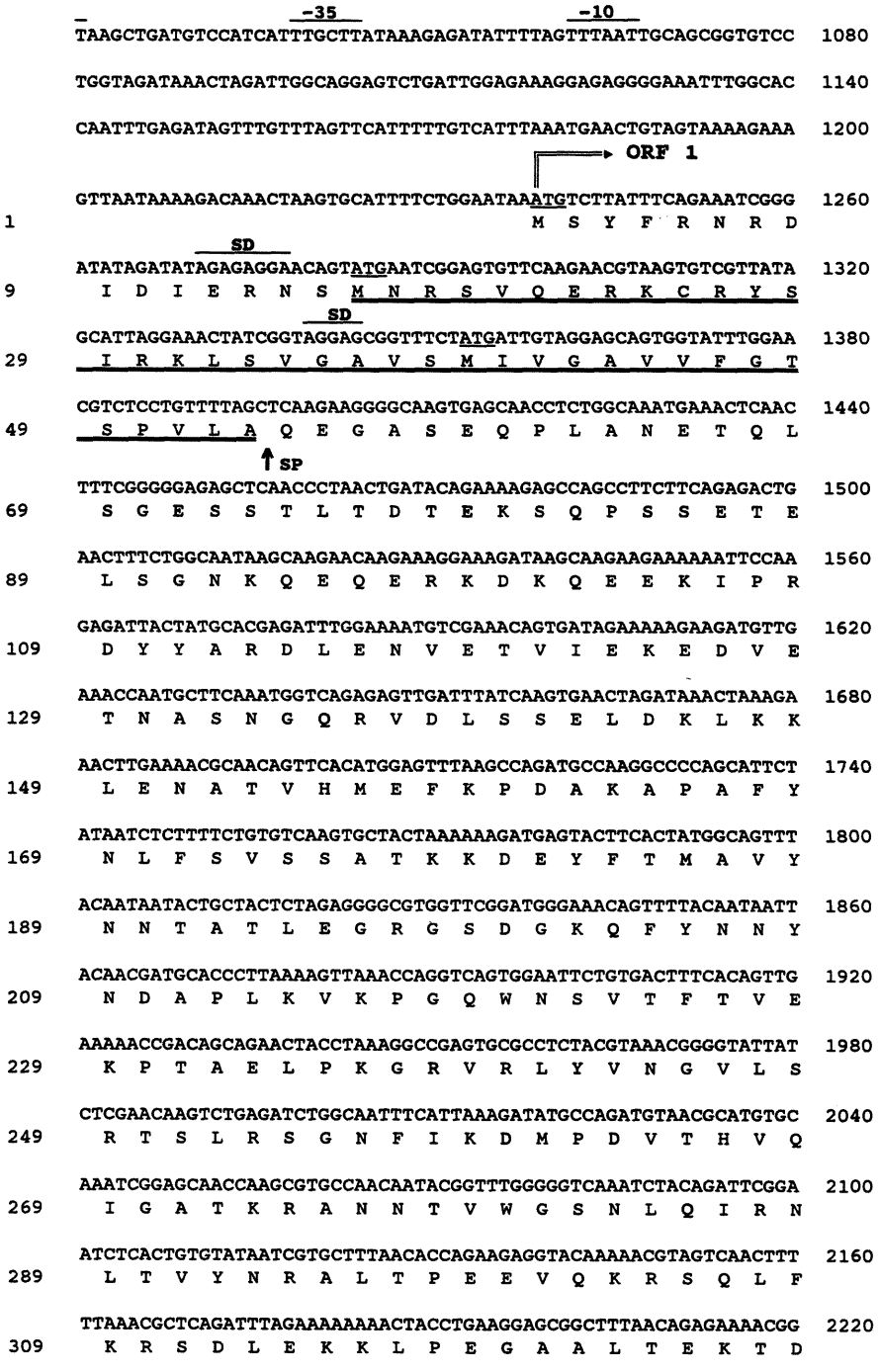
c/ Presence of a ribosome binding site (Shine-Dalgarno sequence). Almost all ribosome binding sites have a sequence such as AGGA or GAGG centered 8 to 13 nucleotides upstream from the start codon. This region is recognized by the 30S ribosome subunit before protein synthesis starts [Shine & Dalgarno 1974].

d/ Presence of a promotor sequence, where the σ subunit of the RNA polymerase specifically binds for the initiation of the RNA synthesis. Two consensus sequences for this promoter have been found at -35 (T_{82} T_{84} G_{78} A_{65} C_{54} A_{45}) and -10 (T_{80} A_{95} T_{45} A_{60} A_{50} T_{96}) basepairs from the translation start codon [Rodriguez & Chamberling 1982].

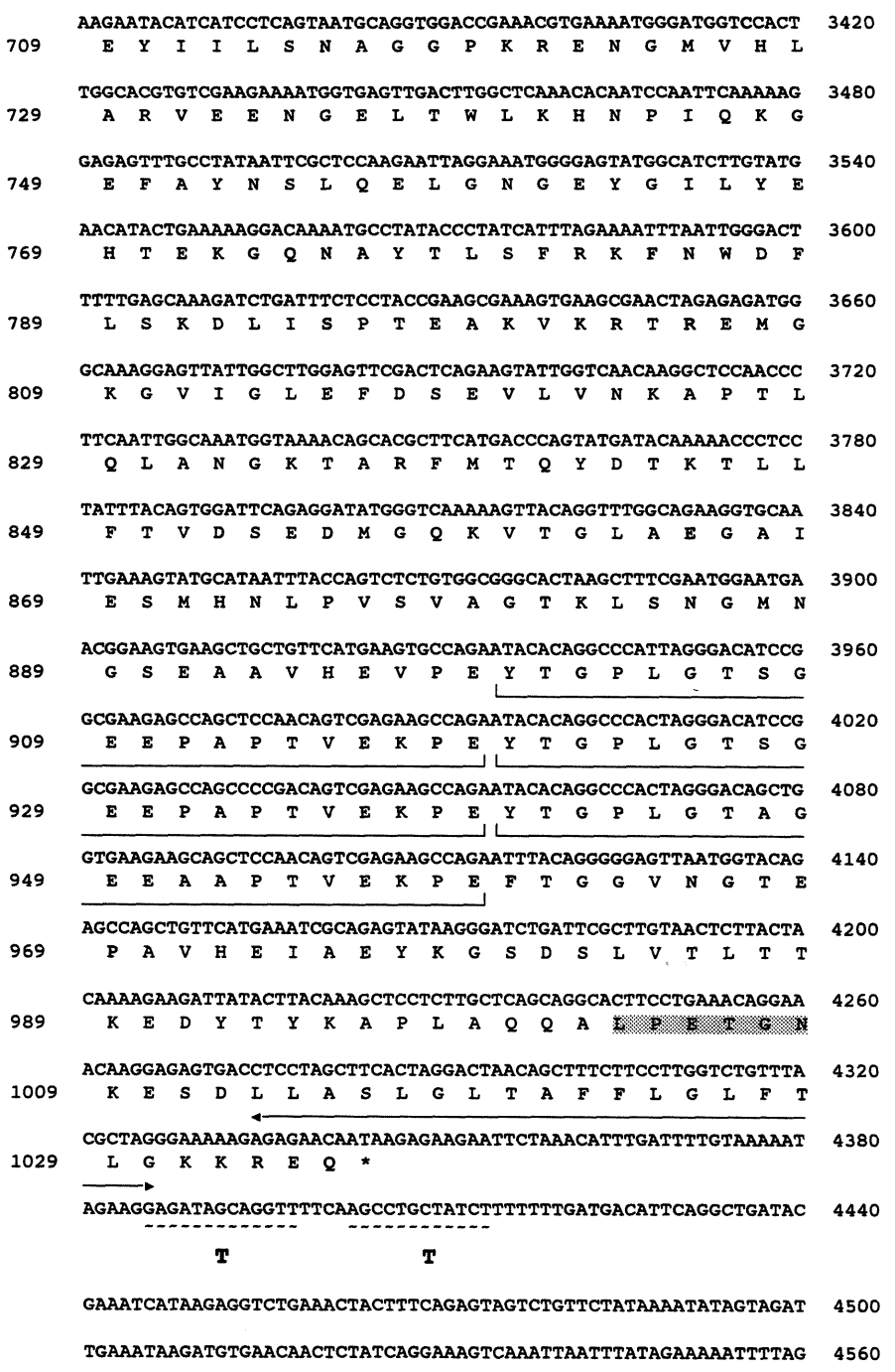
e/ Presence of a transcription terminator sequence. At this point, the RNA polymerase stops polymerizing nucleotides into RNA, releases the RNA chain, and leaves the DNA. At some sites, termination requires an accessory protein, *rho*-factor (*rho*-dependent terminators), whereas at other sites, the core enzyme alone carries out all the steps (*rho*-independent terminators) [Roberts 1969].

Computer analysis, using BIOSEQ in the Leicester University VAX/VMS Cluster, revealed the presence of four ORFs (ORF1, ORF2, ORF3 and ORF4) [FIGURE 5.8] in the 5440 bp DNA fragment from pMC2150 and pMC4170.


GATCAGGACAGTCAAATCGATTTCTAACAATGTTTTAGAAGTAGATGTGTACTATTCTAG 60
TTTCAATCTATTATATTTATAGAATTTTTTGTTGCTAGATTTGTCAAATTGCTTAAAATA 120
ATTTTTTTCAGAAAGCAAAAGCCGATACCTATCGAGTAGGGTAGTTCTTGCTATCGTCAG 180
ORF 2
box A
GCTTGTCTGTAGGTGTTAACACTTTTCAAAAATCTCTTCAAACAACGTCAGCTTTGCCTT 240
1 V L T L F K N L F K Q R Q L C L
box B1
GCCGTATATATGTTACTGACTTCGTCAGTTCTATCTGCCACCTCAAAACGGTGTTTTGAG 300
17 A V Y M L L T S S V L S A T S K R C F E
box B2
CTGACTTCGTCAGTTCTATCCACAACCTCAAAACAGTGTTTTGAGCTGACTTCGTCAGTT 360
37 L T S S V L S T T S K Q C F E L T S S V
box B3
box B4
CTATCCACAACCTCAAAACAGTGTTTTGAGCTGACTTTGTCAGTCTTATCTACAACCTCA 420
57 L S T T S K Q C F E L T L S V L S T T S
box C
AAACAGTGTTTTGAGCATCATGCGGCTAGCTTCTTAGTTTGCTCTTTGATTTTCATTGAG 480
77 K Q C F E H H A A S F L V C S L I F I E
TATAAAAACAGATGAGTTTCTGTTTTCTTTTTATGGACTATAAATGTTCAGCTGAAACTA 540
97 Y K N R *
CTTTCAAGGACATTATTATATAAAAGAATTTTTTGAAACTAAAATCTACTATATTACACT 600
-35
-10
ATATTGAAAGCGTTTTAAAAATGAGGTATAATAAATTTACTAACACTTATAAAAAGTGAT 660
AGAATCTATCTTTATGTATATTTAAAGATAGATTGCTGTAAAAATAGTAGTAGCTATGCG 720
ORF 3
AAATAACAGATAGAGAGAAGGGATTGAAGCTTAGAAAAGGGGAATAATATGATATTTAAG 780
1 M I F K
GCATTCAAGACAAAAAAGCAGAGAAAAAGACAAGTTGAACTACTTTTGACAGTTTTTTTC 840
5 A F K T K K Q R K R Q V E L L L T V F F
GACAGTTTTCTGATTGATTTATTTCTTCACTTATTTGGGATTGTCCCCTTTAAGCTGGAT 900
25 D S F L I D L F L H L F G I V P F K L D
AAGATTCTGATTGTGAGCTTGATTATATTTCCCATTATTTCTACAAGTATTTATGCTTAT 960
45 K I L I V S L I I F P I I S T S I Y A Y
-35
-10
GAAAAGCTATTTGAAAAAGTGTTCGATAAGGATTGAGCAGGAAGTATGGTGTAAATAGCA 1020
65 E K L F E K V F D K D *


-35
-10
TAAGCTGATGTCCATCATTTGCTTATAAAGAGATATTTTAGTTTAATTGCAGCGGTGTCC 1080
TGGTAGATAAACTAGATTGGCAGGAGTCTGATTGGAGAAAGGAGAGGGGAAATTTGGCAC 1140
CAATTTGAGATAGTTTGTTTAGTTCATTTTTGTCATTTAAATGAACTGTAGTAAAAGAAA 1200
ORF 1
GTTAATAAAAGACAAACTAAGTGCATTTTCTGGAATAAATGTCTTATTTCAGAAATCGGG 1260
1 M S Y F R N R D
SD
ATATAGATATAGAGAGGAACAGTATGAATCGGAGTGTTCAAGAACGTAAGTGTCGTTATA 1320
9 I D I E R N S M N R S V Q E R K C R Y S
SD
GCATTAGGAAACTATCGGTAGGAGCGGTTTCTATGATTGTAGGAGCAGTGGTATTTGGAA 1380
29 I R K L S V G A V S M I V G A V V F G T
CGTCTCCTGTTTTAGCTCAAGAAGGGGCAAGTGAGCAACCTCTGGCAAATGAAACTCAAC 1440
49 S P V L A Q E G A S E Q P L A N E T Q L
SP
TTTCGGGGGAGAGCTCAACCCTAACTGATACAGAAAAGAGCCAGCCTTCTTCAGAGACTG 1500
69 S G E S S T L T D T E K S Q P S S E T E
AACTTTCTGGCAATAAGCAAGAACAAGAAAGGAAAGATAAGCAAGAAGAAAAAATTCCAA 1560
89 L S G N K Q E Q E R K D K Q E E K I P R
GAGATTACTATGCACGAGATTTGGAAAATGTCGAAACAGTGATAGAAAAAGAAGATGTTG 1620
109 D Y Y A R D L E N V E T V I E K E D V E
AAACCAATGCTTCAAATGGTCAGAGAGTTGATTTATCAAGTGAACTAGATAAACTAAAGA 1680
129 T N A S N G Q R V D L S S E L D K L K K
AACTTGAAAACGCAACAGTTCACATGGAGTTTAAGCCAGATGCCAAGGCCCCAGCATTCT 1740
149 L E N A T V H M E F K P D A K A P A F Y
ATAATCTCTTTTCTGTGTCAAGTGCTACTAAAAAAGATGAGTACTTCACTATGGCAGTTT 1800
169 N L F S V S S A T K K D E Y F T M A V Y
ACAATAATACTGCTACTCTAGAGGGGCGTGGTTCGGATGGGAAACAGTTTTACAATAATT 1860
189 N N T A T L E G R G S D G K Q F Y N N Y
ACAACGATGCACCCTTAAAAGTTAAACCAGGTCAGTGGAATTCTGTGACTTTCACAGTTG 1920
209 N D A P L K V K P G Q W N S V T F T V E
AAAAACCGACAGCAGAACTACCTAAAGGCCGAGTGCGCCTCTACGTAAACGGGGTATTAT 1980
229 K P T A E L P K G R V R L Y V N G V L S
CTCGAACAAGTCTGAGATCTGGCAATTTCATTAAAGATATGCCAGATGTAACGCATGTGC 2040
249 R T S L R S G N F I K D M P D V T H V Q
AAATCGGAGCAACCAAGCGTGCCAACAATACGGTTTGGGGGTCAAATCTACAGATTCGGA 2100
269 I G A T K R A N N T V W G S N L Q I R N
ATCTCACTGTGTATAATCGTGCTTTAACACCAGAAGAGGTACAAAAACGTAGTCAACTTT 2160
289 L T V Y N R A L T P E E V Q K R S Q L F
TTAAACGCTCAGATTTAGAAAAAAAACTACCTGAAGGAGCGGCTTTAACAGAGAAAACGG 2220
309 K R S D L E K K L P E G A A L T E K T D

```
     ACATATTCGAAAGCGGGCGTAACGGTAAACCAAATAAAGATGGAATCAAGAGTTATCGTA  2280
329    I  F  E  S  G  R  N  G  K  P  N  K  D  G  I  K  S  Y  R  I

     TTCCAGCACTTCTCAAGACAGATAAAGGAACTTTGATCGCAGGTGCAGATGAACGCCGTC  2340
349    P  A  L  L  K  T  D  K  G  T  L  I  A  G  A  D  E  R  R  L

     TCCATTCGAGTGACTGGGGTGATATCGGTATGGTCATCAGACGTAGTGAAGATAATGGTA  2400
369    H  S  S  D  W  G  D  I  G  M  V  I  R  R  S  E  D  N  G  K

     AAACTTGGGGTGACCGAGTAACCATTACCAACTTACGTGACAATCCAAAAGCTTCTGACC  2460
389    T  W  G  D  R  V  T  I  T  N  L  R  D  N  P  K  A  S  D  P

     CATCGATCGGTTCACCAGTGAATATCGATATGGTGTTGGTTCAAGATCCTGAAACCAAAC  2520
409    S  I  G  S  P  V  N  I  D  M  V  L  V  Q  D  P  E  T  K  R

     GAATCTTTTCTATCTATGACATGTTCCCAGAAGGGAAGGGAATCTTTGGAATGTCTTCAC  2580
429    I  F  S  I  Y  D  M  F  P  E  G  K  G  I  F  G  M  S  S  Q

     AAAAAGAAGAAGCCTACAAAAAAATCGATGGAAAAACCTATCAAATCCTCTATCGTGAAG  2640
449    K  E  E  A  Y  K  K  I  D  G  K  T  Y  Q  I  L  Y  R  E  G

     GAGAAAAGGGAGCTTATACCATTCGAGAAAATGGTACTGTCTATACACCAGATGGTAAGG  2700
469    E  K  G  A  Y  T  I  R  E  N  G  T  V  Y  T  P  D  G  K  A

     CGACAGACTATCGCGTTGTTGTAGATCCTGTTAAACCAGCCTATAGCGACAAGGGGGATC  2760
489    T  D  Y  R  V  V  V  D  P  V  K  P  A  Y  S  D  K  G  D  L

     TATACAAGGGTAACCAATTACTAGGCAATATCTACTTCACAACAAACAAAACTTCTCCAT  2820
509    Y  K  G  N  Q  L  L  G  N  I  Y  F  T  T  N  K  T  S  P  F

     TTAGAATTGCCAAGGATAGCTATCTATGGATGTCCTACAGTGATGACGACGGGAAGACAT  2880
529    R  I  A  K  D  S  Y  L  W  M  S  Y  S  D  D  D  G  K  T  W

     GGTCAGCGCCTCAAGATATTACTCCGATGGTCAAAGCCGATTGGATGAAATTCTTGGGTG  2940
549    S  A  P  Q  D  I  T  P  M  V  K  A  D  W  M  K  F  L  G  V

     TAGGTCCTGGAACAGGAATTGTACTTCGGAATGGGCCTCACAAGGGACGGATTTTGATAC  3000
569    G  P  G  T  G  I  V  L  R  N  G  P  H  K  G  R  I  L  I  P

     CGGTTTATACGACTAATAATGTATCTCACTTAAATGGCTCGCAATCTTCTCGTATCATCT  3060
589    V  Y  T  T  N  N  V  S  H  L  N  G  S  Q  S  S  R  I  I  Y

     ATTCAGATGATCATGGAAAAACTTGGCATGCTGGAGAAGCGGTCAACGATAACCGTCAGG  3120
609    S  D  D  H  G  K  T  W  H  A  G  E  A  V  N  D  N  R  Q  V

     TAGACGGTCAAAAGATCCACTCTTCTACGATGAACAATAGACGTGCGCAAAATACAGAAT  3180
629    D  G  Q  K  I  H  S  S  T  M  N  N  R  R  A  Q  N  T  E  S

     CAACGGTGGTACAACTAAACAATGGAGATGTTAAACTCTTTATGCGTGGTTTGACTGGAG  3240
649    T  V  V  Q  L  N  N  G  D  V  K  L  F  M  R  G  L  T  G  D

     ATCTTCAGGTTGCTACAAGTAAAGACGGAGGAGTGACTTGGGAGAAGGATATCAAACGTT  3300
669    L  Q  V  A  T  S  K  D  G  G  V  T  W  E  K  D  I  K  R  Y

     ATCCACAGGTTAAAGATGTCTATGTTCAAATGTCTGCTATCCATACGATGCACGAAGGAA  3360
689    P  Q  V  K  D  V  Y  V  Q  M  S  A  I  H  T  M  H  E  G  K
```


AAGAATACATCATCCTCAGTAATGCAGGTGGACCGAAACGTGAAAATGGGATGGTCCACT 3420
709 E Y I I L S N A G G P K R E N G M V H L
TGGCACGTGTCGAAGAAAATGGTGAGTTGACTTGGCTCAAACACAATCCAATTCAAAAAG 3480
729 A R V E E N G E L T W L K H N P I Q K G
GAGAGTTTGCCTATAATTCGCTCCAAGAATTAGGAAATGGGGAGTATGGCATCTTGTATG 3540
749 E F A Y N S L Q E L G N G E Y G I L Y E
AACATACTGAAAAAGGACAAAATGCCTATACCCTATCATTTAGAAAATTTAATTGGGACT 3600
769 H T E K G Q N A Y T L S F R K F N W D F
TTTTGAGCAAAGATCTGATTTCTCCTACCGAAGCGAAAGTGAAGCGAACTAGAGAGATGG 3660
789 L S K D L I S P T E A K V K R T R E M G
GCAAAGGAGTTATTGGCTTGGAGTTCGACTCAGAAGTATTGGTCAACAAGGCTCCAACCC 3720
809 K G V I G L E F D S E V L V N K A P T L
TTCAATTGGCAAATGGTAAAACAGCACGCTTCATGACCCAGTATGATACAAAAACCCTCC 3780
829 Q L A N G K T A R F M T Q Y D T K T L L
TATTTACAGTGGATTCAGAGGATATGGGTCAAAAAGTTACAGGTTTGGCAGAAGGTGCAA 3840
849 F T V D S E D M G Q K V T G L A E G A I
TTGAAAGTATGCATAATTTACCAGTCTCTGTGGCGGGCACTAAGCTTTCGAATGGAATGA 3900
869 E S M H N L P V S V A G T K L S N G M N
ACGGAAGTGAAGCTGCTGTTCATGAAGTGCCAGAATACACAGGCCCATTAGGGACATCCG 3960
889 G S E A A V H E V P E Y T G P L G T S G
GCGAAGAGCCAGCTCCAACAGTCGAGAAGCCAGAATACACAGGCCCACTAGGGACATCCG 4020
909 E E P A P T V E K P E Y T G P L G T S G
GCGAAGAGCCAGCCCCGACAGTCGAGAAGCCAGAATACACAGGCCCACTAGGGACAGCTG 4080
929 E E P A P T V E K P E Y T G P L G T A G
GTGAAGAAGCAGCTCCAACAGTCGAGAAGCCAGAATTTACAGGGGGAGTTAATGGTACAG 4140
949 E E A A P T V E K P E F T G G V N G T E
AGCCAGCTGTTCATGAAATCGCAGAGTATAAGGGATCTGATTCGCTTGTAACTCTTACTA 4200
969 P A V H E I A E Y K G S D S L V T L T T
CAAAAGAAGATTATACTTACAAAGCTCCTCTTGCTCAGCAGGCACTTCCTGAAACAGGAA 4260
989 K E D Y T Y K A P L A Q Q A L P E T G N
ACAAGGAGAGTGACCTCCTAGCTTCACTAGGACTAACAGCTTTCTTCCTTGGTCTGTTTA 4320
1009 K E S D L L A S L G L T A F F L G L F T
CGCTAGGGAAAAAGAGAGAACAATAAGAGAAGAATTCTAAACATTTGATTTTGTAAAAAT 4380
1029 L G K K R E Q *
AGAAGGAGATAGCAGGTTTTCAAGCCTGCTATCTTTTTTTGATGACATTCAGGCTGATAC 4440
T T
GAAATCATAAGAGGTCTGAAACTACTTTCAGAGTAGTCTGTTCTATAAAATATAGTAGAT 4500
TGAAATAAGATGTGAACAACTCTATCAGGAAAGTCAAATTAATTTATAGAAAAATTTTAG 4560

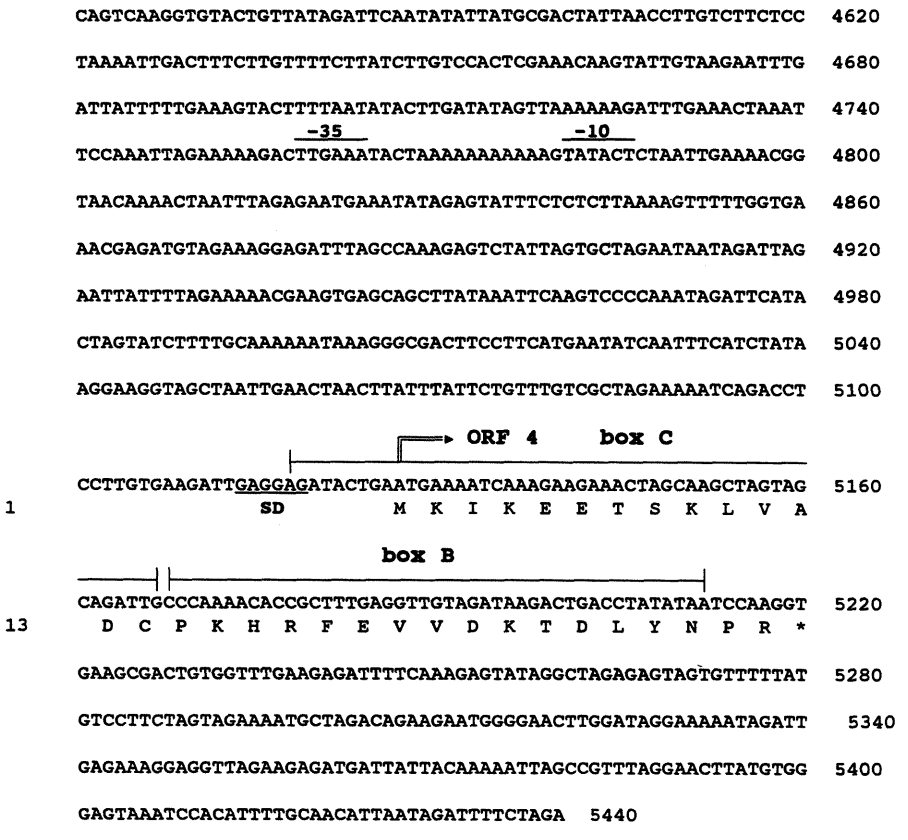


FIGURE 5.8. Sequence from pMC2150 and pMC4170. The start of the ORFs is indicated by arrows. Putative -35 and -10 promotor sequences are also indicated. Possible Shine Dalgarno sequences (SD) are marked. Transcription terminators are represented by a broken line. Unusual DNA sequences (box elements) are defined. Signal peptide is double underlined with an arrow in the predicted recognition site for the signal peptidase (SP). Aspartic boxes are underlined with the conserved sequences in bold. The three amino acid repeats are underlined near the end of ORF 1. Close to this C-terminus there is present a surface anchor motif in bold, followed by an arrow underlined hydrophobic domain.

