Biochemical and Structural Characterization of *Streptococcus pyogenes C5a Peptidase*

Pania Dawn Mouat

2002

Acknowledgements

I dedicate this thesis to the memory of my father Magnus Charles Hay Mouat.

I would like to thank my supervisors Dr Jakki Cooney, Dr Todd Kagawa, and Dr Paul O'Toole for giving me their expertise and the opportunity to be involved in a very interesting research project.

I would especially like to thank Dr Kathryn Stowell who took over supervision for the write-up of my thesis for her encouragement, advice, and support.

I would also like to thank my seven year old daughter, Ruby, for putting up with having a student for a mother for a large part of her childhood, and my mother for her love and support.

Thank you also to the Palmerston North Medical Research Foundation and the Alumni Foundation who awarded scholarships for work on this project.

Abstract

Streptococcus pyogenes, also known as Group A Streptococci, is a common causative agent of bacterial infections of the human upper respiratory tract, skin, and soft tissue. Non-suppurative sequelae of *S. pyogenes* infections include rheumatic fever, rheumatic heart disease, and acute glomerulonephritis. Recently there has been a resurgence of rheumatic fever and rheumatic heart disease as well as an increase in aggressive streptococcal disease such as toxic shock syndrome and necrotizing fasciitis. *S. pyogenes* produce a formidable array of virulence factors, one of which is C5a peptidase. The human complement factor C5a is a potent chemoattractant, macrophage activator, and an anaphylatoxin. The C5a peptidase of both Group A Streptococci and Group B Streptococci cleave C5a within its polymorphonuclear neutrophil binding site rendering it inactive. Mouse infection models have demonstrated a functional C5a peptidase assists colonization by retarding the infiltration of phagocytes to the foci of infection. C5a peptidase is a multidomain cell surface subtilisin-like serine protease (subtilase) with an Asp, His, and Ser catalytic triad. Comparative sequence analysis shows C5a peptidase has considerable sequence homology to *Lactococcus lactis* PrtP, both of which are highly specific endopeptidases. Whereas the subtilisins in general show broad substrate specificity profiles, the cell envelope proteinases of lactic acid bacteria demonstrate remarkable substrate specificity. The greater specificity of the cell envelope proteinases is held attributable to changes in variable regions within the structurally conserved regions and the presence of the A-domain, both of which have been demonstrated to modify specificity in PrtP proteinases. The aims of this project were to study the structural and biochemical properties of the C5a peptidase of *Streptococcus pyogenes*. C5a peptidase and variants were cloned, expressed, purified, and assayed for activity in under-agarose lymphocyte migration assays and *in vitro* digestion assays. Absence of activity was found in a C5a peptidase variant in which the A-domain was absent. Purified recombinant C5a peptidase and derivatives were screened for crystallization conditions. Crystallization conditions were found for recombinant C5a peptidase. To combat both the increasing incidence of *S. pyogenes* associated diseases, and increasing antibiotic resistance, new chemotherapeutic agents are required. This study was designed to elucidate the structural and biochemical basis of the substrate specificity of C5a peptidase, which will assist the design of potent inhibitors of this powerful virulence factor.

Abbreviations

AGN	acute glomerulonephritis
bp	base pair
BSA	bovine serum albumin
C5P	C5a peptidase
cDNA	complementary DNA
CEP	lactic acid bacteria cell envelope proteinase
CO_2	carbon dioxide
DMSO	dimethyl sulphoxide
dNTPs	deoxynucleotide triphosphates
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
ELISA	enzyme linked immunosorbent assay
FPLC	fast protein liquid chromatography
GAS	group A streptococci
GBS	group B streptococci
HEPES	N-[2-hydodroxyethyl]piperazine-N'-[2-ethane sulfonic acid]
HPLC	high pressure liquid chromatography
IEFPLC	ion exchange FPLC
kb	kilobase
LAB	lactic acid bacteria
MAC	membrane attack complex
MEM	Eagle's minimal essential medium
MS	mass spectrometry
Mw	molecular weight
oligo	oligonucleotide
PAGE	polyacrylamide gel electrophoresis
pAMPSF	4-aminophenylmethane sulfonyl fluoride hydrochloride
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pI	isoelectric point

PMSF	phenyl methyl-sulfonyl fluoride
PMN	polymorphonuclear neutrophil
PNK	polynucleotide kinase
RF	rheumatic fever
RPHPLC	reverse phase HPLC
SDS	sodium dodecylsulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SEFPLC	size exclusion FPLC
SpeB	streptococcal cysteine proteinase
S. pyogenes	*Streptococcus pyogenes*
STSS	streptococcal toxic shock syndrome
TAE	tris acetate EDTA
Tris	tris (hydroxymethyl)-aminomethane
UV	ultra violet light
VR	variable region within structurally conserved catalytic domain of C5P
ZnAc	zinc acetate
ZnCl	zinc chloride

List of figures

Page

List of tables

Table of contents

Chapter 2 Materials and Methods

Chapter 6 Summary and General Discussion

Introduction

1.1 Medically significant Streptococci

Streptococci are Gram-positive bacteria of the *Streptococcaciae* family. Medically significant Streptococci can be categorised either on the basis of haemolysis on blood agar (complete haemolysis – beta; partial haemolysis – alpha; no haemolysis – gamma), or by their Lancefield Group (labelled alphabetically A to S) typing which is designated by the reaction of anti-sera with group-specific cell wall associated carbohydrate antigen (Madigan et al., 1997; Mims et al., 1998).