5.3.1. Neuraminidase ORF (ORF1)

ORF1 has a size of 3,105 bp starting from the ATG at position 1239 and finishing at the TAA at position 4344 [FIGURE 5.8]. This ORF would code for a protein of 114 KD. Two other ATGs, located at positions 1284 and 1353 are in the same reading frame as the ATG at 1239. They would give two proteins of maximum predicted sizes of 112 KD and 110KD. Only these last two ATGs are associated with a putative Shine Dalgarno sequences [FIGURE 5.8]. Although the first ATG has no typical Shine Dalgarno sequence, this fact does not exclude it from being the real translation starting point, since the presence of unusual Shine-Dalgarno sequences in other genes from *S. pneumoniae* has been observed [De la Campa *et al* 1987].

Two putative -10 and -35 promoters sequences have been found at -172 and -218 from the first ATG. Although very distant from this ATG, this situation resembles the one for the pneumolysin gene, also from the pneumococcus, having a putative promoter at position -164 from the translation start codon [Walker *et al* 1987].

Downstream to the stop codon there is a putative transcription terminator [FIGURE 5.8] that could form the following loop:

```
 U C
U  A
U  G
 G:C
 G:C
 A:U
 C:G          ΔG = -18.9 Kcal/mol
 G:C
 A:U
 U:A
 A:U
 G:C
.....A:U U U U U U U U -3'OH
```

This structure resembles that of a typical *rho*-independent transcription terminator as it possesses the two most common features of these terminators. Firstly, it is 50% GC rich, and could therefore form a stable

loop. Secondly, it has a poly-U tail, making the RNA-DNA hybrid very unstable, causing rapid RNA chain release from the DNA duplex without the requirement of a *rho* factor [Platt & Bear 1983].

5.3.2. ORF2

ORF2 is 300 bp in length. It starts with GUG codon located at position 193, and it finishes in a TGA stop codon at position 493 [FIGURE 5.8]. Detailed analysis of this sequence has failed to find any putative promoter, Shine-Dalgarno or terminator sequences. This sequence spans a region of tandem repeats that will be mentioned below.

5.3.3. ORF3

ORF3 falls between ORF2 and the neuraminidase gene. It starts with an ATG codon at position 769, and finishes in a TGA stop codon at position 994, therefore the maximum predicted size for this ORF is 225 bp [FIGURE 5.8]. No potential ribosome binding site has been found associated with ORF3. There are putative -30 and -10 promotor-like sequences at positions -137 and -160. No apparent transcription terminator has been found.

5.3.4. ORF4

This ORF, although only 93 bp long, falls within a sequence, that will be discussed below. This ORF starts with an ATG codon at position 5127, finishing in a TGA stop codon at position 5220 [FIGURE 5.8]. There is a putative ribosome binding site present. Also, a putative promotor sequences have been found at positions -341 and -363. There are no apparent transcription terminators associated to ORF4.

5.3.5. Presence of large tandem repeats in sequences flanking the neuraminidase ORF

During the process of DNA sequence determination from pMC2150, a region of DNA which gave rise to compressions of the sequence on the DNA sequencing gels was observed. The sequence between these compressions was repeated four times [FIGURE 5.9]. When the DNA sequence was resolved,

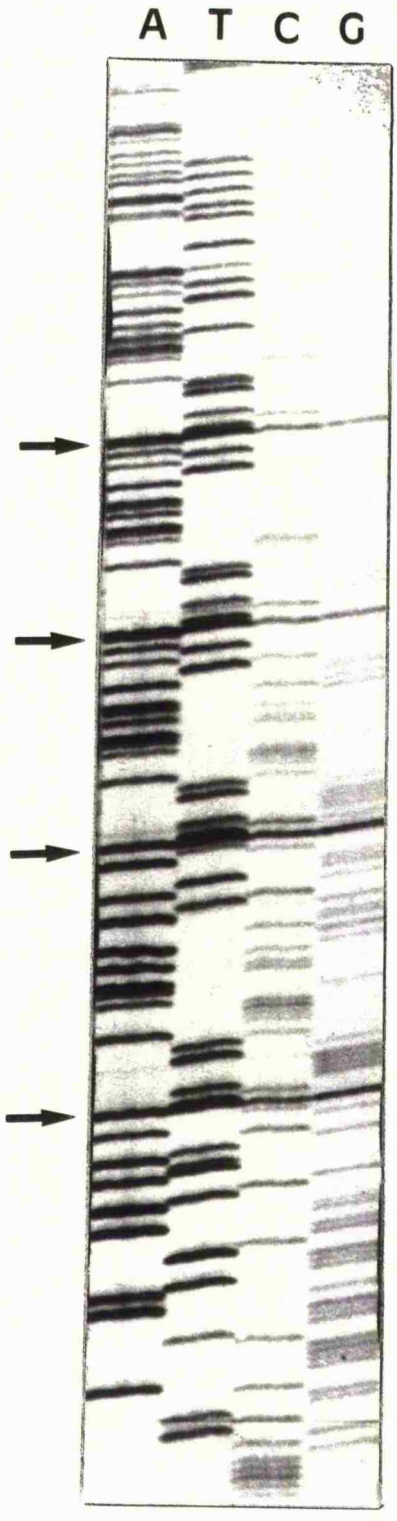


FIGURE 5.9. **DNA compression in sequencing gels of pneumococcal DNA corresponding to the box elements.**

Compressions are indicated by arrows. Sequences between arrows correspond to single DNA repeated elements.

further analysis showed that the regions of compression corresponded to four 45 bp tandem direct repeats located 752 bp upstream to the ORF1. The alignment of these repeats is shown in FIGURE 5.10. This region is spanned by ORF2.

```
256     CTGACTTCGTCAGTTCTATCTGCCACCTCAAAACGGTGTTTTGAG     300
        |||||||||||||||||||||  | ||||||||||| ||||||||||
301     CTGACTTCGTCAGTTCTATCCACAACCTCAAAACAGTGTTTTGAG     336
        |||||||||||||||||||||||||||||||||||||||||||||
337     CTGACTTCGTCAGTTCTATCCACAACCTCAAAACAGTGTTTTGAG     390
        ||||||| ||||||  |||| ||||||||||||||||||||||||
391     CTGACTTTGTCAGTCTTATCTACAACCTCAAAACAGTGTTTTGAG     435
```

FIGURE 5.10. Alignment of the 45 bp DNA repeats found in ORF2.

A single DNA repeat was subjected to comparison with the whole DNA sequence determined from pMC2150 and pMC4170 by using BIOSEQ (Leicester University Sequence Analysis Software). It was found that a single repeat was also present downstream to the neuraminidase gene, although on the opposite strand [FIGURE 5.8:box B].

As will be described later, these repeats were subject to a database search for homologies to other sequences.

5.4. ANALYSIS OF THE NEURAMINIDASE PREDICTED PROTEIN SEQUENCE

5.4.1. Determination of the hydrophobicity profile

The neuraminidase hydrophobicity profile was determined using the method of Kyte-Doolittle [Kyte & Doolittle 1982]. This was achieved by using PEPPLOT from the University of Wisconsin Sequence Analysis Software. For this analysis the 114 KD predicted protein encoded by the 3.1 Kb of ORF1 was selected. The window used was 9. The resulting plot is shown in FIGURE 5.11. This plot reveals a very hydrophilic protein with hydrophobic regions at both the N-terminus, running from amino acids 32 to 49, and C-terminus, from amino acids 1013 to 1030.

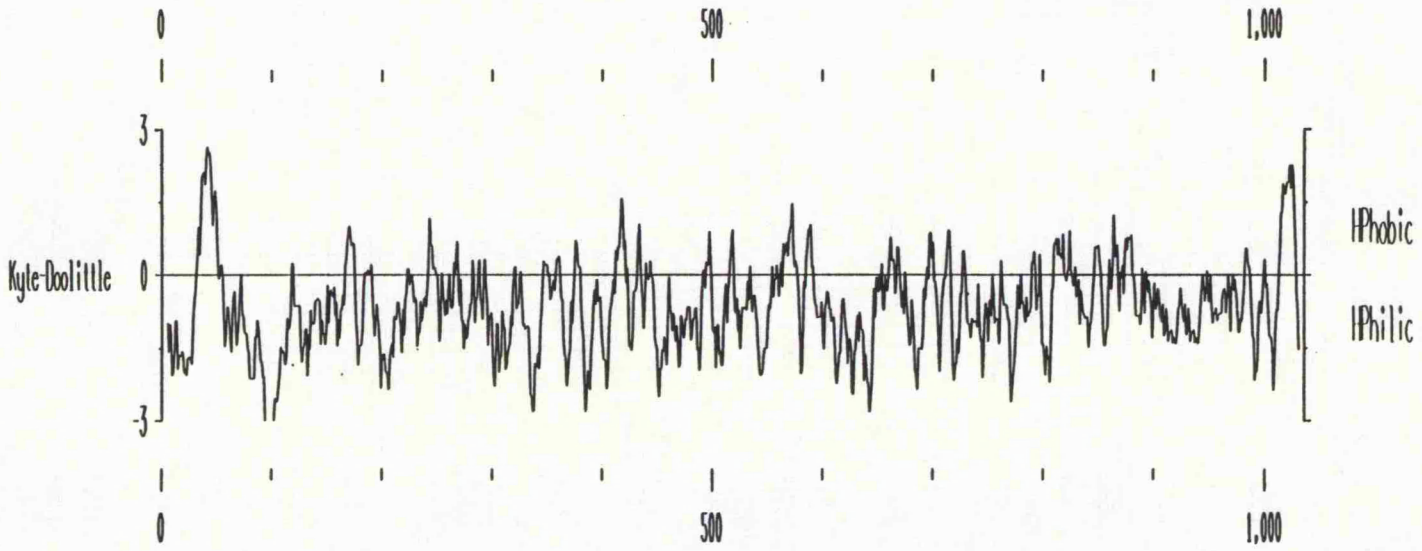


FIGURE 5.11. **Kyte-Doolittle analysis of the predicted neuraminidase protein sequence.**
Plot above the line represents hydrophobic regions whereas plot below the line represents hydrophilic domains. Numbers under and over the plots represent amino acids from the protein sequence. The window size used was 9.

5.4.2. Presence of a potential signal peptide sequence in the neuraminidase predicted amino acid sequence

Signal sequences are those amino acid sequences that have an influence on the folding of the precursor protein. Other function of the signal sequence include recognition by factors involved in protein translocation and interaction with the membrane [Austen & Westwood 1991]. These signal peptides have the following structure [FIGURE 5.12]:

a) An N-terminal region containing one or more positively charged amino acids [Silhavy *et al* 1983].

b) A core rich in hydrophobic residues, especially those with the potential for forming α-helices (Phe, Leu, Val, Ile, Met, Trp) but also with some helix-disrupting amino acids like proline and glycine. This region can be between 9 and 18 amino acids long and terminates with one of the helix-disrupting amino acids [Austen & Westwood 1991].

c) A hydrophilic carboxy-terminus region of approximately 4 to 6 amino acids, defining the cleavage region for the signal peptidase which has a very high degree of conservation, conforming to the "-1,-3 rule". Essentially, the sequence of the cleavage site is A-X-B (where B is Ala, Gly or Ser; A is any of B or Leu, Val or Ile) [von Heijne 1983].

Analysis of the N-terminus of neuraminidase revealed the presence of a putative signal peptide that starts at the second methionine of the predicted amino acid sequence. The structure of this signal peptide is shown in FIGURE 5.12. The features previously described can be observed in this peptide:

a) An N-terminal region with a net positive charge.

b) A hydrophobic region of 18 amino acids most of which potentially form α-helix. This region corresponds to the N-terminal hydrophobic domain showed in the Kyte-Doolittle analysis [FIGURE 5.11]. The hydrophobic

core also includes three Gly and ends in a Pro which potentially disrupts α-helices.

c) After the Pro that terminates the hydrophobic region there is present the "-1,-3" sequence represented by Val-Leu-Ala, in which -1 is Val and -3 is Ala.

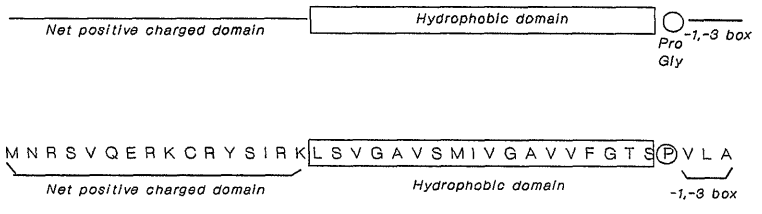


FIGURE 5.12. Structure of the putative signal peptide found in the N-terminus of the neuraminidase sequence. Above this sequence there is represented the structure of a consensus signal peptide.

5.4.3. Presence of a neuraminidase sequence motif in the predicted amino acid sequence of pneumococcal neuraminidase

Comparison of all the microbial neuraminidase sequences has revealed the presence of a common motif (see introduction). This motif has been called "Aspartic box" [Roggentin *et al* 1989] and has the following sequence: -Ser-X-Asp-X-Gly-X-Thr-Trp-.

Detailed analysis of the predicted protein from ORF1 has shown the presence of this "aspartic box" four times in the sequence [FIGURE 5.8] resembling most other microbial neuraminidases [FIGURE 5.13]. The distance between the second and third motifs is very conserved in all these neuraminidases except in the one from *T. cruzi* [FIGURE 5.13].

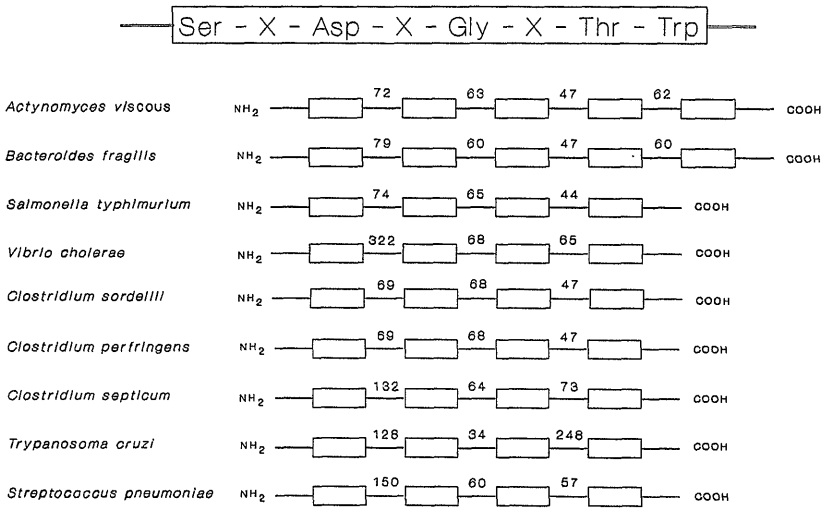


FIGURE 5.13. Presence of the aspartic box in all the microbial neuraminidases sequenced. Boxes represent aspartic boxes, and numbers represent distance in amino acids between boxes (see introduction for references).

5.4.4. Presence of putative surface protein anchor motif in the C-terminus of neuraminidase

Kyte-Doolittle analysis has shown the presence of a hydrophobic domain in the C-terminus of the protein [FIGURE 5.11]. This region has similar features to the C-terminal hydrophobic domains of some Gram-positive surface proteins. These proteins have in common the presence of a consensus anchor motif structure that is shown in FIGURE 5.14 A. This motif consists of up to seven charged amino acids, composed of a mixture of both negative and positive charged residues, which are found at the C-terminus. Immediately N-terminal to this short charged region there is a segment of 15 to 22 predominantly hydrophobic amino acids. Beginning about 1 to 6 amino acids N-terminal to the hydrophobic domain is found a hexapeptide with the consensus sequence LPXTGX that is highly conserved among all the Gram-

positive surface proteins analyzed. Finally, continuing towards the N-terminus there is a region that spans about 50 to as much as 125 residues with a high percentage of proline/glycine and threonine/serine residues [Fischetti *et al* 1991].

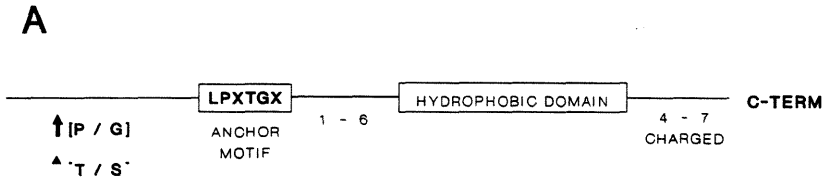


FIGURE 5.14. Presence of a Gram-positive surface protein anchor motif.

A. Consensus structure of a Gram-positive surface protein anchor motif [Fischetti *et al* (1991)].

B. Structure of the C-terminus of neuraminidase resembling the Gram-positive surface anchor motifs. The shadowed box indicates the presence of the three 20 amino acid repeats, within the 100 amino acid P/G-T/S rich region, found near the C-terminus [FIGURE 5.8]

Detailed analysis of this structure in the C-terminus of the neuraminidase shows a structure that resembles that of surface proteins from Gram-positive bacteria. This structure, shown in FIGURE 5.14 B, is as follows. A region of 5 charged amino acids in the C-terminus is preceded by a hydrophobic domain of 18 amino acids. Ten amino acids N-terminal to the hydrophobic domain there is the sequence LPETGN. Continuing towards the N-terminus a region of 100 residues containing 27% proline/glycine and 19% threonine/serine is present. This last region contains a sequence of 20 amino acids tandemly repeated three times. Identical repeats have not been found in any other C-terminal region of proteins from Gram-positive organisms [FIGURE 5.8.][FIGURE 5.15].

```
899   Y T G P L G T S G E E P A P T V E K P E   918
      | | | | | | | | | | | | | | | | | |
919   Y T G P L G T S G E E P A P T V E K P E   938
      | | | | | | |   | | |   | | | | | | |
939   Y T G P L G T A G E E A A P T V E K P E   958
```

FIGURE 5.15. Sequence alignment of the amino acid sequence repeats found near the C-terminus of neuraminidase.

Therefore the C-terminus of the predicted neuraminidase amino acid sequence clearly resembles the motif found in C-terminal sequences of many surface proteins from Gram-positive bacteria. The alignment of these sequences is shown in FIGURE 5.16.

5.5. DATABASE SEARCHES.

5.5.1. DNA DATABASES SEARCH.

As described above, four 45 pb tandem repeats were found 752 bp upstream to the neuraminidase gene. The role of these sequences is not clear. A clue about a possible function for these sequences might be obtained from homologous sequences in the DNA databases.

The EMBL databases were searched with the University of Wisconsin Sequence Analysis Software choosing FASTA which uses the method of Pearson & Lipman (1988) to search for similarities between one sequence (the query) and any group of sequences. Word sizes of 1, 3 and 6 were selected for these searches. In all the searches either the four 45 bp repeats or a single 45 bp repeat was used as query sequences. Homology to upstream regions of three pneumococcal genes was found. These three genes were:

a/ *hexB* which codes for a protein involved in DNA mismatch repair [Prudhomme *et al* 1989].

b/ *comA* which codes for a protein involved in competence for transformation in the pneumococcus [Hui & Morrison 1989].

c/ *lytA* which codes for pneumococcal autolysin [García *et al* 1986].

		REFERENCES
Protein H	**LPSTGE** TAN PFFTAAALTVMATAGVAAVV KRKEEN	Gomi *et al* (1990)
Protein G	**LPTTGE** GSN PFFTAAALAVMAGAGALAVA SKRKED	Olsson *et al* (1987)
Protein A	**LPETGE** EN PLIGTTVFGGLSLALGAALLAG RRREL	Guss *et al* (1984)
M6,M5	**LPSTGE** TAN PFFTAAALTVMATAGVAAVV KRKEEN	Hollingstead *et al* (1989), Miller *et al* (1988)
M55	**LPSTGE** ATN PFFTAAALTVMATAGVAAVV KRKEEN	Hollingstead *et al* (1987)
M19,M24,M30	**LPSTGE** ATN PFFTAAALTVMATAGVAAVV KRKEEN	Hollingstead *et al* (1977), Mouw *et al* (1988)
M49,ARP4	**LPSTGE** TAN PFFTAAAATVMVSAGMLAL KRKEEN	Frithz *et al* (1989), Haanes & Cleary (1989)
T6	**LPSTGS** IGT YLFKAIGSAAMIGAIGIYIV KRRKA	Schneewind *et al* (1990)
PAc	**LPNTGV** TNN AYMPLLGIIGLVTSFSLLGL KAKKD	Okahashi *et al* (1989)
Sec10	**LPQTGE** QQS IWLTIIGLLMAAGTI KNKKRKKNS	Fischetti *et al* (1991)
FcRA	**LPSTGE** ETTN PFFTA...	Heath & Cleary (1898)
wapA	**LPSTGE** Q AGLLLTTVGLVIVAVAGVYFY RTRR	Ferretti *et al* (1989)
FnBp	**LPETGG** EESTNK MLFGGLFSILGLALL RRNKKNHKA	Signas *et al* (1989)
wg2	**LPKTGE** TTER PAFGFLGVIVVILMGVLGL KRKQREE	Kok *et al* (1988)
Asc10	**LPKTGE** KQN VLLTVVGSLAAMLGLAGLGF KRRKETK	Fischetti *et al* (1991)
spaP	**LPNTGV** TNN AYMPLLGIIGLVTSFSLLGL KAKKD	Kelly *et al* (1990)
SCP	**LPTTND** KDTNR LHLLKLVMTTFFFGLVAHIF KTKQKETKK	Chen & Cleary (1990)
nanA	**LPETGN** KESD LLASLGLTAFFLGLFTLG KKREQ	This thesis

FIGURE 5.16. Alignements of all the Gram-positive consensus surface anchor sequences. Anchor motifs are in bold whereas the hydrophobic regions are in shadow.

In each the region of homology with the pneumococcal neuraminidase associated sequences were upstream to those genes [FIGURE 5.17].

Further database searches were done using the same program and parameters as before. For these searches the DNA comprising ORF2 was used since this ORF spans the DNA repeat region of interest. From these searches it was found that not only the repeat region, but also their flanking sequences, were homologous to upstream sequences of *comA* and *hexB* [FIGURE 5.17].

A search for the presence of homologous sequences to ORF2 in the flanking regions of the pneumolysin gene (*ply*) [Walker *et al* 1987] was done. In this sequence comparison the same parameters as before were applied but using BIOSEQ (Leicester University Sequence Analysis Software). DNA homology to ORF2 was also found downstream to *ply*.

Homologous sequences to these were discovered downstream to the neuraminidase ORF. An identical search to the one with pneumolysin revealed that the flanking sequences were also homologous.

Three types of repeated sequences have been found with different length and sequence amongst them. A single repeat of each type have been named "box". These sequences are as follows:

1/ A 59 bp sequence called box A.

2/ A 45 bp sequence that can be unique or tandemly repeated up to four times, called box B (box B1 to box Bn).

3/ A 50 bp sequence called box C.

The distribution of these repeated sequences in the flanking regions of the pneumococcal genes mentioned above is shown in FIGURE 5.17. This figure also shows the distance between the ORFs of the different genes as well as the direction of the ORFs with respect to the boxes.

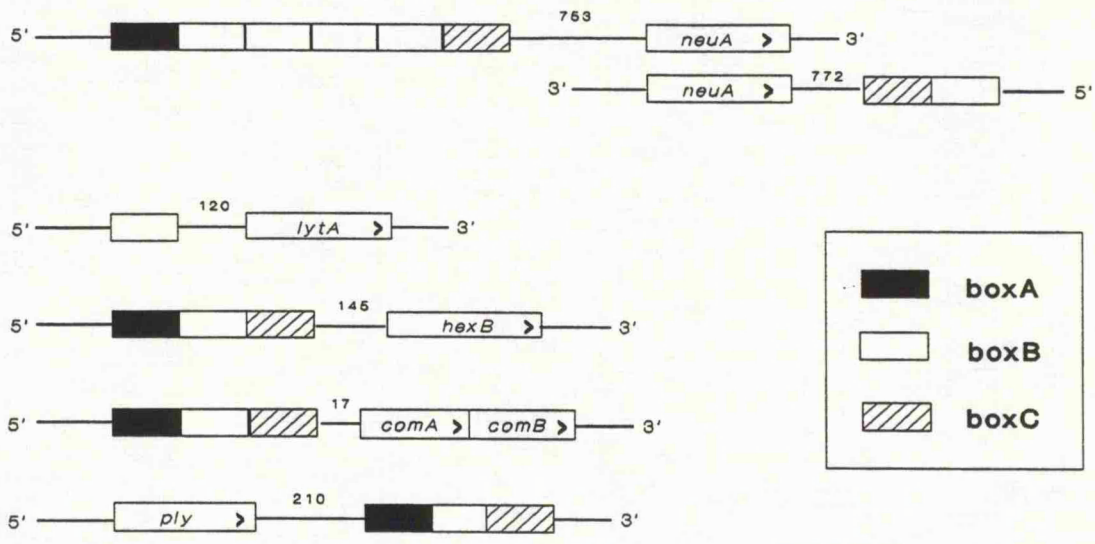


FIGURE 5.17. Representation of the arrangement of boxA, boxB and boxC sequences in the flanking regions of the pneumococcal genes: *neuA*, *lytA*, *hexB*, *comA* and *ply*. Numbers represent nucleotide distance between boxes and ORFs. Arrows symbolize the direction of the translation of the ORF.

An alignment of the box elements found in the vicinity of the different pneumococcal genes mentioned above is shown in FIGURE 5.18. From this alignment a consensus sequence for each box element was constructed using information obtained from J.P. Claverys (personal communication) [FIGURE 5.18]. Not all the boxes are present in the flanking regions of the genes already mentioned. A large variation in the length boxA and boxB near the pneumolysin gene, with respect to the consensus sequence, is shown in FIGURE 5.18.

box A

```
hexB    TAATACTCTTCaAAAATCTC.TTCAAACCACGTCAGCGTCGtCTTGCCGTAGATA.GTTg
comA    TtATACTCCTCGAAAATCaaaTTCAAACCACGTCAaCGTCGCCTTGCCGTAtATATGTgA
ply     TtATACTCTTaGAAAATCTC.TTCAAACCAtGTCAGC.....................
neuA    TAAcACTcTTCaAAAATCTC.TTCAAACaACGTCAGCtTtGCCTTGCCGTAtATATGTTA
        .       .       .       .       .       .       .
        1       10      20      30      40      50      59
CONS    TAATACTCTTCGAAAATCTC.TTCAAACCACGTCAGCGTCGCCTTGCCGTAGATATGTTA
```

box B

```
hexB    tTGACTTtGTCAGTTtTATCTACAACCTCAAAACAGTG..TTTTGAa
comA    CTGACTTCGTCAGTcCTATCTACAACCTCAAAACAGTG..TTTTGAG
lytA    CaGACcTtGTCAGTcCTATtTACAgtgTCAAAAtAGTGcgTTTTGAa
ply     ..............cCTATCcgCAACCTCAAAACAGTG..TTTTGAG
neuA1   CTGACTTCGTCAGTTCTATCTgCcACCTCAAAACgGTG..TTTTGAG
neuA2   CTGACTTCGTCAGTTCTATCcACAACCTCAAAACAGTG..TTTTGAG
neuA3   CTGACTTCGTCAGTTCTATCcACAACCTCAAAACAGTG..TTTTGAG
neuA4   CTGACTTtGTCAGTctTATCTACAACCTCAAAACAGTG..TTTTGAG
neuA*   tTataTagGTCAGTcTTATCTACAACCTCAAAgCgGTG..TTTTGgG
        .       .       .       .       .   .
        1       10      20      30      40  45
CONS    CTGACTTCGTCAGTTCTATCTACAACCTCAAAACAGTG..TTTTGAG
```

box C

```
hexB    CtgtCTGCGcCTAGCTTtCTAGTTTGCTCTTTGATTTTtATTGAGTATAc
comA    CAgCCTGCGGCTAGtTTCCTAGTTTGCTCTTTGATTTTCATTGAGTATtA
ply     CAACCTGCGGCTAGCTTtCTAGTTTGCTCTTTGATTTTtATTGAGTATtA
neuA    CAtCaTGCGGCTAGCTTCtTAGTTTGCTCTTTGATTTTCATTGAGTATAA
neuA*   CAAtCTGCtaCTAGCTTgCTAGTTTctTCTTTGATTTTCATTCAGTATct
        .       .       .       .       .       .
        1       10      20      30      40      50
CONS    CAACCTGCGGCTAGCTTCCTAGTTTGCTCTTTGATTTTCATTGAGTATAA
```

FIGURE 5.18. **Alignment of the DNA sequences from the homologous sequences found in the flanking regions of several pneumococcal genes.**
Abbreviations of the genes are described in the text. neuA* represents the homologous sequences found downstream to the neuraminidase gene. Nucleotides in large letters represent conserved bases with respect to the consensus sequence (CONS), whereas nucleotides in lower case letters represent unconserved bases. Box A, B and C, in the different genes, represent contiguous sequences.

5.5.2. PROTEIN DATABASES SEARCH

ORF2 and ORF4 [FIGURE 5.8] have not been analyzed further, as database searches with their predicted amino acid sequences did not show any significant homology to any already described protein. Also in the case of ORF3 work analyzing the predicted amino acid sequence was being done by Claverys *et al* (personal communication). Therefore, protein sequence analysis was concentrated on the neuraminidase ORF (ORF1).

For this protein sequence analysis, the 1035 amino acids, comprising the complete sequence of neuraminidase, were used. The analysis was done by using the University of Wisconsin Molecular Sequence Analysis Software. Several searches were done using different databases. When searching the Prokaryotes database, which is a part of the SWISSPROT protein data base, the option TFASTA was selected. TFASTA uses the method of Pearson & Lipman (1988) to search for similarities between a query peptide sequence and any group of translated nucleotide sequences. TFASTA translates the nucleotide sequence in all six frames before performing the comparison and each translated reading frame is treated as a separate sequence to be searched. A word size of 1 was selected. A word size is the largest number of short perfect matches for each comparison. Therefore, a word size of one would apply the lowest stringency in these searches. Using these options, homology between the pneumococcal neuraminidase and those from *Bacteroides fragilis*, *Clostridium perfringens* and *Clostridium sordellii* was found. This is shown in FIGURES 5.19, 5.20 and 5.21. In the case of *B. fragilis* the homologous region covers 122 amino acids. This includes two of the aspartic boxes found in other neuraminidases and mentioned previously. The neuraminidases from *C. perfringens* and *C. sordellii* cover an overlapping region of homology with the pneumococcal neuraminidase [FIGURE 5.20, 5.21]. The homologous region from *C. perfringens* covers 123 amino acids whereas the region from *C. sordellii* only covers 58 amino acids. In this case, only one aspartic box falls in the region of homology.

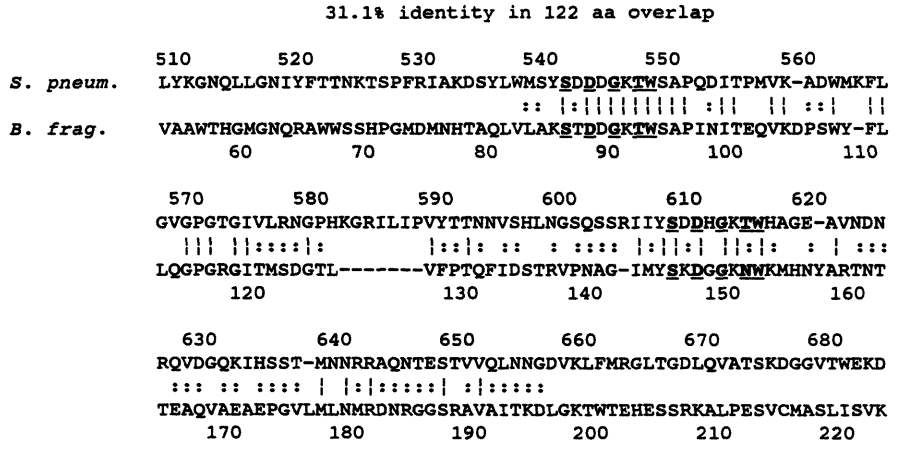


FIGURE 5.19. **Alignment of the homologous region of pneumococcal neuraminidase with the neuraminidase from *Bacteroides fragilis*.**
Underlined bold letters correspond to "aspartic boxes". Numbers represent the position in the amino acid sequence.

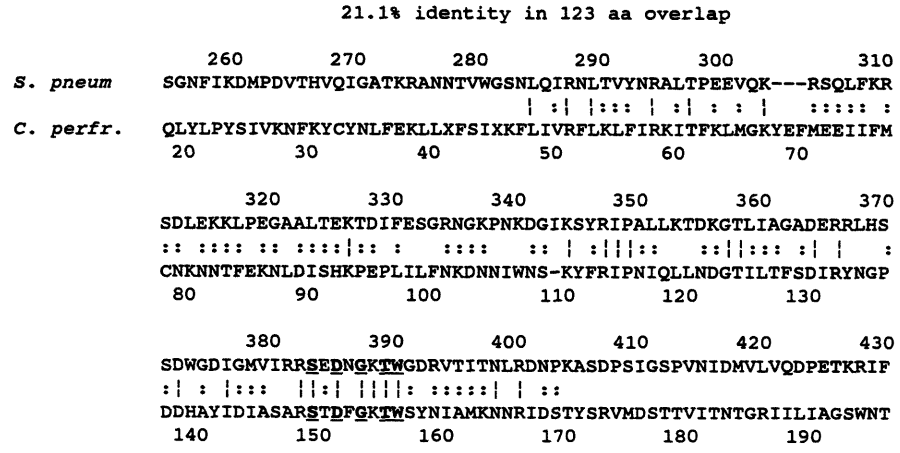


FIGURE 5.20. **Alignment of the homologous region of pneumococcal neuraminidase with the neuraminidase from *Clostridium perfringens*.**
Underlined bold letters correspond to "aspartic boxes". Numbers represent the position in the amino acid sequence.

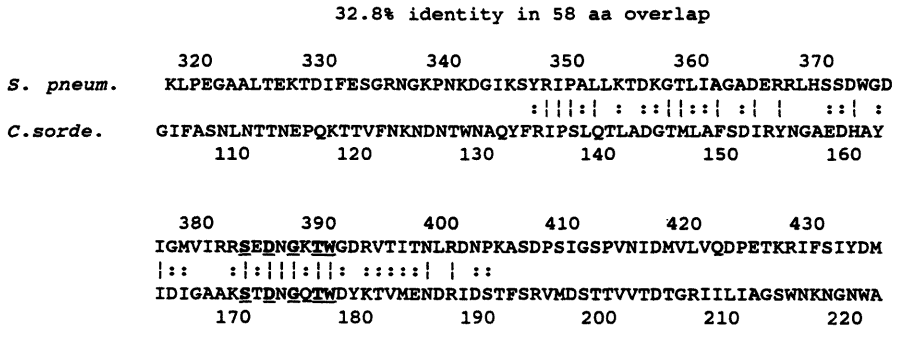


FIGURE 5.21. **Alignment of the homologous region of pneumococcal neuraminidase with the neuraminidase from *Clostridium sordellii*.**
Underlined bold letters correspond to "aspartic boxes". Numbers represent the position in the amino acid sequence.

Finally, the PIR1 protein sequence database was searched using FASTA, with a word size of 1. Interestingly, a small homology (39 amino acids) to the Influenza A virus neuraminidase was found. None of the aspartic boxes were contained in the region of homology [FIGURE 5.22]. Also, the homology did not fall within the active site of the viral neuraminidase [Air & Laver 1989].

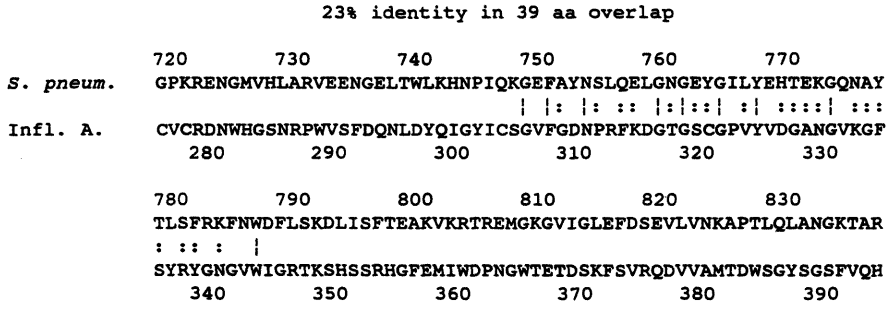


FIGURE 5.22. **Alignment of the homologous region of pneumococcal neuraminidase with the neuraminidase from Influenza A virus.**
Numbers represent the position in the amino acid sequence.

A search of EMBL databases for the presence of bacterial and eukaryotic neuraminidase DNA sequences revealed the presence of the three neuraminidases already mentioned together with the one from *Salmonella typhimurium* LT2, *Vibrio cholerae* and *Trypanosoma cruzi* Y. To search for sequence homologies between the neuraminidase amino acid sequences from all these organisms, an attempt to align all the sequences was made. The program used was CLUSTALV. This program compares amino acid sequences from different proteins aligning the regions of highest homology. It was found that only bacterial neuraminidases could be analyzed. These alignments are shown in FIGURE 5.23. Homology between these proteins was found not only in the aspartic boxes, but also in the region where Hoyer *et al* (1992) found the presence of an invariant glycine from bacterial, viral and eukaryotic sialidases. This region falls in the sequence between the second and third aspartic boxes of bacterial and eukaryotic neuraminidases [FIGURE 5.13][FIGURE 5.23].

CLUSTALV ALIGNMENTS

```
NEURA     MSYFRNRDIDIERNSMNRSVQERKCRYSIRKLSVGAVSMIVGAVVFGTSPVLAQEGASEQ
VIBRIO    MR-FKN----VKKTAL--------------MLAMFGMATSSNAALFDYNATGDTEFDSPA
CSSIAC    MKKF------------------------IKILKVLSMAIVLSACN---------------
CFSIAL    M-----------------------------------------CN---------------
BFNANH    ----------------------------------DVGLSRST---------------
SALMON    M---------------------------------------TVE---------------

NEURA     PLANETQLSGESSTLTDTEKSQPSSETELSGNKQEQERKDKQEEKIPRDYYARDLENVET
VIBRIO    KQGWMQDNTNNGSGVLTNADGMPAWLVQGIGGRAQWTYSLSTNQHAQASSFGWRMTTEMK
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     VIEKEDVETNASNG-QRVDLSSELDKLKKLENATVHMEFKPDAKAPAFYNLFSVSSATKK
VIBRIO    VLSGGMITNYYANGTQRVLPIISLDS-----SGNLVVEFEGQT-----------------
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     DEYFTMAVYNNTATLEGRGSDGKQFYNNYNDAPLKVKPGQWNSVTFTVEKPTAELPKGRV
VIBRIO    -----------GRTVLATGTAATE-YHKFE---LVFLPGSNPSASF--------------
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     RLYVNGVLSRTSLRSGNFIKDMPDVTHVQIGATKRANNTVWGSNLQIRNLTVYNRALTPE
VIBRIO    --YFDGKLIRDNIQPTASKQNMIVWGNGSSNT--DGVAAYRDIKFEIQGDVIFR---GPD
CSSIAC    -------------INGIFASNLNTTNEPQKTTVFNKNDNTWNAQ--------Y-------
CFSIAL    -------------KNNTFEKNLDISHKPEPLILFNKDNNIWNSK--------Y-------
BFNANH    -------------DGG----------------------KTWEK-----------------
SALMON    -------------KSVVFKAEGEHFTDQKGNTIVGSGSGG-TTK--------Y-------

NEURA     EVQKRSQLFKRSDLEKKLPEGAALTEKTDIFESGRNGKPNKDGIKSYRIPALLK--TDKG
VIBRIO    ------------------------------------------------RIPSIVASSVTPG
CSSIAC    ------------------------------------------FRIPSLQT--LADG
CFSIAL    ------------------------------------------FRIPNIQL--LNDG
BFNANH    ------------------------------------------MRLP----------
SALMON    ------------------------------------------FRIPAMCT--TSKG

NEURA     TLIAGADERRLHSSDWGDI---GMVIRRSEDNGKTWGDRVTITNLRDNPKASDPSIGSPV
VIBRIO    VVTAFAEKRVGGGDPGALSNTNDIITRTSRDGGITWDTELNLTEQINVSDEFDFSDPRPI
CSSIAC    TMLAFSDIRYNGAEDHAYI---DIGAAKSTDNGQTWDYKTVMENDR-----IDSTFSRVM
CFSIAL    TILTFSDIRYNGPDDHAYI---DIASARSTDFGKTWSYNIAMKNNR-----IDSTYSRVM
BFNANH    --LAFGEFGGLPAGQNGVG---DPSILVDTKTNNVW---------------VVAAWTHGM
SALMON    TIVVFADARHNTASDQSFI---DTAAARSTDGGKTWNKKIAIYNDR-----VNSKLSRVM
```


NEURA NIDMVLVQDPETKRIFSIYDMFPEGKGIFGMSSQKEEAYKKIDGKTYQ---ILYREGEKG
VIBRIO Y-------DPSSNTVLVSYARWPTD------AAQNGDRIK-----PWMPNGIFY-----S
CSSIAC DSTTVVTDT-GR--IILIAGSWNK----------NGN-W----------
CFSIAL DSTTVITNT-GR--IILIAGSWNT----------NGN-W----------
BFNANH GNQR----------AW----------W----------
SALMON DPTCIVANIQGRETILVMVGKWNN----------NDKTW----------
NEURA AYTIRENGTVYTPDGKATDYRVVVDPVKPAYSDKGDLYKGNQLLGNIYFTTNKTSPFRIA
VIBRIO VYDV----------ASGNWQAPIDVTDQV---KERSFQIAGWGGSELY--RRNTSLNSQ
CSSIAC --------AS-STTSLRSDWSVQ----------
CFSIAL --------AM-TTSTRRSDWSVQ----------
BFNANH --------SSHPGMDMNHTAQLV----------
SALMON --------GAYRDKAPDTDWDLV----------
NEURA KDSYLWMSYSDDD-GKTWSAPQDIT----------PMVKADWMKFLGVGPGTGIVLRNGP
VIBRIO QD---WQSNAKIRIVDGAANQIQVADGSRKYVVTLSIDESGGLVANLNGVSAPIILQSEH
CSSIAC ------MVYSDDN-GETWSDKVDLTTNKAR-IKNQPSNTIGW----LAGVGSGIVMSD--
CFSIAL ------MIYSDDN-GLTWSNKIDLTKDSSK-VKNQPSNTIGW----LGGVGSGIVMDD--
BFNANH ------LAKSTDD-GKTWSAPINITEQ----VKDPS-----WY-FLLQGPGRGITMSD--
SALMON ------LYKSTDD-GVTFS-KVETNIHDIV-TKNGTISAM------LGGVGSGLQLND--
NEURA HKGRILIPVYTTNNVSHLNGSQSSRIIYSDDHGKT-WHAGEAVNDNRQVDGQKIHSSTMN
VIBRIO AKVHS-FHDYEL----QYSALNHTTTLFVDGQQITTWAGEVSQENNIQFGNADAQID---
CSSIAC --GTIVMPI-QIALRENNANNYYSSVIYSKDNGET-WTMGNKVPD----------
CFSIAL --GTIVMPA-QISLRENNENNYYSLIIYSKDNGET-WTMGNKVPN----------
BFNANH --GTLVFPT-QFI---DSTRVPNAGIMYSKDGGKN-WKM-HNYAR----------
SALMON --GKLVFPV-QMVRTKNITTVLNTSFIYSTD-GIT-WSLPSGYCE----------
NEURA NRRAQNTESTVVQLNNGDVKLFMRGLTGDLQVA-TSKDGGV-TWEKDIKRYPQVKDVYVQ
VIBRIO ---GRLHVQKIVLTQQGHNLVEFDAFYLAQQTPEVEKDLEKLGWTK-IKTGNTM-SLYGN
CSSIAC ---PKTSENMVIELD-GALIMSSRNDGKNYRASYISYDMGS-TWE----VYDPLHN----
CFSIAL ---SNTSENMVIELD-GALIMSTRYDYSGYRAAYISHDLGT-TWE----IYEPLNG----
BFNANH ---TNTTEAQVAEAEPGVLMLNMRDNRGGSRAVAITKDLGK-TWT----EHESSRK----
SALMON ---GFGSENNIIEFN-ASLVNNIRN--SGLRRSFETKDFGK-TWT----EFPPMDK----
NEURA MSAIHTMHEGKEYIILSNAGGPKRENGMVHLARVEENGELTWLKHNPIQKGEFAYNSLQE
VIBRIO AS----------VNPGPGHG-----ITLTRQQN------ISGSQNGRLIYPAIVLDR
CSSIAC ----------KISTGNGSGCQGSFIKVTAKDGHRL-GFISAPKNTKGGYVRDNI--
CFSIAL ----------KILTGKGSGCQGSFIKATTSNGHRI-GLISAPKNTKGEYIRDNI--
BFNANH ----------AL---PESVCMASLISVKAKD----------
SALMON ----------KV-DNRNHGVQGS--TITIPSGNKLVAAHSSAQNKNNDYTRSDI--
NEURA LGNGEYGI-LYEHTEKGQNAYTLSFR-KFNWDFLSKDLISPTEAKVKRTREMGKGVIGLE
VIBRIO F--FLNVMSIYSDDGGSNWQTGSTLPIPFRWK---------SSSILETLEPSEADMVELQ
CSSIAC ---TVYMID-FDDLSKGIRELCSPYP-E----------DGNSSGGGYSCLS
CFSIAL ---AVYMID-FDDLSKGVQEICIPYP-E----------DGNKLGGGYSCLS
BFNANH ----------NVLGKDLLIFSNPNT-T----------KG----------R
SALMON ---SLYAHNLYSGEVKLIDAF---YP-K----------VGNASGAGYSCLS

CLUSTALV ALIGNMENTS

```
NEURA     MSYFRNRDIDIERNSMNRSVQERKCRYSIRKLSVGAVSMIVGAVVFGTSPVLAQEGASEQ
VIBRIO    MR-FKN----VKKTAL--------------MLAMFGMATSSNAALFDYNATGDTEFDSPA
CSSIAC    MKKF------------------------IKILKVLSMAIVLSACN---------------
CFSIAL    M---------------------------------------CN---------------
BFNANH    ------------------------------------DVGLSRST---------------
SALMON    M---------------------------------------TVE---------------

NEURA     PLANETQLSGESSTLTDTEKSQPSSETELSGNKQEQERKDKQEEKIPRDYYARDLENVET
VIBRIO    KQGWMQDNTNNGSGVLTNADGMPAWLVQGIGGRAQWTYSLSTNQHAQASSFGWRMTTEMK
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     VIEKEDVETNASNG-QRVDLSSELDKLKKLENATVHMEFKPDAKAPAFYNLFSVSSATKK
VIBRIO    VLSGGMITNYYANGTQRVLPIISLDS-----SGNLVVEFEGQT-----------------
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     DEYFTMAVYNNTATLEGRGSDGKQFYNNYNDAPLKVKPGQWNSVTFTVEKPTAELPKGRV
VIBRIO    -----------GRTVLATGTAATE-YHKFE---LVFLPGSNPSASF--------------
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     RLYVNGVLSRTSLRSGNFIKDMPDVTHVQIGATKRANNTVWGSNLQIRNLTVYNRALTPE
VIBRIO    --YFDGKLIRDNIQPTASKQNMIVWGNGSSNT--DGVAAYRDIKFEIQGDVIFR---GPD
CSSIAC    -------------INGIFASNLNTTNEPQKTTVFNKNDNTWNAQ--------Y--------
CFSIAL    -------------KNNTFEKNLDISHKPEPLILFNKDNNIWNSK--------Y--------
BFNANH    -------------DGG---------------------KTWEK-----------------
SALMON    -------------KSVVFKAEGEHFTDQKGNTIVGSGSGG-TTK--------Y--------

NEURA     EVQKRSQLFKRSDLEKKLPEGAALTEKTDIFESGRNGKPNKDGIKSYRIPALLK--TDKG
VIBRIO    ---------------------------------------------RIPSIVASSVTPG
CSSIAC    ---------------------------------------------FRIPSLQT--LADG
CFSIAL    ---------------------------------------------FRIPNIQL--LNDG
BFNANH    ---------------------------------------------MRLP----------
SALMON    ---------------------------------------------FRIPAMCT--TSKG
                                                          *.*

NEURA     TLIAGADERRLHSSDWGDI---GMVIRRSEDNGKTWGDRVTITNLRDNPKASDPSIGSPV
VIBRIO    VVTAFAEKRVGGGDPGALSNTNDIITRTSRDGGITWDTELNLTEQINVSDEFDFSDPRPI
CSSIAC    TMLAFSDIRYNGAEDHAYI---DIGAAKSTDNGQTWDYKTVMENDR-----IDSTFSRVM
CFSIAL    TILTFSDIRYNGPDDHAYI---DIASARSTDFGKTWSYNIAMKNNR-----IDSTYSRVM
BFNANH    --LAFGEFGGLPAGQNGVG---DPSILVDTKTNNVW--------------VVAAWTHGM
SALMON    TIVVFADARHNTASDQSFI---DTAAARSTDGGKTWNKKIAIYNDR-----VNSKLSRVM
            ..        .              *                        .
```

```
NEURA    NIDMVLVQDPETKRIFSIYDMFPEGKGIFGMSSQKEEAYKKIDGKTYQ---ILYREGEKG
VIBRIO   Y-------DPSSNTVLVSYARWPTD------AAQNGDRIK-----PWMPNGIFY-----S
CSSIAC   DSTTVVTDT-GR--IILIAGSWNK-----------------NGN-W-------------
CFSIAL   DSTTVITNT-GR--IILIAGSWNT-----------------NGN-W-------------
BFNANH   GNQR----------------AW----------------------W-------------
SALMON   DPTCIVANIQGRETILVMVGKWNN-----------------NDKTW-------------

NEURA    AYTIRENGTVYTPDGKATDYRVVVDPVKPAYSDKGDLYKGNQLLGNIYFTTNKTSPFRIA
VIBRIO   VYDV-----------ASGNWQAPIDVTDQV---KERSFQIAGWGGSELY--RRNTSLNSQ
CSSIAC   --------AS-STTSLRSDWSVQ-------------------------------------
CFSIAL   --------AM-TTSTRRSDWSVQ-------------------------------------
BFNANH   --------SSHPGMDMNHTAQLV-------------------------------------
SALMON   --------GAYRDKAPDTDWDLV-------------------------------------

NEURA    KDSYLWMSYSDDD-GKTWSAPQDIT----------PMVKADWMKFLGVGPGTGIVLRNGP
VIBRIO   QD---WQSNAKIRIVDGAANQIQVADGSRKYVVTLSIDESGGLVANLNGVSAPIILQSEH
CSSIAC   ------MVYSDDN-GETWSDKVDLTTNKAR-IKNQPSNTIGW----LAGVGSGIVMSD--
CFSIAL   ------MIYSDDN-GLTWSNKIDLTKDSSK-VKNQPSNTIGW----LGGVGSGIVMDD--
BFNANH   ------LAKSTDD-GKTWSAPINITEQ----VKDPS-----WY-FLLQGPGRGITMSD--
SALMON   ------LYKSTDD-GVTFS-KVETNIHDIV-TKNGTISAM------LGGVGSGLQLND--

NEURA    HKGRILIPVYTTNNVSHLNGSQSSRIIYSDDHGKT-WHAGEAVNDNRQVDGQKIHSSTMN
VIBRIO   AKVHS-FHDYEL----QYSALNHTTTLFVDGQQITTWAGEVSQENNIQFGNADAQID---
CSSIAC   --GTIVMPI-QIALRENNANNYYSSVIYSKDNGET-WTMGNKVPD---------------
CFSIAL   --GTIVMPA-QISLRENNENNYYSLIIYSKDNGET-WTMGNKVPN---------------
BFNANH   --GTLVFPT-QFI---DSTRVPNAGIMYSKDGGKN-WKM-HNYAR---------------
SALMON   --GKLVFPV-QMVRTKNITTVLNTSFIYSTD-GIT-WSLPSGYCE---------------

NEURA    NRRAQNTESTVVQLNNGDVKLFMRGLTGDLQVA-TSKDGGV-TWEKDIKRYPQVKDVYVQ
VIBRIO   ---GRLHVQKIVLTQQGHNLVEFDAFYLAQQTPEVEKDLEKLGWTK-IKTGNTM-SLYGN
CSSIAC   ---PKTSENMVIELD-GALIMSSRNDGKNYRASYISYDMGS-TWE----VYDPLHN----
CFSIAL   ---SNTSENMVIELD-GALIMSTRYDYSGYRAAYISHDLGT-TWE----IYEPLNG----
BFNANH   ---TNTTEAQVAEAEPGVLMLNMRDNRGGSRAVAITKDLGK-TWT----EHESSRK----
SALMON   ---GFGSENNIIEFN-ASLVNNIRN--SGLRRSFETKDFGK-TWT----EFPPMDK----

NEURA    MSAIHTMHEGKEYIILSNAGGPKRENGMVHLARVEENGELTWLKHNPIQKGEFAYNSLQE
VIBRIO   AS-------------VNPGPGHG-----ITLTRQQN------ISGSQNGRLIYPAIVLDR
CSSIAC   --------------KISTGNGSGCQGSFIKVTAKDGHRL-GFISAPKNTKGGYVRDNI--
CFSIAL   --------------KILTGKGSGCQGSFIKATTSNGHRI-GLISAPKNTKGEYIRDNI--
BFNANH   --------------AL---PESVCMASLISVKAKD-------------------------
SALMON   --------------KV-DNRNHGVQGS--TITIPSGNKLVAAHSSAQNKNNDYTRSDI--

NEURA    LGNGEYGI-LYEHTEKGQNAYTLSFR-KFNWDFLSKDLISPTEAKVKRTREMGKGVIGLE
VIBRIO   F--FLNVMSIYSDDGGSNWQTGSTLPIPFRWK---------SSSILETLEPSEADMVELQ
CSSIAC   ---TVYMID-FDDLSKGIRELCSPYP-E-------------------DGNSSGGGYSCLS
CFSIAL   ---AVYMID-FDDLSKGVQEICIPYP-E-------------------DGNKLGGGYSCLS
BFNANH   -----------NVLGKDLLIFSNPNT-T--------------------KG----------R
SALMON   ---SLYAHNLYSGEVKLIDAF---YP-K-------------------VGNASGAGYSCLS
```

```
NEURA     FDSEVLVNKAPTLQLANGKTARFMTQYDTKTLLFTVDSEDMGQKVTGLAEGAIESMHNLP
VIBRIO    NGDLLLTARLDFNQIVNG------VNYSPRQQFLSKDGGITWSLLEANNANVFSNISTGT
CSSIAC    FN------------------------------------------------------------
CFSIAL    FK------------------------------------------------------------
BFNANH    YNTTI---------------------------------------------------------
SALMON    YRKNV---------------------------------------------------------

NEURA     VSVAGTKLSNGMNGSEAAVHEVPEYTGPLGTSGEEPAPTVEKPEYTGPLGTSGEEPAPTV
VIBRIO    VDASITRFEQS-DGSHFLL-----FTNPQGN----------------PAGTNGRQN----
CSSIAC    ------------------------------------------------------------
CFSIAL    ------------------------------------------------------------
BFNANH    ------------------------------------------------------------
SALMON    ------------------------------------------------------------

NEURA     EKPEYTGPLGTAGEEAAPTVEKPEFTGGV--NGTEPAVHEIAEYKGSDSLVTLTTKEDYT
VIBRIO    --------LGL----------WFSFDEGVTWKGPIQLVNGASAYSD---IYQLDSENAI-
CSSIAC    ----------------------------------------------DGKLSIL-YEANGNIE
CFSIAL    ----------------------------------------------NNHLGIV-YEANGNIE
BFNANH    ----------------------------------------------KISL-------DGGVT
SALMON    ----------------------------------------------DKKHCMLSMKPMEVLS

NEURA     YKAPLAQQALPETGNKESDLLASLGLTAFFLGLFTLGKKREQ
VIBRIO    --------VIVETDNSNMRIL-RMPITLLKQKL-TLSQN---
CSSIAC    YKD----------------------LTDYYLSIEN--NKKLK
CFSIAL    YQD----------------------LTPYYSLINK--Q----
BFNANH    WSP---------------------------------EHQL-
SALMON    SRT----------------------LAVIY----------Q
```

FIGURE 5.23. CLUSTALV amino acid sequence alignment of bacterial neuraminidases.

Codes represent neuraminidases from: NEURA (*S. pneumoniae*); VIBRIO (*V. cholerae*); CSSIAC (*C. sordellii*); CFSIAL (*C. perfringens*); BFNANH (*B. fragilis*); SALMON (*S. typhimurium*). The symbol "*" represents identities, whereas "." represents conservative replacements between the aligned sequences. Single underlined sequences represent regions corresponding to the Aspartic Box. Double underlines identify a conserved region present in the highly conserved glycine residue (indicated by "▼"), found in neuraminidases from bacteria, eukaryotes and viruses [Hoyer *et al* 1992]

5.6. LOCATION OF NEURAMINIDASE WITHIN THE PNEUMOCOCCUS

Analysis of the predicted pneumococcal neuraminidase sequence showed that this enzyme had a putative signal peptide sequence. This indicates that this enzyme might be secreted from the cell. Further analysis has shown that in the C-terminus of this protein there is present a typical Gram-positive surface anchor motif. Although previous reports showed that neuraminidase is actively secreted from the pneumococcus during the logarithmic phase of growth [Lee *et al* 1966], knowing that there are two distinct neuraminidase genes in *S. pneumoniae*, a more detailed analysis should be done to determine whether one of the two neuraminidases is anchored, at any stage of growth, to the surface of the pneumococcus. For this purpose a cell fractionation experiment was needed to isolate pneumococcal cell membranes that could be tested for neuraminidase activity.

Before cell fractionation experiments were done, a growth curve of the pneumococcus was done to determine the optimal time point for harvesting the largest number of intact pneumococcal cells before autolysis occurred.

5.6.1. *Streptococcus pneumoniae* GROWTH CURVE.

The growth curve of the pneumococcal strain GB05B type III was done as it produces a large amount of neuraminidase and also has both neuraminidase genes. Samples were taken every hour for 12 hours. Both culture supernatant and cell-associated neuraminidase was measured. Bacterial numbers (CFU) were taken at every time point. The results are shown in TABLE 5.2.

At nine hours of growth, the number of CFU reached the maximun. At that point neuraminidase activity was maximum, correlating with the end of the stationay phase. Therefore, to obtain the largest amount of membranes from intact cells, bacteria needed to be harvested after nine hours of growth.

	NEURAMINIDASE ACTIVITY		
Time point	Supernatant	Pellet	CFU.ml^{-1}
T = 0	0	0	3.0 x 10^5
T = 1	0	0	4.7 x 10^5
T = 2	0	0	2.9 x 10^6
T = 3	0	0	8.7 x 10^6
T = 4	0	0	5.6 x 10^7
T = 5	395	475	1.7 x 10^8
T = 6	430	1670	5.2 x 10^8
T = 7	3230	25080	7.5 x 10^8
T = 8	3890	45210	9.6 x 10^8
T = 9	3880	47295	1.0 x 10^9
T = 10	3735	49220	7.2 x 10^8
T = 11	4275	47550	5.4 x 10^8
T = 12	5050	46565	4.3 x 10^8

TABLE 5.2. **Results from a growth curve of GB05B.**
Samples were taken every hour for 12 hours. At every time point culture supernatants and cell pellets were assayed for neuraminidase activity (expressed in fluorimeter units per ml of culture). Colony forming units (CFU) were also measured at every time point.

5.6.2. PNEUMOCOCCAL CELL FRACTIONATION

S. pneumoniae GB05B was grown for nine hours, at which point the CFU.ml^{-1} count was 4.2 x 10^9. Cell fractionation was carried out as described in Materials and Methods. In summary, cells were harvested, washed and disrupted. The disrupted cell preparation was centrifuged and the supernatant was centrifuged twice more until no visible pellet of unbroken cells was found. The final supernatant was centrifuged at high speed to collect the cell membranes which were washed three times. The following samples were assayed for neuraminidase activity :

1. The culture supernatant.
2. The two supernatants from the cell pellet washes.
3. The supernatants and pellets from the centrifugations after cell disruption.
4. The supernatants and membrane pellets from the high speed spin of the broken cell suspensions and the three subsequent washes.

These results are shown in TABLE 5.3.

FRACTION	NEURAMINIDASE ACTIVITY
Culture supernatant	13,000
Supernatant 1st PBS wash	2,130
Supernatant 2nd PBS wash	5
1st supernatant 5K	48,422
1st pellet 5K	2,321
2nd supernatant 5K	44,385
2nd pellet 5K	437
1st supernatant 45K	42,204
1st pellet 45K	1,488
2nd supernatant 45K	1,457
2nd pellet 45K	58
3rd supernatant 45K	22
3rd pellet 45K	36
4th supernatant	0
MEMBRANES	18

TABLE 5.3. Results from neuraminidase assays of all the various stages in the fractionation of the pneumococcal cells. The neuraminidase activity is in fluorimeter units (FU) per ml of original culture.

Very little neuraminidase activity was obtained from the pneumococcal cell membranes after the cell fractionation. It was not clear if such a low activity was due to the background of the assay. Most of the neuraminidase was released into the supernatant of the broken cells. This suggests that, either neuraminidase was present in the cytoplasm, or else, it was associated, to some extent, to the cell surface but released during the cell fractionation procedure.

To distinguish between these possibilities an alterative experiment, using electron microscopy of immunogold stained bacteria, was designed to study the location of the pneumococcal neuraminidase.

5.6.3. ELECTRON MICROSCOPY STUDIES

To study the subcellular location of neuraminidase, an electron microscopy experiment was designed. For this experiment an immunogold staining of the pneumococcus using polyclonal anti-neuraminidase antibodies was used. The polyclonal antibodies used were raised against a purified neuraminidase from the pneumococcus [Lock *et al* 1988]. These antibodies react with the cloned neuraminidases from LambdaEMBL301-Neu1 (see Part IV) and pJCP301 [Berry *et al* 1988]. The pneumococcal strain used was GB05B. Since a neuraminidase-negative mutant of pneumococcus does not exist, another bacterium, lacking this enzyme activity, was used as a negative control for the immunogold staining. It is acknowledged that this is an important control. The bacterium chosen was *E. coli* JM101. As a positive control for this staining *S. pneumoniae* GB05B was used and stained with polyclonal anti-type III capsular antibodies. The *E. coli* strain used was harvested from an overnight culture whereas GB05B was collected from a nine hour culture.

The electron micrographs from these experiments are shown in FIGURES 5.23 to 5.25. The *E. coli* negative control did not show any reaction with antibody by the anti-neuraminidase antibodies [FIGURE 5.23]. In contrast, specific staining of pneumococcal capsule from the strain GB05B was detected when anti-capsular antibodies were used [FIGURE 5.24]. Staining of the pneumococcal strain GB05B with anti-neuraminidase antibodies revealed recognition of the surface of the pneumococcus, with only very little staining of the cytoplasm [FIGURE 5.25].

No conclusions should be taken from these results due to the lack of an appropriate negative control which for this experiment is essential. However, although there may be more cytoplasmic neuraminidase that has not been detected by the method used, the reaction of the anti-neuraminidase antibodies with the surface of the pneumococcus [FIGURE 5.25] is consistent with a surface localization for neuraminidase.

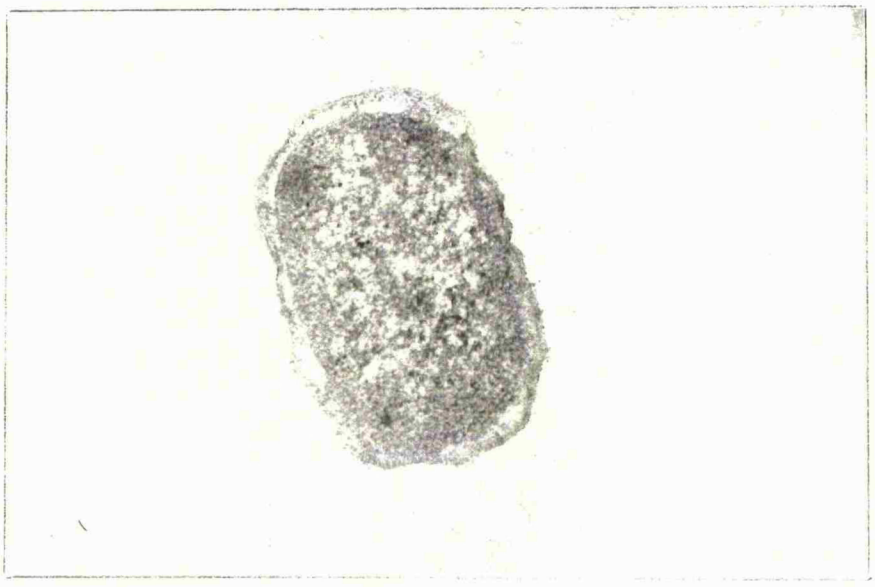


FIGURE 5.23. **Electron microscopy of immunogold stained *E. coli* with anti-neuraminidase antibodies.**

From 100 cells examined none showed immunogold staining.

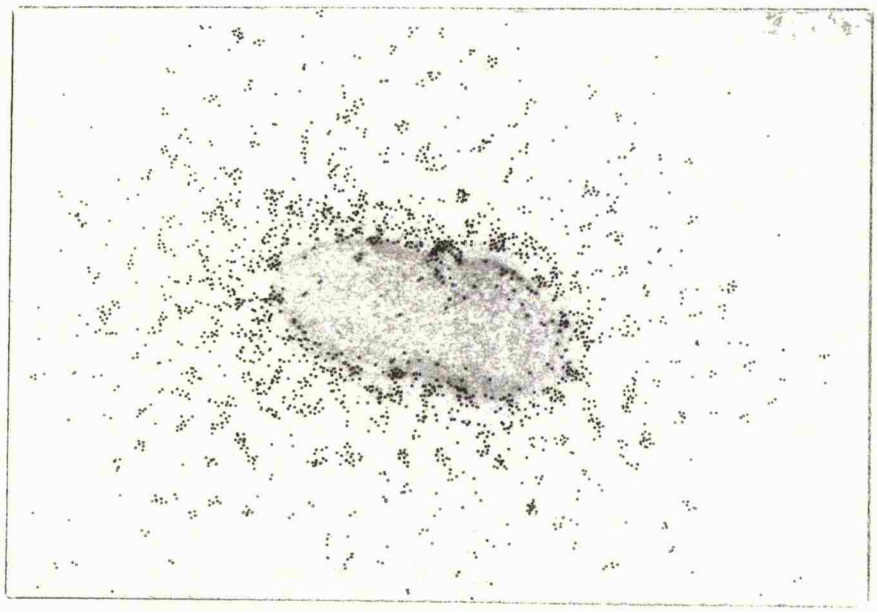


FIGURE 5.24. **Electron microscopy of immunogold staining of *S. pneumoniae* GBO5B using anti-capsule antibodies.**

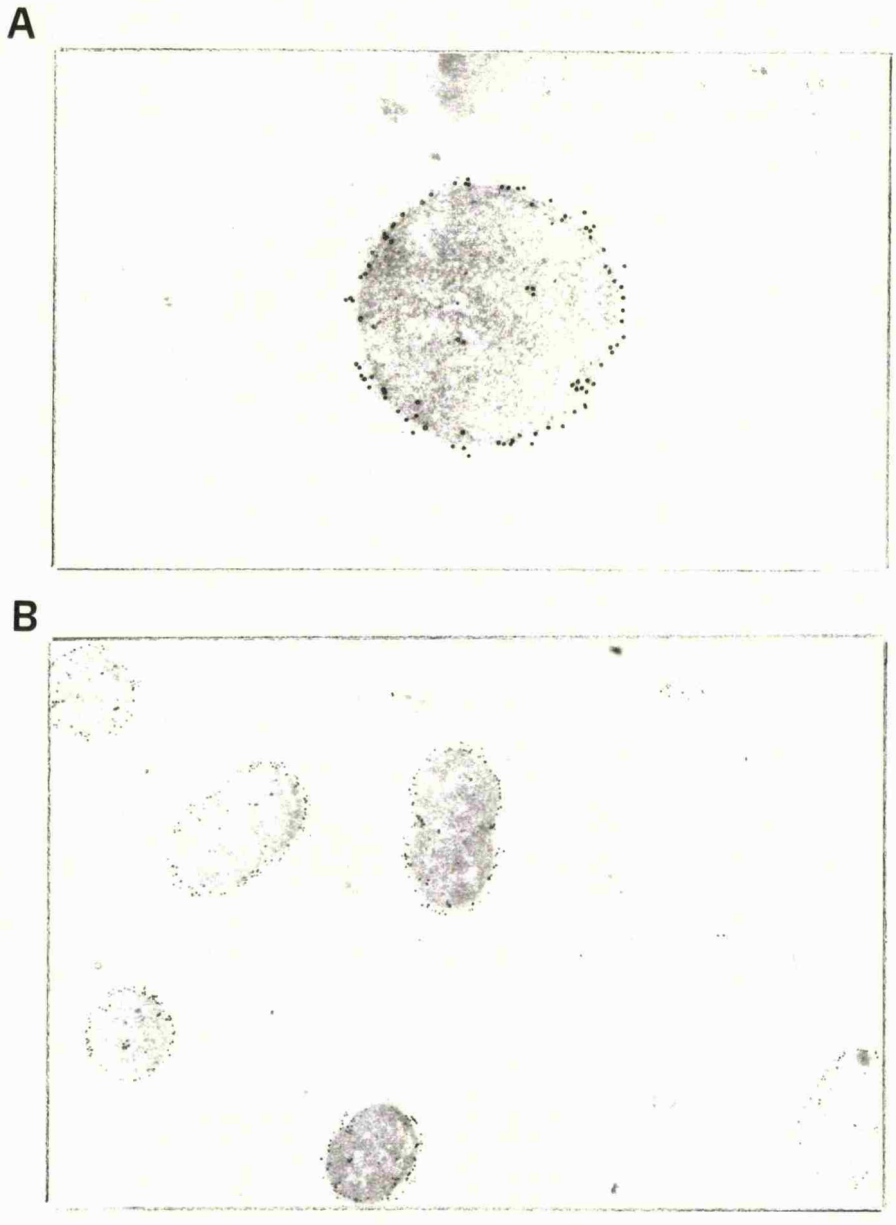


FIGURE 5.25. **Electron microscopy of *S. pneumoniae* GB05B immunogold stained with anti-neuraminidase antibodies.**

A. A single bacterium with antibody bound at the cell surface

B. Several pneumococci with the presence of anti-neuraminidiase antibodies at the cell surface.

PART VI

DISCUSSION

6.1. CONSTRUCTION OF A LambdaEMBL301 PNEUMOCOCCAL GENOMIC LIBRARY

Streptococcus pneumoniae remains one of the major causes of bacterial pneumonia and meningitis in man. It produces several factors that may be of relevance during pneumococcal infection, including IgA1 protease and neuraminidase enzymes. Attempts to clone the genes coding for these enzymes have been made previously with little success [Berry *et al* 1988; Pratt & Boulnois 1987]. The main problem found when cloning these genes was the instability of the recombinant DNA encoding for these activities [Berry *et al* 1988; Pratt & Boulnois 1987]. Both IgA1 protease and neuraminidase genes, when cloned in cosmid vectors were unstable. Similar problems were found during attempts to clone the gene for IgA1 protease from a plasmid-based pneumococcal genomic library in *E. coli* [Pratt 1988]. Lambda vectors have been used successfully to clone pneumococcal genes where plasmid vectors had not [Walker *et al* 1987]. Also, cloning with Lambda vectors has the advantage over cosmid and plasmid vectors that proteins in the host cell will be released by phage lysis and secretion will be no longer required. Therefore, it was decided to attempt to select the genes for IgA1 protease and neuraminidase from a lambda pneumococcal genomic library. LambdaEMBL301 [Lathe *et al* 1987] was the vector of choice as it allows cloning of large DNA fragments and, having a large polylinker, offers a greater range of possibilities for the subcloning of recombinant DNA.

Normally when constructing genomic DNA libraries size fractionation of chromosomal DNA partial digests is included prior to ligation into the vector. However, partially digested chromosomal DNA from the pneumococcus could not be fractionated on sucrose gradients as only small amounts of DNA, which could not be ligated to the vector, were recovered using this technique. Also, size selection using low melting point agarose could not be used as it involves mechanical shearing that can break large DNA fragments [Langridge *et al* 1980]. Partially digested DNA was therefore cloned directly, without size selection, as it then gave a high ligation efficiency. Although this could cause the adjacent cloning of DNA sequences from distant parts of the chromosome, Southern blot analysis with

different restriction enzyme digests of both chromosomal and cloned DNA would discern between contiguous and rearranged DNA sequences in the chromosome.

6.2. CLONING OF AN IgA1 PROTEASE GENE FROM *S. pneumoniae*

Some of the IgA1 protease genes already cloned have been shown to express a very low level of the enzyme in *E. coli* when compared to the parental strains [Koomey *et al* 1982; Bricker *et al* 1983]. Therefore, in screening for expression of the enzyme, a very sensitive method would be required.

The most ideal screening protocol would involve the use of DNA probes, as this would not depend on the expression of the recombinant protein. Unfortunately, no homology by DNA hybridization has been detected between the *iga* gene from *S. sanguis* and the chromosome of *S. pneumoniae* [Gilbert *et al* 1988]. Lack of homology has also been found between the IgA1 protease gene from *S. sanguis* and those from the gonococcus or *H. influenzae*. Therefore, it would have been unlikely to find a DNA probe from any of the cloned *iga* genes that would hybridize with the gene from the pneumococcus. Consequently, screening for IgA1 protease expression was required.

Several techniques have been developed to test for the activity of this enzyme although not all of them are suitable for detection of small amounts of the protease in screenings involving large number of recombinants. Mehta *et al* (1973) developed an immunoelectrophoresis technique for the detection of IgA1 protease activity. A very sensitive method, involving high performance gel-permeation chromatography, was used by Mortensen & Kilian (1984). This method, although very sensitive, was not ideal for testing large number of samples. Other procedures such as the immunoblotting technique developed by Tsuji *et al* (1989) were not suitable for either detecting very low enzyme activity or studying large numbers of samples. Since none of these methods were appropriate for the screening of the LambdaEMBL301 pneumococcal genomic library, a technique developed by Gilbert & Plaut (1983), consisting of a colony overlay of ^{125}I-IgA1 bound to anti-IgA1

polyacrylamide beads suspended in agar was attempted. This method was unsuccessful due to large background signals, possibly corresponding to diffusion of the beads, carrying the labelled IgA1, in the agar. A modification of the method used by Mulks *et al* (1982) was therefore adopted. This extremely sensitive method consists in detection of IgA1 protease activity by SDS-polyacrylamide gel electrophoresis of ^{125}I-labelled human IgA1 that had been incubated with the samples to be tested. To avoid the hazard of ^{125}I, ^{35}S was used as the radioisotope and, to decrease the number of samples to be screened, the recombinants were assayed in pools of five. The recombinant LambdaEMBL301-iga1, isolated from one of the pools, expressed IgA1 protease activity. To confirm this activity as specific for IgA1 other immunoglobulins should be tested for protease cleavage. No further studies were carried out with LambdaEMBL301-iga1 as analysis of the neuraminidase recombinant (LambdaEMBL301-Neu1) was in progress. However, future work with LambdaEMBL301-iga1 will be orientated towards the understanding of the role that IgA1 protease plays in the pneumococcus. Determination of pneumococcal IgA1 protease predicted amino acid sequence would allow comparisons between this enzyme and other IgA1 proteases from different organisms. These studies could reveal useful information, such as the location of regions of homology amongst these enzymes, that could contribute to the understanding of the structure and function of the pneumococcal IgA1 protease. This information will be required, once a suitable animal model for the study of this enzyme *in vivo* is found, to understand the mechanism of action of IgA1 protease during pneumococcal infection.

6.3. CLONING OF A NEURAMINIDASE GENE FROM *S. pneumoniae*

A recombinant expressing neuraminidase activity was isolated from a pneumococcal cosmid library in *E. coli* by Berry *et al* (1988). This neuraminidase gene was cloned from a cosmid pneumococcal genomic library in *E. coli*. This clone suffered a spontaneous deletion, after which the insert DNA was reduced to 3 Kb. The deleted recombinant named pJCP301, expressed very low neuraminidase activity (0.6% of that from the parental pneumococcal strain) and sometimes this activity was lost. It was also found

that this clone corresponded to a contiguous DNA fragment in the chromosome.

From the screening of the LambdaEMBL301 pneumococcal genomic library, a stable recombinant expressing neuraminidase activity was isolated. For the subcloning of the recombinant neuraminidase gene a suitable vector had to be found in order to avoid instability problems. It has been shown that there are many pneumococcal DNA sequences that closely resemble the canonical *E. coli* promoters. These sequences occur frequently in the pneumococcal chromosome as it has an average A-T content of 61% [Marmur & Doty 1962]. Some of them have shown strong promoter activity in *E. coli* [Morrison & Jaurin 1990]. Such an activity could interfere with plasmid stability [Stassi & Lacks 1982]. To overcome this problem a vector containing bidirectional transcription terminators was designed for the cloning of pneumococcal DNA [Chen & Morrison 1987]. This vector named pJDC9 was used for the subcloning of the neuraminidase gene contained in LambdaEMBL301-Neu1. Recent studies have revealed that removal of transcription terminators from some cloning vectors results in reduced copy number for the strongest pneumococcal promoter-acting sequences but not in altered plasmid stabilities [Dillard & Yother 1991]. Therefore, the instability problems observed when cloning pneumococcal DNA may be related to other problems different from promoter strength.

From the neuraminidase subcloning the recombinant with the smallest insert expressing the highest enzyme activity (pMC2150) was selected for study. The restriction map of the insert was different to that of the neuraminidase-expressing clone pJCP301 [Berry *et al* (1988)]. As these two clones were originated from two different pneumococcal types it was possible that there were more than one neuraminidase gene and that the distribution of these genes was type specific. Also, there could be a single neuraminidase gene showing a polymorphic DNA restriction map depending on the pneumococcal type. Finally, at least two distinct pneumococcal neuraminidase genes could be simultaneously present in different serotypes. To distinguish between these possibilities DNA hybridization studies were done.

Dot blot and Southern blot hybridizations showed that the recombinant DNA from pMC2150 and pJCP301 were present in the pneumococcus. The distribution of this cloned DNA was shown to be serotype independent although to confirm this, a large number of the 83 different pneumococcal capsular serotypes described [Lee *et al* 1991] should be examined. This type of study was outside the scope of this thesis.

The purified neuraminidase from Lock *et al* (1988) was used to raise polyclonal antibodies. These antibodies reacted in Western blot from SDS-PAGE gels with the products pJCP301 [Berry *et al* 1988] and Lambda EMBL301-Neu1. These results indicate that although no homology has been seen in DNA hybridization studies between these two recombinants, their products might share some homology. This could be expected as both clones express the same enzyme activity. In contrast, the purified neuraminidase could contain a mixture of the enzymes from pJCP301 and LambdaEMBL301-Neu1, especially as they may have similar molecular weights. In this case, the differences in molecular weight between the two neuraminidases would have not been detected in a 10% SDS-PAGE gel which was the test for homogeneity of the enzyme purification. Therefore, the cross-reaction observed in Western blot between the products of the two recombinants could be due to the presence of antibodies raised against both enzymes.

The recombinant pJCP301 was derived by spontaneous deletion of a cosmid clone and lacked the full neuraminidase gene although it expressed some enzyme activity [Berry *et al* 1988]. The cloning of the full length gene from the LambdaEMBL301 library, using an internal fragment from pJCP301 as a probe, was attempted. Two identical recombinants (LambdaEMBL301-Neu2 and LambdaEMBL310-Neu3) hybridized with this probe but failed to express neuraminidase activity. When this gene was cloned it was suggested that the expression of this enzyme in pJCP301 was due to an in-frame fusion of the distal end of the neuraminidase gene with the proximal end of the vector *tet* gene, therefore, it would not be expressed from its own promoter [Berry *et al* 1988]. The lack of neuraminidase activity in LambdaEMBL301-Neu2 and

Lambda EMBL301-Neu3 might have been due to the absence of either promoter sequences or the full length neuraminidase gene in these clones. Also, if the promoter sequences have been cloned they may not be active in *E. coli*.

For many years there has been a dispute about the existence of more than one neuraminidase in *S. pneumoniae*. Early purification attempts of this protein from the pneumococcus suggested the presence of several neuraminidase isoenzymes in this organism (see introduction). In contrast, Lock *et al* (1988) speculated that these isoenzymes could be breakdown products from a single neuraminidase as this enzyme is very susceptible to proteolytic degradation. For the first time, this thesis raises the possibility of the presence of more than one neuraminidase gene in the pneumococcus.

These neuraminidases may have different substrate specificities, as in the case of the purified neuraminidases from *Arthrobacter ureafaciens* [Uchida *et al* 1979] and *Streptococcus viridans* II [von Nicolai *et al* 1988]. They may also have different subcellular locations in the pneumococcus. One of them could be secreted, being responsible for attachment of the pneumococcus to specific receptors in the mucosal surfaces that could be exposed after neuraminidase treatment [Anderson *et al* 1983]. The other neuraminidase could be cytoplasmic being released by autolysis and playing a different role. Also, as expression of neuraminidase has been shown to be induced by free and conjugated sialic acids [Kelly *et al* 1966; Kelly & Greiff 1970], each neuraminidase gene may be expressed at different stages of pneumococcal infection by specific sialic acids present in the multiple environments where the pneumococcus is found during the various stages of infection.

The possibility of differential expression of more than one pneumococcal neuraminidase and for different substrate specificities may be important in the various phases of pneumococcal infection. To study the role of each neuraminidase in the pneumococcus, the construction of specific neuraminidase- negative isogenic mutants would be required. However, the presence of additional neuraminidase genes in the pneumococcus cannot be

dismissed, and would further complicate the study of isogenic mutants.

6.4. DNA SEQUENCE DETERMINATION AND ANALYSIS OF THE NEURAMINIDASE GENE FROM LambdaEMBL301-Neu1

DNA sequence analysis of the neuraminidase producing recombinant pMC2150 revealed that this clone contained a large ORF which was likely to be the neuraminidase gene as its predicted protein sequence contained the "aspartic box" motif found in other neuraminidases (see Introduction). However this ORF lacked the C-terminal translation stop codon, and another recombinant containing the whole neuraminidase sequence (pMC4170) was isolated to complete the sequence determination. Southern blot analysis revealed that the neuraminidase gene, as well as the flanking ORFs determined from pMC2150 and pMC4170 corresponded to a contiguous fragment of DNA in the chromosome of *S. pneumoniae*.

The neuraminidase gene from pMC2150 lacked the coding sequence for the last 233 amino acids of this enzyme. This C-terminal region comprises the three 20 amino acid repeats and the surface anchor motif found in the predicted neuraminidase amino acid sequence. The expression of neuraminidase activity from pMC2150 suggests that, this C-terminal region might not required for the activity of the enzyme.

The availability of pMC2150 encoding for an enzyme lacking the C-terminus, and pMC4170 having the full length enzyme would allow the study of the influence the C-terminus has in the cellular location of neuraminidase in the pneumococcus. This could be attempted by subcloning the neuraminidase from those two recombinants into shuttle vectors such us pVA838 [Macrina *et al* 1982] that replicate in the pneumococcus. The new recombinants could be transformed individually into *S. pneumoniae* neuraminidase isogenic negative mutants, and the location of the enzyme expressed from these recombinants could be investigated. Unfortunately, the possibility of several neuraminidase enzymes in the pneumococcus would make these type of experiments much more complex.

Analysis of the DNA sequence has revealed that the neuraminidase is

preceded by two -35 and -10 putative promoter sequences located some distance from the first putative translation codon of the gene. Promoters located some distance from the coding sequence have been also found in front of the pneumococcal genes coding for pneumolysin (*ply*)[Walker *et al* 1987], competence factor (*comA*)[Hui & Morrison 1991] and autolysin gene (*lytA*)[Diaz & García 1990]. In the case of *lytA*, two promoters have also been found [Diaz & García 1990]. In general, the presence of a long leader region is very uncommon and is usually involved in regulatory mechanisms of gene expression such us the formation of secondary structures at the mRNA level where certain proteins can bind and regulate translation of downstream located genes [Friensen *et al* 1983]. Expression of pneumococcal neuraminidase has been suggested to be regulated by the presence of neuraminic acid in the culture media [Kelly *et al* 1966; Kelly & Greiff 1970]. If the expression of neuraminidase activity is under such regulation, the large leader sequence found in front of the neuraminidase gene might be involved in this process although this is still far from being proved.

There are three possible ATG translation start codons at the beginning of the neuraminidase gene but only the second and third are associated with typical Shine-Dalgarno sequences. Although the presence of unusual Shine-Dalgarno sequences has been found in *S. pneumoniae* [De la Campa *et al* 1987], it seems unlikely that translation would be initiated from the first ATG, due to the location of the putative secretion signal peptide associated with the second ATG. Whether translation starts from the second or third ATG is not known. If translation starts from the second ATG, a full length signal peptide for neuraminidase would be synthesized, whereas translation initiation from the third ATG would produce an enzyme with a deficient signal peptide. This may allow the pneumococcus to produce two forms of the enzyme, one cytoplasmic and one exported, depending on which ATG is employed. This would agree with what has been found during the study of neuraminidase expression during growth of the pneumococcus, where enzyme activity was partly present in the culture supernatant, and partly associated with the cell. However, these possibilities are difficult to analyze due to the possible presence of more than one neuraminidase gene product.

Analysis of the predicted neuraminidase sequence has revealed the presence of a motif in the C-terminus of the protein that is present in many other surface-associated proteins from Gram-positive organisms. This motif comprises an anchoring region with a hydrophobic domain in the C-terminus of the protein followed by charged residues and preceded by the consensus sequence LPXTGX. This motif has been postulated to be responsible for the anchoring of proteins from Gram-positive bacteria to the cell membrane possibly through the hydrophobic domain [Schneewind *et al* 1991]. N-terminal to this region there is usually a region rich in proline, glycine, serine and threonine [Fischetti *et al* 1991]. This region, because of the proximity to the anchoring region, it is thought to span the peptidoglycan of the cell wall [Pancholi & Fischetti 1988]. Also, it has been hypothesized that the prolines and glycines from this region, with their ability to initiate bends and turns within proteins, may allow this part of the molecule to weave its way through the highly crosslinked peptidoglycan, thus stabilizing the protein within the cell wall [Pancholi & Fischetti 1988].

Protein anchoring via the motif described above seems transitional for some of the proteins containing this motif as they have been seen to be present both on the cell surface and then released into the culture media. This is the case of the cell surface antigen I/II from *S. mutants* [Kelly *et al* 1989], the protein G from *Streptococcus* G148 [Olsson *et al* 1987], the wall-associated protein A from *S. mutans* [Ferretti *et al* 1989], the streptococcal M protein [Pancholi & Fischetti 1989], and the cell wall proteinase Wg2 from *S. cremoris* [Kok *et al* 1988]. These proteins share the presence of signal peptides located at the N-termini of the molecules. Some of them are very hydrophilic with major hydrophobic domains located in the N-termini and C-termini which correspond to the signal peptide and membrane anchor motif respectively [Pancholi & Fischetti 1989; Olsson *et al* 1987; Ferretti *et al* 1989].

The pneumococcal neuraminidase from pMC4170 shares with the proteins described above most of the features mentioned. The hydrophobicity plot of neuraminidase suggests a hydrophilic protein with two major hydrophobic

domains, one in the N-terminus, corresponding to the signal peptide, and the other in the C-terminus, corresponding to the membrane anchor motif. Release of neuraminidase into the culture media has also been observed during the growth of the pneumococcus. These results would agree with the findings from the predicted amino acid sequence of the neuraminidase from pMC4170. However, the possibility of the expression of more than one neuraminidase in the pneumococcus makes these results questionable.

Engelman and Steitz (1981) proposed the "Helical Hairpin Hypothesis" for spontaneous insertion of proteins into and across membranes without involvement of transmembrane proteins. They proposed that during protein synthesis, the nascent polypeptide chain folds in an aqueous environment to form an antiparallel pair of helices, although is it not known whether cytoplasmic proteins are required for the correct folding of this polypeptide. These two helices may, in general, occur anywhere within the polypeptide, but for those proteins that contain a processed leader peptide, one of the two helices will be formed by the leader peptide. Insertion of the helical hairpin into the bilayer will initiate secretion if the second helix is polar. Secretion of a newly synthesized protein will continue until or unless a stretch of nonpolar residues is encountered [FIG. 6.1 (1 to 5)]. This model could be applied not only to the pneumococcal neuraminidase, but also to the proteins from Gram-positive organisms that, having the membrane anchor motif are released into the media (mentioned above). The signal peptide could form a hairpin with a helix of the nascent protein which would be immediately inserted in the cytoplasmic membrane initiating secretion of the newly synthesized protein. This secretion could stop when the hydrophobic helix, corresponding to the anchor motif, would form part of the hairpin. The positive charges found C-terminal to this hydrophobic helix could act as stop-transfer sequence [Dalbey 1990] contributing to the termination of translocation of the protein across the membrane. At this stage cleavage of the signal peptide by a signal peptidase may occur, and the protein could remain anchored to the cytoplasmic membrane by the hydrophobic domain of the C-terminal membrane anchor motif. The region preceding this hydrophobic domain would be likely to be located within the peptidoglycan

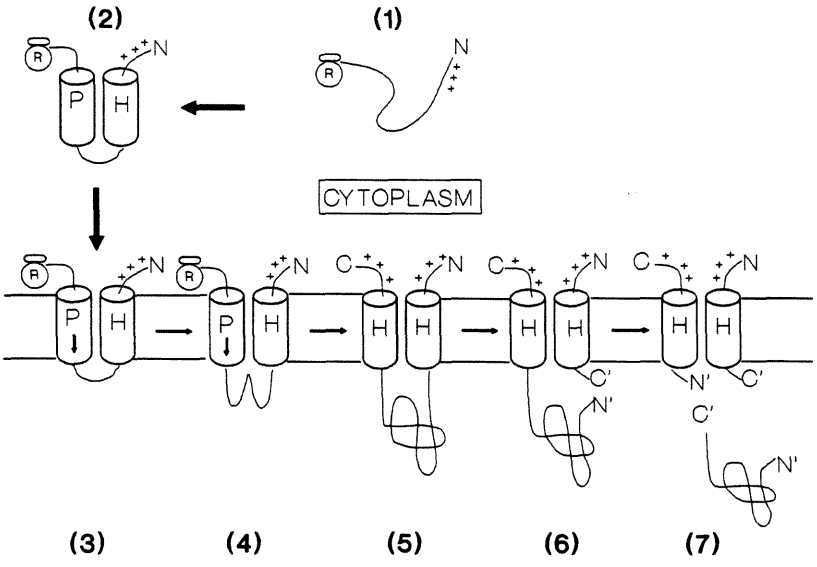


FIGURE 6.1. **Secretion model for some proteins from Gram-positive bacteria based on the "helical hairpin" hypothesis [Engelman & Steitz 1981].**
The cotranslational aspect of the process is represented by the presence of the initially soluble ribosome (R). (1) Initial polypeptide insufficiently long for stable folding. (2) After 40-50 residues are in the nascent chain, marginally stable folding in the helical hairpin occurs. (3) The hairpin, constituting of a polar (P) and a hydrophobic (H) helix, inserts spontaneously in the membrane. (4) Synthesis continues; the polypeptide passes through the membrane as a helix and folds outside it. (5) Translocation stops when two hydrophobic helices form the helix. (6) A signal peptidase releases the C-terminus from the membrane. (7) A membrane protease cleaves after the C-terminal hydrophobic region releasing the protein from the membrane.

of the cell wall, as the high percentage of proline and glycine found in this region would have the ability to initiate bends and turns of the protein, allowing this part of the molecule to weave its way through the highly crosslinked peptidoglycan [Pancholi & Fischetti 1988]. To explain the release

of these proteins into the growth media Pancholi and Fischetti (1989) have identified a membrane associated protease able to release M protein from cell membranes by cleavage of this protein before the threonine of the conserved sequence LPXTGX. Since this sequence is present in the C-terminal membrane anchor motif of all the proteins mentioned above, it is possible that a membrane-associated protease would cleave within this sequence releasing the protein into the growth culture [FIG. 6.1]. This postulated mechanism for neuraminidase secretion would imply the presence of two different molecular weight enzyme products in each culture corresponding to processed (supernatant) and unprocessed neuraminidase (cell-associated) respectively. Sequencing data can, therefore, be used to propose some basis for the secretion and cellular location of neuraminidase. However, validation of this hypothesis is complicated by the possible expression of more than one neuraminidase in the pneumococcus.

Although protein sequence analysis of the pneumococcal neuraminidase indicates that this enzyme might be associated with the surface of the pneumococcus through an anchor motif, the lack of neuraminidase activity from purified pneumococcal membranes is inconsistent with these predictions. The technique used for the purification of these membranes may not have been appropriate. Release of the neuraminidase from the membranes may have occurred either by mechanical sheering during the disruption of the cells, or by non-specific protease degradation. Alternatively, neuraminidase could have been liberated from the pneumococcal cell surface by the action of a specific proteolytic enzyme. This would raise the possibility of two different roles for pneumococcal neuraminidase, one related to its surface location and the other related to the secreted protein.

Very preliminary data from electron microscopy studies have suggested the presence of neuraminidase in the surface of the pneumococcus. This data agrees with the information obtained from the protein sequence analysis of this neuraminidase.

Electron microscopy studies have shown very little neuraminidase in the

cytoplasm of pneumococcus. The lack of a suitable negative control in these studies makes difficult the interpretation of this result. Newly synthesized neuraminidase might have been immediately inserted in the cell membranes after protein synthesis began and the postulated hairpin was formed. This would make the detection of high amounts of neuraminidase in the cytoplasm unlikely. However, degradation of the cytoplasmic enzyme during treatment of bacteria for electron microscopy studies can not be discarded.

Although sequencing data have suggested some basis for the secretion and cellular location of neuraminidase, these hypothesis cannot be tested until systems are developed to distinguish between the possible neuraminidase activities within the pneumococcus.

Comparison of the neuraminidases from many organisms has shown the presence of a consensus sequence that has been named "Aspartic box". This consensus sequence is repeated four or five times in these proteins (see Introduction). Analysis of the pneumococcal neuraminidase predicted protein sequence has revealed the presence of these aspartic boxes repeated four times in the molecule. The distance between the second and third of these consensus sequences is very conserved within all these neuraminidases, except in the case of *Trypanosoma cruzi*. In the region separating the second and third aspartic boxes, a conserved sequence comprising an invariant glycine residue, found not only in bacterial and eukaryotic, but also in viral neuraminidases, has been located. The homology found between all these neuraminidases indicates a common evolutionary origin for these enzymes, as has been suggested by Hoyer *et al* 1992. Recent studies carried out by Warner *et al* (1992) in the neuraminidase from *Salmonella typhimurium* LT2 have revealed that the active site of this enzyme is located near the C-terminus of the molecule outside the region of aspartic boxes. These results suggests that the aspartic boxes might not fall within the active site of this enzyme. However, the presence of these conserved sequences in most neuraminidases may play a role in the secondary structure of these proteins. Repetition of the aspartic boxes in neuraminidase may facilitate interaction with the repeated units that constitute oligosaccharides. These motifs could

also be involved in a secondary enzymatic activity, common to all the neuraminidases and different to the sialidase one. This secondary activity could have been present in the original genes from which neuraminidases might have evolved. Still, much more information about these neuraminidases is required before these possibilities can be analyzed.

6.5. PRESENCE OF DNA BOX ELEMENTS IN THE PNEUMOCOCCUS.

During the sequence analysis of DNA flanking the neuraminidase gene the presence of repeated sequences upstream of the neuraminidase gene was revealed. These sequences, named boxes (see Part V), are present in the flanking regions of other pneumococcal genes. These sequences were identified simultaneously by other laboratories. A collaboration was established with these laboratories. As a result of this joint work some preliminary data has already been published [Martin *et al* 1992]. A detailed experimental study of the role of these boxes is outside the scope of this thesis. However, computer analysis of these sequences have beginning to reveal some interesting features.

Analysis of the DNA boxes has revealed the presence of three different box elements (box A, box B and box C). The distribution of these elements varies in the flanking regions of different pneumococcal genes. Study of the predicted secondary structure of these boxes, using a single consensus box element of each, shows the putative formation of a stem-loop structure [FIG 6.2]. In several cases where box elements differ from the consensus, a compensating base change is seen in the complementary arm of the predicted stem-loop [FIG. 6.2] implying that maintenance of the secondary structure is important. This also reinforces the motion that this secondary structure might be real. Although FIGURE 6.2. shows the stem-loop in DNA, a similar structure might be present at RNA level.

Repeated sequences have been also found in other bacteria. The first of these sequences to be described were the Repetitive Extragenic Palindromic sequences (REP)[Higgins *et al* 1982], also called Palindromic

Units(PU)[Gilson *et al* 1984]. These sequences are about 35 nucleotides long, include an inverted repeat, and can occur singly or in multiple adjacent copies. They might occupy up to 1% of the genomes of *Escherichia coli* and *Salmonella typhimurium*. REP sequences are within transcribed sequences but

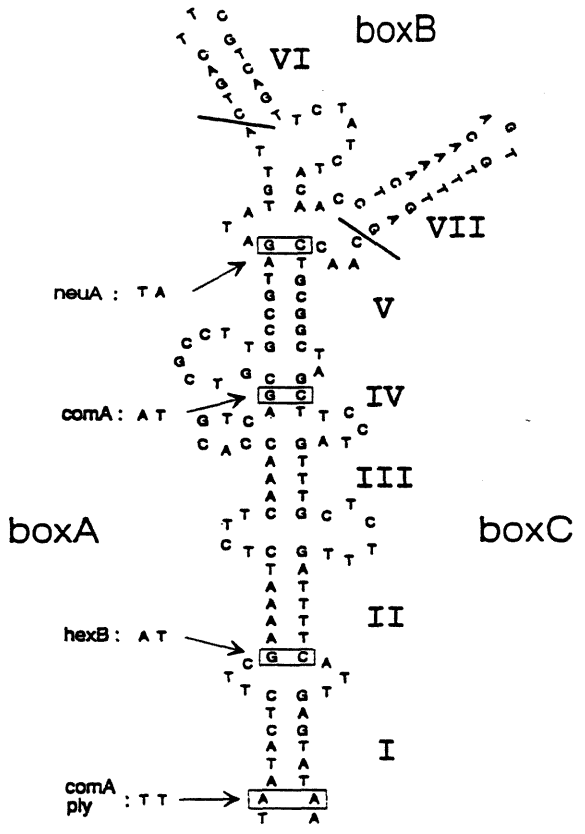


FIGURE 6.2. Predicted stem-loop structure of the consensus DNA boxes sequences containing a single copy of each box. ΔG=-42.9 Kcal mol^{-1}. Arrows mark the location of compensating base changes maintaining the structure in the indicated sequence. Helices are numbered I to VII. Straight lines separate boxA from boxB and boxC.[Figure from Martin *et al* 1992].

external to coding sequences [Stern *et al* 1984]. Gilson *et al* (1991) described a Bacterial Interspersed Mosaic Element (BIME) located within the Palindromic Units. Also present in enterobacteria are the Intergenic Repeated Units (IRUS)[Sharples & Lloyd 1990], called Enterobacterial Repetitive Intergenic Consensus (ERIC) by Hulton *et al* (1991). IRUS (or ERIC) sequences are 124-127 bp long, can form stem-loops and are not always transcribed.

Several roles have been attributed to these repeated sequences at both DNA and RNA level. REP sequences have been reported to stabilize mRNA by protecting it, through their secondary structure, from 3' to 5' nuclease degradation [Newbury *et al* 1987; Newbury *et al* 1987b; Stern *et al* 1988; Belasco & Higgins 1988]. Several proteins such as DNA gyrase and DNA polymerase I have been found to interact with REPs. It has been postulated that the binding of these proteins could mediate gene rearrangement through recombination events within the chromosome [Yang & Ames 1988; Gilson *et al* 1990]. In *S. typhimurium* it has been reported that distant REP sequences can recombine fusing the *hisD* gene to distant foreign promoters. These recombination processes have been postulated as a possible mechanism of evolution [Shyamala *et al* 1990].

The pneumococcal box elements have been found either upstream or downstream to certain pneumococcal genes (see chapter V). The box associated to the autolysin gene *lytA* is the only one known to be transcribed [Diaz & García 1990]. None of the other DNA boxes have been demonstrated to be within transcripts, therefore it would be difficult at this stage to attribute a role in RNA stability to these sequences as has been assigned to the repetitive elements mentioned above.

DNA sequence analysis of the boxes flanking the gene coding for neuraminidase activity has revealed the presence of two ORFs (ORF2 and ORF4) spanning the location of these boxes. It would be unlikely for ORF2 to code for a protein as no promoter or Shine-Dalgano sequences have been seen associated with this ORF. In contrast, these two elements have been

found for ORF4 although the synthesis of a polypeptide from this ORF still needs to be proved.

Although 24 pneumococcal genes were present in the databases when computer analysis of the box elements was done, these elements have only been found associated with either genes involved in competence or genes possibly related to virulence. Involved in competence are *comA* and *hexB*. The *comA* gene controls competence for genetic transformation [Chandler & Morrison 1988]. Competence has been described as a transitory cell state that is accompanied by a global shift of protein synthesis to the production of a few competence specific proteins [Morrison & Baker 1979]. The *HexB* gene encodes an essential component of the mismatch repair system which eliminates base pair mismatches arising in heteroduplex DNA during transformation [Lacks 1989]. The rest of the genes found near the DNA boxes (*ply*, *lytA*, *nanA*) could be related to virulence (see introduction). The *lytA* gene has also been considered to be associated with competence, as it is known that cells developing competence exhibit increased susceptibility to lysis in certain buffers [Seto & Tomasz 1975]. Also, a relationship between expression of pneumolysin and autolysin could be established as it is known that pneumolysin, having a cytoplasmic location, is released from pneumococcus by autolysis [Johnson 1980; Smyth & Duncan 1978]. Additionally, pneumococcal autolysin negative mutants have been shown to have decreased virulence in mice [Berry *et al* 1986b]. Expression of genes identified with competence and virulence might be cell growth dependent or regulated by environmental signals. Therefore, competence may not be completely independent of virulence. It could be speculated that the presence of the box elements in the flanking regions of some pneumococcal genes may play a role in the regulation of expression of all these genes. The possible role of *neuA* in virulence and its location near such box elements suggests that its expression could be regulated by these elements.

Finally, the pneumococcal DNA boxes, like the enterobacterial REP sequences, may be implicated in gene rearrangement and evolution . In *S. typhimurium* it has been seen that REP sequences can fuse genes to distant foreign promoters by recombination events [Shyamala *et al* 1990]. A similar

process may occur in the pneumococcus. During the cloning of neuraminidase gene from pJCP301, a large spontaneous deletion in the pneumococcal insert, that did not affect cosmid vector sequences, occurred [Berry *et al* 1988]. The possible presence of box elements in this pneumococcal insert could have been involved in that deletion event.

Although several hypothesis to explain the role of the pneumococcal boxes could be made, experimental data is required to discern between the multiple theories. For this purpose, specific mutations in the boxes to study alterations in the expression of genes associated with them, fusion of these sequences to already characterized genes, search for specific proteins that might bind to these boxes, and many other experiments will have to be designed to elucidate the role that these DNA elements might play in the pneumococcus.

6.6. PROSPECTS

In terms of IgA1 protease, being specific for human IgA1 it is difficult to study its role in virulence due to the lack of a suitable animal model. However the cloned IgA1 protease gene (*iga*) will allow a more detailed analysis of the structure and function of this enzyme. The cloned protease can be overexpressed in high expression vectors followed by purification of the enzyme. Production of monoclonal antibodies with the purified protease will allow determination of epitopes of the molecule that are involved in the recognition of human IgA1. DNA sequence determination of the pneumococcal IgA1 protease will enable comparison of this *iga* gene with other genes with the same activity from other organisms. The information obtained from this sequence comparison and the monoclonal antibody epitopes studies will allow the use of site-directed mutagenesis for the analysis of the structure and function of this enzyme.

With regard to neuraminidase, the first target will be to resolve the question of how many neuraminidases are expressed in the pneumococcus. The location of these enzymes will have to be resolved. Also the possibility of differential regulation of these enzymes will have to be investigated. Some of the essential tools to be used in the investigation of these questions will involve systematic cloning of neuraminidases, purification of these enzymes and construction of specific neuraminidase negative isogenic mutants.

Once these questions have been answered the possibility of a role in virulence for these enzymes will be analyzed. For this purpose, the molecular biology postulates (see introduction) will have to be answered.

To analyze the role of the pneumococcal intergenic DNA box elements, the effect of the deletion of specific boxes has in the expression of some of the genes associated with them will be studied. Also, specific recognition of these box elements by pneumococcal proteins will be analyzed. Variation of the structure and distribution of these sequences in different pneumococcal

serotypes will be investigated. All these preliminary experiments will provide some information about the role of these DNA elements in the pneumococcus.

REFERENCES

Ada, G.L., French, E.L. and Lind, P.E. (1961)
Purification and properties of neuraminidase from *Vibrio cholerae*. J. Gen. Microb. 24:409-421.

Adlam, C., Knight, J.M., Mugridge, A., Williams, J.M. and Lindon, J.C. (1987)
Production of colominic acid by *Pasteurella haemolytica* serotype A2 organisms. FEMS Microbiol. Lett. 42:23-25.

Air, G.M. and Graeme, L. (1989)
The neuraminidase of influenza virus. Proteins: Structure, Function and Genetics 6:341-356.

Andersson, B., Dahmén, J., Frejd, T., Leffler, H., Magnusson, G., Noori, G. and Edén, C.S. (1983)
Identification of an active disaccharide unit of a glycoconjugate receptor for pneumococci attaching to human pharyngeal epithelial cells. J. Exp. Med. 158:559-570.

Ashwell, G. and Morell, A.G. (1974)
Ther role of surface carbohydrates in the hepatic recognition and transport of the circulating glycoproteins. Adv. Enzymol. 41:99-128.

Austen, B.M. and Westwood, O.M.R. (1991)
Bacterial protein translocation. In, Protein targeting and secretion. D. Rickwood and D. Male Eds. IRL Press. Oxford University Press. 13-23.

Austrian R. (1981)
Some observations on the pneumococcus and on the current status of pneumococcal disease and its prevention. Rev. Infect. Dis. 3(suppl):S1-S17.

Austrian R. (1984)
Pneumococcal infections. In, Bacterial vaccines, Rene Germanier (ed) Academic Press, London. 257-288.

Austrian R. (1986)
Some aspects of pneumococcal carrier state. J. Antimicrobial Chemotherapy 18(suppl. A):25-45.

Baker, C.J. and Kasper, D.L. (1976)
Identification of sialic acid in polysaccharide antigens of group B

streptococcus. Infect. Immun. 13:284-288.

Bardardottir, E., Jonsson, S., Jonsdottir, I., Sigfusson, A. and Valdimarsson, H. (1990)
IgG subclass response and opsonization of *Streptococcus pneumoniae* after vaccination of healthy adults. J. Infect. Dis. 162:482-488.

Berry, A.M., Paton, J.C., Glare, E.M., Hansman, D. and Catcheside, D.E.A. (1988)
Cloning and expression of the pneumococcal neuraminidase gene in *Escherichia coli*. Gene 71:299-305.

Berry, A.M., Yother, J., Briles, D.E., Hansman, D. and Paton, J.C. (1989a)
Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. Infect. Immun. 57(7):2037-2042.

Berry, A.M., Lock, R.A., Hansman, D. and Paton, J.C. (1989b)
Contribution of autolysin to virulence of *Streptococcus pneumoniae*. Infect. Immun. 57:2324-2330.

Bhakdi, S., Tranum-Jensen, J. and Szeigoleit, S. (1985)
Mechanisms of membrane damage by streptolysin-O. Infect. Immun. 47:52-60.

Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983)
Buffer gradient gels and ^{25}S label as aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA 80:3963-3965.

Blake, M., Holmes, K.K. and Swanson, J. (1979)
Studies on gonococcus infection.XVII. IgA1-cleaving protease in vaginal washings from women with gonorrhoea. J. Infect. Dis. 139(1):89-92

Boulnois, G.J. (1992)
Pneumococcal proteins and the pathogenesis of disease caused by *Streptococcus pneumoniae*. J. Gen. Microb. 138:249-259.

Brodbeck, U., Gentinetta, R. and Lundin, S.J. (1973)
Multiple forms of a cholinesterase from body muscle of plaice (*Pleuronectes platessa*) and possible role of sialic acid in cholinesterase reaction specificity. Acta Chem. Scand. 27:561.

Brown, E.J., Hosea, S.W., Hammer, C.H., Burch, C.G. and Frank, M.M. (1982)
A quantitavive analysis of the interactions of antipneumococcal antibody and complement in experimental pneumococcal bacteremia. J. Clin. Invest. 69:85-98.

Bruyn, A.W. and van Furth, R. (1991)
Pneumococcal polysaccharide vaccines: indications, efficacy and recomendations. J. Clin. Microbiol. infect. Dis. 10(11):897-910.

Burnet, F.M. and Stone, J.D. (1947)
The receptor-destroying enzyme of *V. cholerae*. Austr. J. Exp. Biol. Med. Sci. 25:227-233.

Brandtzaeg, P. (1974)
Mucosal and glandular distibution of immunoglobulin components: differential localization of free and bound SC in secretory epithelial cells. J. Immunol. 112(4):1553-1559.

Bricker, J., Mulks, M.H., Plaut, A.G., Moxon, E.R. and Wright, A. (1983)
IgA1 proteases of *Haemophilus influenzae*: Cloning and characterization in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA 80:2681-2685.

Briles, E.B. and Tomasz, A. (1973)
Pneumococcal Forssman antigen. A choline-containing lipoteichoic acid. J. Biol. Chem. 248:6394-6397.

Cabezas, J.A. (1991)
Some questions and suggestions on the type references of the official nomenclature (IBU) for sialidase(s) and endosialidase. B.J. Letters 278:311-312.

Cámara, M., Mitchell, T.J., Andrew, P.W. and Boulnois, G.J. (1991)
Streptococcus pneumoniae produces at least two distinct enzymes with neuraminidase activity: Cloning of second neuraminidase gene in Escherichia coli. Infect. Immun. 59(8):2856-2858.

Cassidy, J.T., Jourdian, G.W. and Roseman, S. (1965)
The sialic acids. VI. Purification and properties of sialidase from *Clostridium perfringens*. J. Biol. Chem. 240:3501-3506.

Chandler, M.S. and Morrison, D.A. (1988)
Identification of two proteins encoded by *com*, a competence locus of *Streptococcus pneumoniae*. J. Bacteriol. 170(7):3136-3141.

Chen, J.-D. and Morrison, D.A. (1987)
Cloning of *Streptococcus pneumoniae* DNA fragments in *Escherichia coli* requires vectors protected by strong transcription terminators. Gene 55:179-187.

Chen, C.C. and Cleary, P.P. (1990)
Complete nucleotide sequence of the streptococcal C5a peptidase gene of *Streptococcus pneumoniae*. J. Biol. Chem. 265:3161-3167.

Clamp, J.R. and Johnson, I. (1972)
Immunoglobulins. In, Glycoproteins: their composition, structure, and function. A. Gottschalk ed. 2nd ed. Elsevier Publishing Co. Amsterdam 612-652.

Clowes, A.W., Cherry, R.J. and Chapman, D. (1972)
Physical effects of tetanus toxin on model membranes containing ganglioside. J. Mol. Biol. 67:49.

Cole, M.F. and Hale, C.A. (1991)
Cleavage of chimpanzee secretory immunoglobulin A by *Haemophilus influenzae* IgA1 protease. Microb. Pathog. 11:39-46.

Coonrod, J.D. and Yoneda, K. (1981)
Complement and opsonins in alveolar secretions and serum of rats with pneumonia due to *Streptococcus pneumoniae*. Rev. Infect. Dis. 3:310-322.

Coonrod, J.D. and Yoneda, K. (1982)
Comparative role of complement in pneumococcal and staphylococcal pneumonia. Infect. Immun. 37(3):1270-1277.

Costello, A.H., Cisar, J.O., Kolenbrander, P.E. and Gabriel, O. (1979)
Neuraminidase-dependent haemagglutination of human erythrocytes by human strains of *Actinomyces viscous* and *Actinomyces naeslundii*. Infect. Immun. 26:563-572.

Crain, M.J., Waltman II, W.D., Turner, J.S., Yother, J., Talkington, D.F., McDaniel, L.S., Gray, B.M. and Briles, D.E. (1990)
Pneumococcal surface protein A (PspA) is a serologically highly variable

and is expressed by all clinical important capsular serotypes of *Streptococcus pneumoniae*. Infect. Immun. 58(10):3293-3299.

Dalbey, R.E. (1990)
Positively charged residues are important determinants of membrane protein topology. TIBS 15:253-257.

De la Campa, A.G., Kale, P, Springhorn, S.S. and Lacks, S.A. (1987)
Proteins encoded by the *Dpn*II restriction gene cassette. Two methylases and an endonuclease. J. Mol. Biol. 196:457-469.

Deman, J.J. and Bryneel, E.A. (1973)
A method for the quantitative measurement of cell aggregation. Exp. Cell. Res. 81:351.

Demuth, D.R., Golub, E.E. and Malamud, D. (1990)
Streptococcal host interaction. Structural and functional analysis of a *Streptococcus sanguis* receptor for a human salivary glycoprotein. J. Biol. Chem. 265(13):7120-7126.

Devereux, J., Haeberli, P. and Smithies, O. (1984)
A comprehensive set of sequence analysis programs for the VAX. Nucl. Acid Res. 12:387-395.

Diaz, E. and Garcia, J.L. (1990)
Characterization of the transcription unit encoding the major pneumococcal autolysin. Gene 90:157-162.

Dillard, J.P. and Yother, J. (1991)
Analysis of *Streptococcus pneumoniae* sequences cloned into *Escherichia coli*: effect of promoter strength and transcription terminators. J. Bacteriology 173(16):5105-5109.

Dodge, P.R. and Swartz, M.N. (1965)
Bacterial meningitis - A review on selective aspects. II. Special neurologic problems, postmeningitic complications and clinicopathological correlations. New Engl. J. Med. 272(18):954-960.

Drzeniek, R., Scharmann, W. and Balke, E. (1972)
Neuraminidase and *N*-acetylneuramininate pyruvate-lyase of *Pasteurella multicocida*. J. Gen. Microbiol. 72:357-368.

Drzeniek, R. (1972b)
Viral and bacterial neuraminidases. Curr. Top. Microb. Immunol. 59:35-75.

Eisen, D. and Franson, R.C. (1987)
Acide-active neuraminidases in growth media from cultures of pathogenic *Naegleria fowleri* acid in sonicates of rabbit alveolar macrophages. Biochim. Biophys. Acta. 924:369.

Engelman, D.M. and Steitz, T.A. (1981)
The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell 23:411-422.

Falkow, S. (1988)
Molecular Koch's postulates applied to microbial pathogenicity. Rev. Infect. Dis. 10:S274-S276.

Fanger, M.W., Goldstine, S.N. and Shen L. (1983)
Cytofluorographic analysis of receptors for IgA on human polymorphonuclear cells and monocytes and the correlation of receptor expression with phagocytosis. Cell. Immunol. 20(9):1019-1027.

Farr, B. and Mandell, L. (1983)
Gram-positive pneumonia. In, Respiratory Infections: Diagnosis and Management , edited by James E. Penington. Raven Press, New York 217-226.

Feisntein, D. and Franklin, E.C. (1966)
Two antigenically distinguishable subclasses of human A myeloma proteins differing in their heavy chains. Nature 212:1496-1498.

Feldman, C., Munro, N.C., Jeffery, P.K., Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Guerreiro, D. Rohde, J.A.L., Todd, H.C., Cole, P.J. and Wilson, R. (1991)
Pneumolysin induces the salient histologic features of pneumococcal infection in the rat lung *in vivo*. Am. J. Cell Mol. Biol. 5:416-423.

Feldman, C., Read, R., Rutman, A., Jeffery, P.K., Brain, A., Lund, V., Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Todd, H.C., Cole, P.J. and Wilson, R. (1992)
The interaction of *Streptococcus pneumoniae* with intact human respiratory mucosa *in vitro*. Eur. Respir. J. 5:576-583.

Ferretti, J.J., Russell, R.R.B. and Dao, M.L. (1989)
Sequence analysis of the wall-associated protein precursor of *Streptococcus mutants* antigen A. Mol. Microbiol. 3:469-478.

Fischetti, V.A., Pancholi, V. and Schneewind, O. (1991)
Common characteristics of the surface proteins from Gram-Positive Cocci. In, Genetics and molecular biology of Streptococci, Lactococci, and Enterococci. G.M. Dunny, P.P. Cleary and L.L. McKay Ed. American Society for Microbiology. Washington, D.C. 290-294.

Fishman, Y., Bricker, J., Gilbert, J.V., Plaut, A.G. and Wright, A. (1985)
Cloning of the type 1 Immunoglobulin A1 protease from *Neisseria gonorrhoeae* and secretion of the enzyme from *Escherichia coli*. In, The Pathogenic Neisseriae. G.K. Schoolnik Ed. A.S.M. Washington DC 164-168.

Flashner, M.P., Wang, P., Hurley, J. and Tanembaum, S.W. (1977)
Properties of an inducible extracellular neuraminidase from an *Arthrobacter* isolate. J. Bacteriol. 129:1457-1465.

Friensen, J.D., Tropak, M. and An, G. (1983)
Mutations in the *rplJ* leader of *Escheria coli* that abolish feedback regulation. Cell 32:163-167.

Frithz, E., Heden, L.-O. and Lindahl (1989)
Extensive sequence homology between IgA receptor and M protein in *Streptococcus pyogenes*. Mol. Microbiol. 3:1111-1119.

Fujiyama, Y., Kobayashi, K., Senda, S., Benno, Y., Bamba, T. and Hosoda, S. (1985)
A novel IgA protease from *Clostridium sp.* capable of cleaving IgA1 and IgA2 A2m(1) allotype but not IgA2 A2m(2) allotype paraproteins. J. Immunol. 134(1):573-576.

Galen, J.E., Virm, E.R., Lawrisuk, L. and Kaper, J.B. (1990)
Cloning, sequencing, and expression of the gene, nanH, for *Vibrio cholerae* neuraminidase. In, Advances in research on cholera and related diarrheas. Vol 7. R.B. Sack and Y. Zinnaka Ed. K.T.K. Scientific Publishers, Tokyo 143-153.

Galen, J.E., Ketley, J.M., Fasano, A., Richardson, S.H., Wasserman, S.S.

and Kaper, J.B. (1992)
Role of *Vibrio cholerae* neuraminidase in the function of cholera toxin. Infect. Immun. 60(2):406-415.

García, E., García, J-.L., Ronda, C., García, P. and López, R. (1985)
Cloning and expression of the pneumococcal autolysin gene in *Escherichia coli*. Mol. Gen. Genet. 201:225-230.

García, P., García, J.L., García, E. and López, R. (1986)
Nucleotide sequence and expression of the pneumococcal autolysin gene in *Escherichia coli* from its own promoter. Gene 43:265-272.

Gauldie, J., Richards, C. and Lamontagne, L. (1983)
Fc receptors for IgA and other immunoglobulins on resident and activated alveolar macrophages. Cell. Immunol. 20(9):1029-1037.

Giebink, G.S. (1981)
The pathogenesis of pneumococcal otitis media in chinchillas and the efficacy of vaccination in prophylaxis. Rev. Infect. Dis. 3(2):342-353.

Gilbert, J.V. and Plaut, A.G. (1983)
Detection of IgA1 protease activity among multiple bacterial colonies. J. Immunol. Methods 57:247-251.

Gilbert, J.V., Plaut, A.G. and Longmaid, B. (1983)
Inhibition of bacterial IgA proteases by human secretory IgA and serum. J. Mol. Immunol. 20:1039-1049.

Gilbert, J.V., Plaut, A.G., Fishman, Y. and Wright A. (1988)
Cloning of the gene encoding streptococcal immunoglobulin A protease and its expression in *Escherichia coli*. Infect. Immun. 56(8):1961-1966.

Gilbert, J.V., Plaut, A.G. and Wright, A. (1991)
Analysis of the immunoglobulin A protease gene of *Streptococcus sanguis*. Infect. Immun. 59(1):7-17.

Gillespie, S.H. (1989)
Aspects of pneumococcal infection including bacterial virulence, host response and vaccination. J. Med. Microb. 28:237-248.

Glick, G.D., Toogood, P.L., Wiley, D.C., Skehel, J.J. and Knowles, J.R.

(1991)
Ligand recognition by influenza virus. The binding of bivalent sialosides. J. Biol. Chem. 266(35):23660-23669.

Gomi, H. Hozumi, T., Hattori, S., Tagawa, C, Kishimoto, F. and Bjorck, L. (1990)
The gene sequence and some properties of protein H. J. Immunol. 144:4046-4052.

Gorter, A., Hiemstra, P.S., Leijh, P.C.J., van der Sluys, M.E., van der Barselaar, M.T., van Es, L.A. and Daha, M.R. (1987)
IgA- and secretory IgA-opsonized *S. aureus* induce a respiratory burst and phagocytosis by polymorphonuclear leukocytes. Immun. 61:303-309.

Griffiss, J.M. (1983)
Biological function of the serum IgA system. Modulating complement mediated effector mechanisms and conservation of antigenic mass. Ann. N. Y. Acad. Sci. 409:697-707.

Grundy, F.J. Plaut, A.G. and Wright, A. (1990)
Localization of the cleavage site specifically determinant of *Haemophilus influenzae* immunoglobulin A1 protease genes. Infect. Immun. 58(2):320-331.

Grundy, F.J., Plaut, A. and Wright, A. (1987)
Haemophilus influenzae immunoglobulin A1 protease gene: cloning by plasmid integration-excision, comparative analyses, and location of secretion determinants. J. Bacteriol. 169(10):4442-4450.

Guckian, J.C., Christensen, G.D. and Douglas, P.F. (1980)
The role of opsonins in recovery from experimental pneumococcal pneumonia. J. Infect. Dis. 142(2):175-190.

Guss, B., Uhlen, M., Nilsson, B., Lindberg, M., Sjoquist, J. and Sjodahl. (1984)
Region X, the cell-wall-attachment part of staphylococcal protein A. Eur. J. Biochem. 138:413-420.

Guzmán, C.A., Platé, M. and Pruzzo, C. (1990)
Role of neuraminidase-dependent adherence in *Bacteroides fragilis* attachement to human epithelial cells. FEMS Microb. Letters. 71:187-192.

Haanes, E.J. and Cleary, P.P. (1989).
Identification of a divergent M protein gene and an M protein-related gene family in serotype 49 *Streptococcus pyogenes* J. Bacteriol. 171:6397-6408.

Hadfield, C. (1987)
Genomic library cloning with plasmids, lambda and cosmid vectors. In, Gene cloning and analysis. A lab guide G.J. Boulnois Ed. 61-105.

Halpern, M.S. and Koshland, M.E. (1970)
Novel subunit in secretory IgA. Nature 228:1276-1278

Halter, R., Pohlner, J. and Meyer, T.F. (1984)
IgA protease of *Neisseria gonorrhoeae*: isolation and characterization of the gene and its extracellular product. EMBO J. 3(7):1595-1601.

Haksar, A., Maudsley, D.V., Kimmel, G.L., Peron, F.G., Robidoux, W.R.(Jr) and Gahnon, G. (1974).
Adrenocorticotropin stimulation of cyclic adenosine 3'-5'-monophosphate formation in isolated rat adrenal cells. The role of membrane sialic acid. Biochim. Biophys. Acta. 362:356.

Harford, C.G. and Hara, M. (1950)
Pulmonary edema in influenzal pneumonia of the mouse and the relation of fluid in the lung to the inception of pneumococcal pneumonia. J. Exp. Med. 91:245-260.

Heath, D.G. and Cleary, P.P. (1989)
Fc-receptor and M protein genes of group A streptococci are products of gene duplication. Proc. Natl. Acad. Sci. USA. 86:4741-4745.

Henningsen, M., Roggentin, P. and Schauer, R. (1991)
Cloning, sequencing and expression of the sialidase gene from *Actinomyces viscous* DSM 43798. Biol. Chem. Hoppe-Seyler. 372:1065-1072.

Heremans, J.F. (1974)
Immunoglobulin A. In, The antigens. M. Sela Ed. Academic Press, New York 365-504.

Heuermann, D., Roggentin, P., Kleineidam, R.G. and Schauer, R. (1991)
Purification and characterization of a sialidase from *Clostridium chauvoei* NC08596. Glycoconjugate J. 8:95-101.

Hohn, B. (1979)
In vitro packaging of Lambda and cosmid cloning. Methods. Enzymol. 68:299-309.

Hollingshead, S.K., Fischetti, V.A. and Scott, J.R. (1986)
Complete nucleotide sequence of type 6M protein of the group A streptococcus: repetitive structure and a membrane anchor. J. Biol Chem. 261:1677-1686.

Hollingshead, S.K., Fischetti, V.A. and Scott, J.R. (1987)
A highly conserved region present in transcripts encoding heterologous M proteins of group A streptococcus. Infect. Immun. 55:3237-3239.

Horne, D. and Tomasz, A. (1985)
Pneumococcal Forssman antigen: enrichment in mesosomal membranes and specific binding to the autolytic enzyme of *Streptococcus pneumoniae*. J. Bacteriol. 161:18.

Hoyer, L.L., Roggentin, P., Schauer, R. and Vimr, E.R. (1991)
Purification and properties of cloned *Salmonella typhimurium* LT2 sialidase with virus-typical kinetic preference for sialyl α2→3 linkages. J. Biochem. 110:462-467.

Hoyer, L.L., Hamilton, A.C., Steenbergen, S.M. and Vimr, E.R. (1992)
Cloning, sequencing and distribution of the *Salmonella typhimurium* LT2 sialidase gene, *nanH*, provides evidence for interspecies gene transfer. Mol. Microb. 6(7):873-884.

Hughes, R.C. and Jeanloz, R.W. (1964)
The extracellular glycosidases of *Diplococcus pneumoniae*. I. Purification and properties of neuraminidase and a β-galactosidase. Action of the α_1-acid glycoprotein of human plasma. Biochemistry 3(10):1535-1543.

Hui, F.M. and Morrison, D.A. (1991)
Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *com*A, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family. J. Bacteriol. 173(1):372-381.

Humphrey, J.H. (1944)
Hyaluronidase production by the pneumococci. J. Pathol. 56:273-275.

Igarashi, M. and Bando, Y. (1990)
Enhanced efficiency of cell hybridization by neuraminidase treatment. J. Immun. Methods 135:91-93.

Irani, R.J. and Ganapathi, K. (1962)
Occurence of sialic acid in some gram-positive and gram-negative pathogenic bacteria. Nature 195:1227.

Itoh, K. and Kumagai, K. (1980)
Effect of tunicamycin and neuraminidase on the expression of Fc-IgM and -IgG receptors on human lymphocytes. J. Immunol. 124:1830-1836.

Jeanloz, J. and Codington, J.F. (1976)
The biological role of sialic acid at the surface of the cell. In, Biological role of sialic acids. A. Rosenberg & C.-L. Schengrund Ed. Pennsylvania State University, Hershey. Plennum Press. New York & London 201-238.

Jeery, L.M., Kunkel, H.G. and Grey, H.M. (1970)
Absence of disulfide bonds linking the heavy and light chains: A property of a genetic variant of ΓA2 globulins. Proc. Natl. Acad. Sci. 65(3):557-563.

Jennings, H.J., Lugowski, C. and Young, N.M. (1980)
Structure of the complex polysaccharide C-substrate from *Streptococcus pneumoniae* type I. Biochem. 19:4712-4719.

Johnson, M.K. (1977)
Cellular location of pneumolysin. FEMS Microb. Lett. 2:243-245.

Karjalainen, H., Koskela, M., Luotonen, J. and Sipilä, P. (1991)
Secretory antibodies specific to *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Branhamella catarrhalis* in middle ear effusion during acute otitis media. Acta Otolaryn. 111:524-529.

Kedzierska, B., Mikulaszek, E. and Pongonowska-Goldhar, J. (1968)
Immunochemical studies on Salmonella serogroup 48.3. *N*-acetylneuraminic acid as an immunodominant sugar in *S. dahlea*. Bull. Acad. Pol. Sci. Biol. 16:673-676.

Kelly, C., Evans, P., Bergmeir, L., Lee, S.F., Fox, H. Proguiske, A.,

Harris, A.C., Aitken, A., Bleiweis, A.S. and Lerner, T. (1989)
Sequence analysis of a cloned streptococcal surface antigen I/II. FEBS Letters 258:127-132.

Kelly, R.T., Greiff, D. and Farmer, S. (1966)
Neuraminidase activity in *Diplococcus pneumoniae*. J. Bacteriol. 91(2):601-603.

Kelly, R.T., Farmer, S. and Greiff, D. (1967)
Neuraminidase acctivity of clinical isolates of *Diplococcus pneumoniae*.J. Bacteriol. 94:272-273.

Kelly, R. and Greiff, R. (1970)
Toxicity of pneumococcal neuraminidase. Infect. Immun. 2(1):115-117.

Kilian, M., Mestecky, J. and Russell, M.W. (1988)
Defense mechanisms involving Fc-dependent functions of immunolgobulin A and their subversion by bacterial immunoglobulin A proteases. Microb. Rev. 52(2):296-303.

Kilian, M. and Reinholdt, J. (1987)
A hypothetical model for the development of invasive infection due to IgA1 protease-producing bacteria. Adv. Exp. Med. Biol. 216B:1261-1269.

Kilian, M. and Reinholdt, J. (1986)
Interference with IgA defence mechanisms by extracellular bacterial enzymes. Medical Microb. 5:173-208.

Kilian, M. and Thomsen, B. (1983)
Antigenic heterogeneity of immunoglobulin A1 proteases from encapsulated and non-encapsulated *Haemophilus influenzae*. Infect. Immun. 24:126-132.

Kilian, M. (1981)
Degradation of immunoglobulins A1, A2, and G by suspected principal periodontal pathogens. Infect. Immun. 34(3):757-765.

Kilian, M., Mestecky, J., Kulhavy, R., Tomana, M. and Butler, W.T. (1980).
IgA1 proteases from *Haemophilus influenzae*, *Streptococcus pneumoniae*, *Neisseria meningitidis*, and *Streptococcus sanguis*: comparative immunochemical studies. J. Immunol. 124(6):2596-2600.

Klugman, K.P. (1990)
Pneumococcal resistance to antibiotics. Clin. Microb. Rev. 3(2):171-196.

Knecht, J.C., Schifman, G. and Austrian, R.D. (1970)
Some biological properties of pneumococcus type 37 and the chemistry of its capsular polysaccharide. J. Exp. Med. 132:475-487.

Kholi, J. and Grosjean, H.(1981)
Usage of three termination codons: compilation and analysis of the known eucaryotic and prokaryotic translation termination sequences. Mol. Gen. Genetics 182:430-439.

Kok, J., Leenhouts, K.J., Haandrikman, A.J., Ledeboer, A.M. and Venema, G. (1988).
Nucleotide sequence of the cell wall proteinase gene of *Streptococcus cremoris* Wg2. Appl. Environ. Microb. 54:231-238.

Koomey, J.M. and Falkow, S. (1984)
Nucleotide sequence homology between the immunoglobulin A1 protease gene of *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Haemophilus influenzae*. Infect. Immun. 43(1):101-107.

Koomey, J.M., Gill, R.E. and Falkow, S. (1982)
Genetic and biochemical analysis of gonococcal IgA1 protease: cloning in *Eschericha coli* and construction of mutants of gonococci that fail to produce the activity. Proc. Natl. Acad. Sci. USA 79:7881-7855.

Kozak, M. (1983)
Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. 47:1-45.

Kunkel, G., Smith, K., Joslin, G., Natvig, B., Litwin, D. (1969)
Genetic marker of the ГА2 subgroup of ГА immunoglobulins. Nature 223:1247-1248.

Kyte, J. and Doolittle, R.F. (1982)
A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Lacks, S.A. (1989)

Generalized DNA mismatch repair: Its molecular basis in *Streptococcus pneumoniae* and other organisms. In, genetic transformation and expression. L.O. Butler, C.R. Harwood and B.E.B. Moseley Ed. Intercept Ltd. 325-339.

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

Langridge, J., Langridge, P. and Bergquist, P.L. (1980)
Extraction of nucleic acids from agarose gels. Anal. Biochem. 103:264-271.

Lathe, R., Vilotte, J.L. and Clark, A.J. (1987)
Plasmid and bacteriophage vectors for excision of intact inserts. Gene 57:193-201.

Lee, L.T. and Howe, C. (1966)
Pneumococcal neuraminidase. J. Bacteriol. 91(4):1418-1426.

Lee, C.-J., Fraser, B.A., Szu, S. and Lin, K.-T. (1981)
Chemical structure of and immune response to polysaccharides of *Streptococcus pneumoniae*. Rev. Infect. Dis. 3(2):323-331.

Lee, C.-J., Banks, S.D. and Li, J.P. (1991)
Virulence, immunity, and vaccine related to *Streptococcus pneumoniae*. Critical Rev. Microb. 18(2):89-114.

Leprate, R. and Michel-Briand, T. (1980)
Extracellular neuraminidase production by a strain of *Pseudomonas aeruginosa* isolated from cystic fibrosis. Ann. Microb. (Paris) 131:209-222.

Lindh, E. (1975)
Increased resistance of immunolgobulin A dimers to proteolytic degradation after binding of secretory component. J. Immunol. 114(1):284-286.

Lock, R.A., Paton, J.C. and Hansman, D. (1988b)
Purification and immunological characterization of neuraminidase produced by *Streptococcus pneumoniae*. Microb. Pathog. 4:33-43.

Lock, R.A., Paton, J.C. and Hansman, D. (1988)
Comparative efficacy of pneumococcal neuraminidase and pneumolysin as immunogens protective against *Streptococcus pneumoniae*. Microb. Pathog. 5:461-467.

Lowell, G.H., Smith, L.F., Griffiss, J.M. and Brandt, B.L. (1980)
IgA-dependent, monocyte mediated, antibacterial activity. J. Exp. Med. 152:452-457.

Lue, C., Prince, J., Fattom, A., Schneerson, R., Robbins, J.B. and Mestecky, J. (1990)
Antibody-secreting peripheral blood lymphocytes induced by immunization with a conjugate consisting of *Streptococcus pneumoniae* type 12F polysaccharide and diphtheria toxoid. Infect. Immun. 58(8):2547-2554.

Luppi, A. and Cavazzini, G. (1966).
Sialic acid determination in some Corynebacteriaceae. Nuovi. Ann. Ig. Microbiol. 17:183.

Macleod, C.M. and Krauss, M.R. (1950)
Relation of virulence of pneumococcal strains for mice to the quantity of capsular polysaccharide formed *in vitro*. J. Exp. Med. 92:1-9.

Macrina, F.L., Tobian, J.A., Jones, K.R., Evans, R.P. and Clewell, D.B. (1982)
A cloning vector able to replicate in *E. coli* and *Streptococcus sanguis*. Gene 19:345-353.

Magnusson, K.E., Stendahl, O., Stjernström, I. and Edebo, L. (1979)
Reduction of phagocytosis, surface hydrophobicity and charge of *Salmonella typhimurium* 395 MR10 by reaction with secretory IgA (SIgA). Immunol. 36:439-447.

Magnusson, K.E. and Stjernström, I. (1982)
Mucosal barrier mechanisms. Interplay between secretory IgA (SIgA), IgG and mucins on the surface porperties and association of salmonellae with intestine and granulocytes. Immunol. 45:239-248.

Male, C.J. (1979)
Immunoglobulin A1 protease production by *Haemophilus influenzae* and *Streptococcus pneumoniae*. Infect. Immun. 26:254-261.

Mallett, C.P., Boylan, R.J. and Everhart, D.L. (1984)
Competent antigen-binding fragment (Fab) from secretory immunoglobulin A using *Streptococcus sanguis* immunoglobulin A protease. Caries Res. 18:201-208.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982)
Molecular cloning, a laboratory manual. Cold Spring Harbor, USA.

Mansa, B. and Kilian, M. (1986)
Retained antigen-binding activity of Fabα fragment of human monoclonal immunoglobulin A1 (IgA1) cleaved by IgA1 protease. Infect. Immun. 52(1):171-174.

Marmur, J. and Doty, P. (1962)
Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. J. Mol. Biol. 5:109-118.

Martin, B., Humbert, O., Cámara, M., Guenzi, E., Walker, J., Mitchell, T., Andrew, P. Prudhomme, M., Alloing, G., Hakenbeck, R., Morrison, D.A., Boulnois, G.J. and Claverys, J.-P. (1992)
A highly conserved repeated DNA elements located in the chromosome of *Streptococcus pneumoniae*. Nucleic Acids Research 20(13):3479-3483.

McCrea, J.F. (1947)
Modification of red-cell agglutinability by *C. welchii* toxins. Austr. J. Exp. Biol. Med. Sci. 25:127-136.

McDaniel, L.S., Yother, J., Vijayakumar, M., McGarry, L., Guild, W.R. and Briles, D.E. (1987)
Use of insertional inactivation to facilitate studies of biological properties of pneumococcal surface protein A (PspA). J. Exp. Med. 165:381-394.

McGhee, J.R. and Michalek, S.M. (1981)
Immunology of dental caries: microbial aspects and local immunity. Ann. Rev. Microbiol. 25:595-638.

McMenamin, J.B. and Volpe, J.J. (1984)
Bacterial meningitis in infancy: effects of intracranial pressure and cerebral blood flow velocity. Neurology 34:500-504.

Mehta, S.K., Plaut, A.G. and Calvanico, N.J. (1973)

Human IgA production of an Fc fragment by an enteric microbial proteolytic enzyme. J. Immunol 3:1274-1276

Miller, L., Gray, L., Beachey, E.H. and Kehoe, M.A. (1988)
Antigenic variations among group A streptococcal M proteins: nucleotide sequence of the serotype 5 M protein gene and its relationship with genes encoding types 6 and 24 M proteins. J. Biol. Chem. 263:5668-5673.

Milligan, T.W., Mattingly, S.J. and Strauss, D.C. (1980)
Purification and partial characterization of neuraminidase from type III group B streptococci. J. Bacteriol. 144:164-172.

Mitchell, T.J., Andrew, P.W., Boulnois, G.J., Lee, C.-J., Lock, R.A. and Paton, J.C. (1992)
Molecular studies of pneumolysin, the thiol-activated toxin of *Streptococcus pneumoniae* as an aid to vaccine design. In, Bacterial protein toxins. Witholt *et al* Eds. Gustav Fischer, Stuttgart, Jena, New York. Zbl. Bakt. Suppl. 23:429-438.

Molla, A., Kagimoto, T. and Maeda, H. (1988)
Cleavage of immunoglobulin G (IgG) and IgA around the hinge region by proteases from *Serratia marcescens*. Infect. Immun. 56(4):916-920.

Moncla, B.J. and Braham, P. (1989)
Detection os sialidase (neuraminidase) activity in *Actinomyces* species by using 2'-(4-methylumbelliferyl)α-D-*N*-acetylneuraminic acid in a filter paper spot test. J. Clin. Microb. 27(1):182-184.

Morell, A.G., Gregoriadis, G. and Scheinberg, I.H. (1971)
The role of sialic acid in determining the survival of glycoproteins in the circulation. J. Biol. Chem. 246(5):1461-1467.

Moriyama, T. and Barksdale, L. (1967)
Neuraminidase of *Corynebacterium diphtheriae*. J. Bacteriol. 94:1565-1581.

Morrison, D.A. and Baker, M.F. (1979)
Competence for genetic transformation in pneumococcus depends on synthesis of small set of proteins. Nature 282:215-217.

Morrison D.A. and Jaurin, B. (1990)
Streptococcus pneumoniae possesses canonical *Escherichia coli* (sigma 70)

promoters. Mol. Microbiol. 4(7):1143-1152.

Mortensen, S.B. and Kilian, M. (1984)
A rapid method for the detection and quatitation of IgA protease activity by macrobore gel-permeation chromatography. J. Chromatography 296:257-262.

Mouw, A.R., Beachey, E.H. and Burdett, V. (1988)
Molecular evolution of streptococcal M protein: cloning and nucleotide sequence of type 24 M protein gene and relation to other genes of *Streptococcus puogenes*. J. Bacteriol. 170:676-684.

Moxon, E.R., Glode, M.P., Sutton, A. and Robbins, J.B. (1977)
The infant rat as a model of bacterial meningitis. J. Infect. Dis. 136:S186-S190.

Moxon, E.R. (1981)
Experimental infections of animals in the study of *Streptococcus pneumoniae*. Rev. Infect. Dis. 3(2):354-357.

Mufson, M.A, Hughey, D.F., Turner, C.E. and Schiffman, G. (1991)
Revaccination with pneumococcal vaccine of elderly persons 6 years after primary vaccination. Vaccine 9:403-407.

Mulks, M.H., Kornfeld, S.J., Frangione, B. and Plaut, A.G. (1982)
Relationship between the specificity of IgA proteases and serotypes in *Haemophilus influenzae*. J. Infect. Dis. 146(2):266-274.

Mulks, M.H. (1983)
Microbial IgA1 proteases. In, Bacterial Enzymes and Virulence. Eds. Holder I.A. Boca-Raton F.L. CRC Press. Ch.4 81-104.

Müller, H.E. (1971)
Immunoelectrophoretische Untersuchungen zur einwirkung bakterieller enzyme auf menschliche plasmaproteine. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. 217:254.

Musher, D.M., Chapman, A.J., Goree, A., Jonsson, S., Briles, D. and Baughn, R.E. (1986)
Natural and vaccine-related immunity to *Streptococcus pneumoniae*. J. Infect. Dis. 154:245-256.

Musher, D.M. (1992)
Infections caused by *Streptococcus pneumoniae*: clinical spectrum, pathogenesis, immunity, and treatment. Clin. Infect. Dis. 14:801-809.

Myhill, M. and Cook, T.M. (1972)
Extracellular neuraminidase of *Streptomyces albus*. Can. J. Microb. 18:1007-1014.

Nandoskar, M., Ferrante, A., Bates, E.J., Hurst, N. and Paton, J.C. (1986)
Inhibition of human monocyte respiratory burst, degranulation, phospholipid methylation and bacterial activity by pneumolysin. Immunology 59:515-520.

Newbury, S.F., Smith, N.H., Robinson, E.C., Hiles, I.D. and Higgins, C.F. (1987)
Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell 48:297-310.

Ng, S.-S. and Dain, J.A. (1976)
The natural occurence of sialic acids. In, Biological role of sialic acids. Abraham Rosenberg & Cara-Lynne Schengrund Eds. Pennsylvania State Univerrsity, Hershey. Plenum Press. New York & London 59-102.

Ohta, Y., Tsukada, Y. and Sugimori, T. (1989)
Purifications and properties of neuraminidase isoenzymes in *Arthrobacter ureafaciens* mutant. J. Biochem. 106:1086-1089.

Okahashi, N., Sasakawa, S., Yoshikawa, S., Hamada, S. and Koga, T. (1989)
Molecular characterization of a surface protein antigen gene from serotype C *Streptococcus mutants* implicated in dental caries. Mol. Microb. 3:673-678.

Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M. and Uhlen, M. (1987).
Structure and evolution of the repetitive gene encoding streptococcal protein G. Eur. J. Biochem. 168:319-324.

O'Toole, R.D., Goode, L. and Howe, C. (1971)
Neuraminidase activity in bacterial meningitis. J. Clin. Invest. 50:979-985.

O'Toole, R.D. and Stahl, W.L. (1975)
Experimental pneumococcal meningitis. Effects of neuraminidase and other pneumococcal constituents on cerebrospinal fluid in the intact dog. J. Neurol. Sci. 26:167-178.

Pancholi, V. and Fischetti, V.A. (1988)
Isolation and characterization of the cell-associated region of Group A streptococcal M6 protein. J. Bacteriol. 170(6):2618-2624.

Pancholi, V. and Fischetti, V.A. (1989)
Identification of endogenous membrane anchor-cleaving enzyme for group A streptococcal M protein. J. Exp. Med. 170:2119-2133.

Papkoff, H. (1966)
Glycoproteins with biological activity. In, Glycoproteins: their composition, structure, and function. A. Gottschalk Ed. 1st ed. Elsevier Publishing Co., New York 532-557.

Pardoe, G.I. (1970)
The inducible neuraminidase CN-acyl-neuraminyl hydrolase (Ec.3.2.18) of *Klebsiella aerogenes* NCID9479. Path. Microb. 35:361-376.

Pasteur, L. (1881)
Note sur la maladie nouvelle provoquée par la salive d'un enfant mort de la rage. Bulletin de l'Académie de Médicine (Paris) [series 2] 10:94-103.

Paton, J.C. and Ferrante, A. (1983)
Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin. Infect. Immun. 41:1212-1216.

Paton, J.C., Lock, R.A. and Hansman D.J. (1983)
Effect of immunization with pneumolysin on survival time of mice challenged with *Streptococcus pneumoniae*. Infect. Immun. 40(2):548-552.

Paton, J.C., Rowan-Kelly, B. and Ferrante, A. (1984)
Activation of human complement by the pneumococcal toxin pneumolysin. Infect. Immun. 43:1085-1087.

Paton, J.C., Lock, R.A., Lee, C.-J., Li, J.P., Berry, A.M., Mitchell, T.J., Andrew, P.W. and Boulnois, G.J. (1991)

Purification and immunogenicity of genetically toxoided derivatives of pneumolysin and their conjugation to *Streptococcus pneumoniae* type 19F polysaccharide. Infect. Immun. 59:2297-2304.

Pearson, W.R. and Lipman, D.J. (1988)
Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.

Peeters, C.C.A.M., Tenbergen-Meekes, A.-M., Haagmans, B., Evenberg, D., Poolman, J.T., Zegers, B.J.M. and Rijkers, G.T. (1991)
Pneumococcal conjugate vaccines. Immunol. Letters 30:267-274.

Pellegrin, J.-L.J., Ortega-Barria, E., Barza, M., Baum, J. and Pereira, M.E.A. (1991)
Neuraminidase activity in Acanthamoeba species trophozoites and cysts. Investigative ophtalmology and visual science 32(12):3061-3066.

Pereira, M.E.A. and Hoff, R. (1986)
Heterogeneous distribution of neuraminidase activity in strains and clones of *Trypanosoma cruzi* and its possible association with parasite myotropism. Mol. Biochem. Parasitol. 20:183-189.

Pereira, M.E.A., Mejia, J.S., Ortega-Barria, E., Matzilevich, D. and Prioli, R.P. (1991)
The *Trypanosoma cruzi* contains sequences similar to bacterial neuraminidases, YWTD repeats of the low density lipoprotein receptor, and the type III modules of fibronectin. J. Exp. Med. 174:179-191.

Pfaffenbach, G., Lamm, M.E. and Gigli, I. (1982)
Activation of the guinea pig alternative complement pathway by mouse IgA immune complexes.J. Exp. Med. 155:231-247.

Pierce-Cretel, A., Pamblanco, M., Strecker, G., Montreuil, J. and Spik, G. (1981)
Heterogeneity of the glycans *O*-glycosidically linked to the hinge region of secretory immunoglobulins from human milk. Eur. J. Biochem. 114:169-178.

Platt, T. and Bear, D. (1983)
The role of RNA polymerase, Rho factor, and ribosomes in transcription termination. In, Gene function in prokaryotes. J. Beckwith, J. Davies, and J. Gallant. Cold Spring Harbor, N.Y.:Cold Spring Harbor Laboratory. 123-162.

Plaut, A.G. (1983)
The IgA1 proteases of pathogenic bacteria. Ann. Rev. Microbiol. 37:603-622.

Plaut, A.G., Gilbert, J.V. and Wistar Jr., R. (1977)
Loss of antibody activity in human immunoglobulin A exposed to extracellular immunoglobulin A proteases of *Neisseria gonorrhoeae* and *Streptococcus sanguis*. Infect. Immun. 17(1):130-135.

Plaut, A.G., Genco, R.J. and Tomasi Jr., B. (1974)
Isolation of an enzyme from *Streptococcus sanguis* which specifically cleaves IgA. J. Immunol. 113(1):289-291.

Pohlner, J., Halter, R., Beyreuther, K. and Meyer, T.F. (1987)
Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature 325:458-462.

Pratt, S.A. and Boulnois, G.J. (1987)
Molecular cloning of the immunoglobulin A1 protease gene from *Streptococcus pneumoniae*. In, *Streptococcal genetics*. Ferretti, J.J. and Curtiss, R., III. Eds., American Society of Microbiology, Washington D.C. 177-180.

Pratt, S.A. (1988)
Immunoglobulin A1 protease of *Streptococcus pneumoniae*. Doctoral Thesis. Departmet of Microbiology.University of Leicester. U.K.

Prioli, R.P., Mejia, J.S., Aji, T., Aikawa, M. and Pereira, M.E.A. (1991)
Trypanosoma cruzi: localization of neuraminidase on the surface of trypomastigotes. Trop. Med. Parasitol. 42:146-150.

Prudhomme, M., Martin, B., Méjean, V. and J.P. Claverys. (1989)
Nucleotide sequence of *Streptococcus pneumoniae hex*B mismatch repair gene: homology of HexB to MutL of *Salmonella typhimurium* and to PMS1 of *Saccharomyces cerevisiae*. J. Bacteriol. 171:5332-5338.

Reed, K.J., van Epps, D.E. and Williams Jr, R.C. (1979)
Inhibition of human esinophil chemotaxis by IgA paraproteins. Inflammation 3(4):405-416.

Reinholdt, J., Tomana, M., Mortensen, S.B. and Kilian, M. (1990)
Molecular aspects of immunoglobulin A1 degradation by oral streptococci. Infect. Immun. 58(5):1186-1194.

Riley, I.D. and Douglas, M. (1981)
An epidemiologic approach to pneumococcal disease. Rev. Infec. Dis. 3(2):233-245.

Ripley-Petzoldt, M.L., Giebink, G.S., Juhn, S.K., Aeppli, D., Tomasz, A. and Tuomanen, E. (1988)
The contribution of pneumococcal cell wall pathogenesis of experimental otitis media. J. Infect. Dis. 157(2):245-255.

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

Robertson, J., Caldwell, J.R., Castle, J.R. and Waldman, R.H. (1976)
Evidence for the presence of components of the alternative pathway of complement activation in respiratory secretions. J. Immunol. 117:900-903.

Rodriguez, R. and Chamberlin, M. Eds. (1982)
Promoters: structure and function, New York: Praeger.

Roggentin, P., Rothe, B., Lottspeich, F. and Schauer, R. (1988)
Cloning and sequencing of a *Clostridium perfringens* sialidase gene. FEBS Letters 238(1):31-14.

Roggentin, P., Rothe, B., Kaper, J.B., Galen, J., Lawrisuk, L., Virm, E.R. and Schauer, R. (1989)
Conserved sequences in bacterial and viral sialidases. Glycoconjugate 6:349-353.

Romanovska, E. and Watkins, W.M. (1963)
Fractionation of neuraminidases in extracts of *Trichomonas foetus*. Biochem. J. 87:37-38.

Rosenberg, A. and Schengrund, C.-L.(1976)
Circulating sialyl compounds. In, Biological role of sialic acids. A. Rosenberg & C.-L. Schengrund Ed. Pennsylvania State University, Hershey. Pennum Press. New York & London 275-294.

Rothe, B., Roggentin, P., Frank, R., Blöcker, H. and Schauer, R. (1989)
Cloning, sequencing and expression of a sialidase gene from *Clostridium sordellii* G12. J. Gen. Microb. 135:3087-3096.

Rothe, B., Rothe, B., Roggentin, P. and Schauer, R. (1991)
The sialidase gene from *Clostridium septicum*: cloning, sequencing, expression in *Escherichia coli* and identification of conserved sequences in sialidases and other proteins. Mol. Gen. Genet. 226:190-197.

Russel, B.E. and Sherwood, N.P. (1948)
Studies on the Streptococci. II. The role of hyaluronidase in experimental streptococcal infection. J. Infect. Dis. 84:81-87.

Russo, T.A., Thompson, J.S., Godoy, V.G. and Malamy, M.H. (1990)
Cloning and expression of the *Bacteroides fragilis* TAL2480 neuraminidase gene, *nanH*, in *Escherichia coli*. J. Bacteriol. 172(5):2594-2600.

Saito, H., and Miura, K.I. (1963)
Preparation of transforming deoxyribonucleic acid by phenol treatment. Biochim. Biophis. Acta. 72:619-629.

Sánchez-Puelles, J.M., Ronda, C., García, J.L., García, P., López, R. and García, E. (1986)
Searching for autolysin functions. Characterization of a pneumococcal mutant deleted in the lytA gene. Eur. J. Biochem. 158:289-293.

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

Sarnaik, S., Kaplan, J., Schiffmann, G., Bryla, D., Robbins, J.B. and Schneerson, R. (1990)
Studies on pneumococcus vaccine alone or mixed with DTP and on pneumococcus type 6B and *Haemophilus influenzae* type b capsular polysaccharide-tetanus toxoid conjugates in two-to five-year-old children with sickle cell anaemia. Ped. Infect. Dis. J. 9:181-186.

Scanlon, K.L., Diven, W.F. and Glew, R.H. (1989)
Purification and properties of *Streptococcus pneumoniae* neuraminidase. Enzyme 41:143-150.

Schaffner, A., Michel-Harder, C. and Yeginsoy, S. (1991)
Detection of capsular polysaccharide in serum for the diagnosis of pneumococcal pneumonia: clinical and experimental evaluation. J. Infect. Dis. 163:1094-1102.

Schauer, R. (1982)
Sialic acids. Adv. Carbohydr. Chem. Biochem. 40:131-234.

Schloemer, R.H. and Wagner, R.R. (1974)
Sialoglycoprotein of vesicular stomatitis virus. Role of the neuraminic acid in infection. J. Virol. 14:270.

Schneerson, R., Barrera, O., Sutton, A. and Robbins, J.B. (1980)
Preparation, characterization, and immunogenicity of *Haemophilus influenzae* type b polysaccharide-protein conjugates. J. Exp. Med. 152:361-376.

Schneewind, O., Jones, K.F. and Fischetti, V.A. (1990)
Sequence and structural characterization of the trypsin-resitant T6 surface protein of group A streptococci. J. Bacteriol. 172:3310-3317.

Schneewind, O., Pancholi, V. and Fischetti, V.A. (1991)
Surface proteins from Gram-positive cocci have a common motif for membrane anchoring. In, Genetics and molecular biology of Streptococci, Lactococci and Enterococci. G.M. Dunny, P.P. Cleary and L.L. McKay Ed. American Society for Microbiology. 152-154.

Sestini, P., Nencioni, L., Villa, L., Boraschi, D. and Tagliabue, A. (1987)
Antibacterial activity against *Streptococcus pneumoniae* by mouseu lung lymphocytes. Adv. Exp. Med. Biol. 216:517-525.

Sestini, P., Nencioni, L., Villa, L., Boraschi, D. and Tagliabue, A. (1988)
IgA-driven antibacterial activity against *Streptococcus pneumoniae* by mouse lung lymphocytes. Am. Rev. respir. Dis. 137:138-143.

Seto, H. and Tomasz, A. (1976)
Calcium-requiring step in the uptake of deoxyribonucleic acid molecules through the surfaces of competent pneumococci. J. Bacteriol 126:1113-1118.

Shapiro, E.D., Berg, A.T., Austrian, R., Schroeder, D., Parcells, V., Margolis, A., Adair, R.K. and Clemens, J.D. (1991)

The protective efficacy of polyvalent pneummococcal vaccine. New. Engl. J. Med. 325(21):1453-1460.

Shine, J. and Dalgarno, L. (1974)
The 3'-terminal sequence of *E. coli* 16S rRNA: complementary to nonsense triplets and ribosome binding sites. Pro. Nat. Acad. Sci. 71:1342-1346.

Shoberg, R.J. and Mulks, M.H. (1991)
Proteolysis of bacterial membrane proteins by *Neisseria gonorrhoeae* type 2 immunoglobulin A1 protease. Infect. Immun. 59(8):2535-2541.

Shoemaker, N.B. and Guild, W.R. (1974)
Destruction of low efficiency markers is a slow process occuring at heteroduplex stage of transformation. Mol. Gen. Genet. 128:283-290.

Sibille, Y., Delacroix, D.L., Merill, W.W., Chatelain, B. and Vaerman, J.P. (1987)
IgA-induced chemokinesis of human polymorphonuclear neutrophils: requirement of their Fc-alpha receptor. Mol. Immunol. 24(6):551-559.

Signas, C., Raucci, G., Jonsson, K., Lindgren, P., Anatharamaiah, G.M., Hook, M. and Lindberg, M. (1989)
Nucleotide sequence of the gene for fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. Proc. Natl. Acad. Sci. Usa. 86:699-703.

Silhavy, T.J., Benson, S.A. and Emr, S.D. (1983)
Mechanisms of protein localization. Microb. Rev. 47:313-344.

Simmons, R.L., Rios, A., Lundgren, G., Ray, P.K., McKhann, C.F. and Haywood, G.R. (1971)
Immunospecific regression of methylcholanthrene fibrosarcoma with the use of neuraminidiase. Surgery 70(1):38-46.

Simpson, D.A., Hausinger, R.P. and Mulks, M.H. (1988)
Purification, characterization, and comparison of the immunoglobulin A1 protease of *Neisseria gonorrhoeae*. J. Bacteriol. 170(4):1866-1873.

Smith, D.J., Taubman, A.A. and Ebersole, J.L. (1982)
Effect of local immunization with glysocyltransferase on colonization of hamsters by *Streptococcus mutants*. Infect. Immun. 37:656-661.

Smyth, C.J. and Duncan, J.L. (1978)
Thiol-activated (oxygen-labile) cytolysins. In, Bacterial toxins and cell membranes. J. Jeljaszewicz and T. Wadstrom (ed). Academic Press, Inc. (London), Ltd., London. 129-183.

Sondag-Thull, D., Levene, N.A., Levene, C., Manny, N., Liew, Y.W., Bird, G.W.G., Schechter, Y., François-Gérard, Ch. and Blanchard, D. (1989)
Characterization of a neuraminidase from *Corynebacterium aquaticum* responsible for Th polyagglutination. Vox sanguinis 57:193-198.

Southern, E.M. (1975)
Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 89:503-517.

Sørensen, U.B.S., Henrichsen, J., Chen, H.-C. and Szu, S.C. (1990)
Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of *Streptococcus pneumoniae* revealed by immunochemical methods. Microb. Pathog. 8:325-334.

Stahl, W.L. and O'Toole, R.D. (1972)
Pneumococcal neuraminidase: purification and properties. Biochim. Biophys. Acta 268:480-487.

Stassi, D.L. and Lacks, S.A. (1982)
Effect of strong promoters on the cloning in *E. coli* of DNA fragments from *Streptococcus pneumoniae*. Gene 18:319-328.

Steinfort, C., Wilson, R., Mitchell, T., Feldman, C., Rutman, A, Todd, H., Sykes, D., Walker, J., Saunders, K., Andrew, P.W., Boulnois, G.J. and Cole, P.J. (1989)
Effect of *Streptococcus pneumoniae* on human respiratory epithelium in vitro. Infect. Immun. 57(7):2006-2013.

Sternberg, G.M. (1881)
A fatal form of septicaemia in the rabbit, produced by subcutaneous injection of human saliva. An experimental research. National Board of Health Bulletin 2:781-783.

Stenberg, P.E., Levin, J. Baker, G., Mok, Y. and Corash, L. (1991)
Neuraminidase-induced thrombocytopenia in mice: effects on

thrombopoiesis. J. Cell. Physiol. 147:7-16.

Takeshita, T., Takagaki, M., Amano, A., Murakami, Y., Nagata, H. and Shizukuishi, S. (1991)
Purification and properties of neuraminidase from the culture-medium of *Bacteroides loescheii*. Oral MIcrob. Immunol. 6(5):316-319.

Talkington, D.F., Crimmis, D.L., Voellinger, D.C., Yother, J. and Briles, D.E. (1991)
A 43-Kilodalton pneumococcal surface protein, PspA: isolation, protective abilities, and structural analysis of the amino-terminal sequence. Infect. Immun. 59(4):1285-1289.

Tanenbaum, S.W., Gulbinsky, J., Katz, M. and Sun, S.C. (1970)
Separation, purification and some properties of pneumococcal neuraminidase isoenzymes. Biochim. Biophys. Acta 198:242-254.

Tanenbaum, S.W. and Sun, S.C. (1971)
Some molecular properties of pneumococcal neuraminidase isoenzymes. Biochim. Biophys. Acta 229:824-828.

Tenovuo, J., Moldoveanu, Z., Mestecky, J., Pruitt, K.M. and Rahemtulla, B-.M. (1982)
Interaction of specific and innate factors of immunity: IgA enhances the antimicrobial effect of the lactoperoxidase system against *Streptococcus mutants*. J. Immunol. 128(2):726-731.

Tomasi, T.B. (1984)
The secretory immune system. In, Basic and Clinical Immunology 5th Ed. D. Stites, J.D. Stobo, H.H. Fudenberg & J.V. Wells Eds. Lange Medical Publications. Los Altos, California 187-196.

Tomasz, A., Moreillon, P. and Pozzi, G. (1988)
Insertional inactivation of the major autolysin gene of *Streptococcus pneumoniae*. J. Bacteriol. 170(12):5931-5934.

Towbin, H., Staehelin, T. and Gordon, J. (1979)
Electrophoretic transfer of proteins from polysaccharide gels to nitrocellulose sheets: procedure and some applications. Proc. Nat. Acad. Sci. USA. 76:4350-4354.

Tramont, E.C., Ciak, J., Boslego, J., McChesney, D.G., Brinton, C.C. and

Zollinger, W. (1980).
Antigenic specificity of antibodies in vaginal secretions during infection with *Neisseria gonorrhoeae*. J. Infect. Dis. 142:23-31

Troy II, F.A. (1992)
Polysialilation: from bacteria to brains. Glycolbiol. 2(1):5-23.

Tsuji, T., Alborno, R.M., Ehara, M., Honda, T. and Miwatani, T. (1989)
Detection of IgA1 protease from *Haemophilus influenzae* by immunoblotting. Eur. J. Epidemiol. 5(2):199-201.

Tuomanen, E., Liu, H., Hengstler, B., Zak, O. and Tomasz, A. (1985)
The induction of meningeal inflammation by components of the pneumococcal cell wall, J. Infect. Dis. 151(5):859-867.

Uchida, Y., Tsukada, Y and Sugimori, T. (1979)
Enzymatic properties of neuraminidases from *Arthrobacter ureafaciens*. J. Biochem. 86:1573-1585.

Uzman, L.L. and Rumley, M.K. (1956)
Neuraminic acid as a constituent of human cerebrospinal fluid. Proc. Soc. Exp. Biol. Med. 93:497.

Vaerman J-.P. and Heremans, J.F. (1966)
Subclasses of immunoglobulin A based on differences in the alpha polypeptide chains. Sience 153:647-649.

van Epps, D.E., Reed, K. and Williams, Jr. R.C. (1978)
Suppression of human PMN bactericidal activity by human IgA paraproteins. Cell. Immunol. 36:363-376.

van Epps, D.E., Plaut, A., Bernier, G.M. and Williams, Jr. R.C. (1980)
IgA paraprotein inhibition of human neutrophil chemotaxis. Reduced activity following treatment with IgA-specific protease from *Neisseria gonorrhoeae*. Inflam. 4:137-145.

Varki, A. and Diaz, S. (1983)
A neuraminidase from *Streptococcus sanguis* that can release *O*-acetylated sialic acids. J. Biol. Chem. 25:12465-12471.

Vered, M., Simon, S.R. and Janoff, A. (1985)

Subcellular location and further characterization of a new elastase inhibitor from the pneumococci. Infect. Immun. 49(1):52-60.

Vimr, E.R., Lawrisuk, L., Galen, J. and Kaper, J.B. (1988)
Cloning and expression of *Vibrio cholerae* neuraminidase gene *nanH* in *Escherichia coli*. J. Bacteriol. 170(4):1495-1504.

Von Heijne, G. (1983)
Patterns of amino acids near signal-sequence cleavage sites. J. Biochem. 133:17-21.

Von Nicolai, H., Müller, H.E. and Zilliken, F. (1980)
Purification and properties of two neuraminidases from *Streptococcus viridans* II. FEBS letters 117(1):107-110.

Walker, J.W., Allen, R.L., Falmagne, P., Jonhson, M.K. and Boulnois, G.J. (1987).
Molecular cloning, characterisation and complete nucleotide sequence for the gene for pneumolysin, the sulphydryl activated toxin of *Streptococcus pneumoniae*. Infect. Immun. 55:1184-1189.

Walker, J.A., Allen, R.L. and Boulnois, G.J. (1987b)
Towards a genetic analysis of the virulence of *Streptococcus pneumoniae*. In, Streptococcal Genetics. J.J. Ferretti & R. Curtis Eds. American Society for Microbiology 165-171.

Walker, T.A. (1988)
Studies of pneumolysin, the membrane damaging toxin of *Streptococcus pneumoniae*. University of Leicester. U.K. PhD. Thesis.

Waltman, W.D., McDaniel, L.S., Gray, B.M. and Briles, D.E. (1990)
Variations in the molecular weight of PspA pneumococcal surface protein A) among *Streptococcus pneumoniae*. Microb. Pathog. 8:61-69.

Warner, T.G., Harris, R., McDowell, R. and Vimr, E.R. (1992)
Photolabelling of *Salmonella typhimurium* LT2 sialidase. Identification of a peptide with a predicted structural similarity to the active sites of influenza-virus sialidases. Biochem. J. 285:957-964.

Watson, D.A. and Musher, D.M. (1990)
Interruption of capsule production in *Streptococcus pneumoniae* serotype 3 by insertion of transposon Tn*916*. Infect. Immun. 58(9):3135-3138.

Weiss, L. (1963)
Studies on cellular adhesion in tissue-culture. V. some effects of enzymes on cell-detachment. Exp. Cell Res. 30:509.

Weiss, L. (1965)
Studies on cell deformability. I. Effect of surface charge. J. Cell Biol. 2:735-739.

Wenger, J.D., Hightower, A.W., Facklam, R.R., Gaventa, S., Broome, C.V. and the Bacterial Meningitis Study Group. (1990).
Bacterial meningitis in the United States, 1986: Report of a multistate surveillance study. J. Infect. Dis. 162:1316-1323.

Wickner,S. (1978)
Conversion of phage single-stranded DNA to duplex DNA in vitro. In, The single stranded DNA phages. D.T. Denhardt, D. Dressler and D.S. Ray Ed. Cold Spring Harbor, New York. 255-271.

Winkelstein, J.A. and Tomasz, A. (1978)
Activation of the alternative complement pathway by pneumococcal cell wall teichoic acid. J. Immunol. 120:174-178.

Williams, R.C. and Gibbons, R.J. (1972)
Inhibition of bacterial adherance by the secretory immunoglobulin A: a mechanism of antigen disposal. Science 177:697-699.

Williams, W.W., Hickson, M.A., Kane, M.A., Kendal, A.P., Spika, J.S. and Himman, A.R. (1988)
Immunization policies and vaccine coverage among adults: the risk for missed opportunities. Ann. Intern. Med. 108:616.

Wolfenstein-Todel, C., Frangione, B. and Franklin, E.C. (1972)
Structure of immunoglobulin A. II. Sequence around the hinge region and labile disulfide bonds of an immunoglobulin A2 myeloma proteins. Biochemistry 11(21):3971-3975.

Woodruff, J.J. and Gesner, B.M. (1969)
The effects of neuraminidase on the fate of transfused lymphocytes. J. Exp. Med. 129:551-567.

Yanisch-Perron, C., Viera, J. and Messing, J. (1985)
Improved M13 phage cloning vectors and host strains: Nucleotide sequences of M13mp18 and pUC19 vectors. Gene 33:103-119.

Yeung, M.K. and Fernandez, S.R. (1991)
Isolation of a neuraminidase gene from *Actinomyces viscous* T14V. Appl. Environ. Microb. 57(11):3062-3069.

Yoneda, K. and Coonrod. (1980)
Experimental type 25 pneumococcal pneumonia in rats. Am. J. Pathol. 99:231-238.

Yother, J., Forman, C., Gray, B.M., and Briles, D.E. (1982)
Protection of mice from infection with *Streptococcus pneumoniae* by anti-phosphoryl choline antibodies. 36:184-188.

Yother, J., McDaniel, L.S., Crain, M.J., Talkington, D.F. and Briles, D.E. (1991)
Pneumococcal surface protein A: structural analysis and biological significance. In, Genetics and Molecular Biology of Streptococci, Lactococci, and Enterococci. Eds. G.M. Dunny, P.P. Cleary, L.L. McKay. American Society for Microbiology 88-91.

Zwahlen, A., Nydegger, U.E., Vaudaux, P., Lambert, P.-H. and Waldvogel, F.A. (1982)
Complement-mediated opsonic activity in normal and infected human cerebrospinal fluid: early response during bacterial meningitis. J. Infect. Dis. 145(5):635-646.