1.1.1 The rising threat of S. pyogenes infections

Since the 1980s there has been a concurrent resurgence of rheumatic fever and an increase in invasive *S. pyogenes* infections (including streptococcal toxic shock syndrome and necrotizing fasciitis) reported in both the northern and southern hemispheres (Carapetis et al., 1995; Stevens, 1999). Different strains of *S. pyogenes* are associated with these diseases so their increases in incidence are apparently epidemiologically incongruent (Stollerman, 1998b). *S. pyogenes* can acquire transposon-mediated antibiotic resistance to erythromycin, tetracycline, and kanamycin. Furthermore, although bactericidal *in vitro*, penicillin failure in treatment of streptococcal infections is frequently observed (Gillespie, 1998; Stevens, 1999). The molecular mechanisms involved in streptococcal infections are therefore of interest both from a scientific point of view as well as for the development of possible intervention strategies.

1.1.2 Beta-haemolytic streptococci

The most pathogenic streptococci are the beta-haemolytic members: *Streptococcus pyogenes* (Group A Streptococcus or GAS), and *Streptococcus agalactiae* (Group B Streptococcus or GBS). *S. pyogenes* are exclusively human pathogens and are commonly associated with nasopharyngeal infections of children and skin infections of the elderly. Complications connected with *S. pyogenes* infections include the non-suppurative sequelae rheumatic fever (RF) and acute glomerulonephritis (AGN). The diseases associated with *S. pyogenes* are discussed in more detail below. The GBS *S. agalactiae* can cause neonatal meningitis and septicaemia in humans and mastitis in bovines. Other clinically important beta-haemolytic streptococci include streptococci of the Lancefield Groups C and G, which may cause

pharyngitis, and *S. milleri,* which carries the Lancefield Group F or G antigen and has a propensity for abscess formation (Mims et al., 1998).

1.1.3 Alpha-haemolytic streptococci

The most pathogenic alpha-haemolytic streptococci are the *S. pneumoniae*, which are associated with pneumonia, otitis media (middle ear infection), septicaemia and meningitis. Several species of alpha-haemolytic streptococci, which are mostly commensals in the mouth, are known collectively as the 'viridans streptococci'. The viridans member *S. mutans* is associated with dental caries, and several other species are capable of causing bacterial endocarditis (Mims et al., 1998).

1.2 Diseases associated with S. pyogenes infections

The diseases of *S. pyogenes* can be placed in a spectrum: ranging from non-invasive local infections (for example, acute pharyngitis and impetigo) through invasive but localised infections (such as erysipelas) to invasive systemic infections (including necrotizing faciitis, myositis, and Streptococcal Toxic Shock Syndrome (STSS) (Berge, 1997)). *S. pyogenes* is responsible for approximately 10-20% of cases of acute pharyngitis and is the most common cause of bacterial pharyngitis. Whilst generally trivial and effectively treated with penicillin complications of *S. pyogenes* throat infections include peritonsillar abscess (quinsy), otitis media, scarlet fever (characterised by an erythematous rash caused by a lysogenic phage encoded erythrogenic toxin), and the non-suppurative sequelae rheumatic fever (RF), rheumatic heart disease, and acute glomerulonephritis (AGN). RF is an indirect complication in which antibodies formed to antigens in the streptococcal cell wall cross-react with the sarcolemma of the heart, which leads to the formation of granulomas (called Aschoff bodies) and the development of myocarditis or pericarditis. Rheumatic heart disease develops when repeated infections with *S. pyogenes* of different M-types (see section *Antiphagocytic factors*) leads to damage of the heart valves. Circulating immune complexes, composed of *S. pyogenes* antigens and host neutralising antibodies, can cause AGN when these complexes are deposited in the glomeruli, and subsequent activation of the complement and inflammatory systems leads to kidney damage (Mims et al., 1998).

1.3 Virulence factors of S. pyogenes

S. pyogenes is armed with a diverse array of virulence factors These assist the bacterium to evade the host non-specific (innate) and specific (acquired) immune defences (Berge, 1997).

1.3.1 Secreted proteins

S. pyogenes secretes many virulence factors. These include the streptococcal pyrogenic exotoxins: SpeA SpeC and SpeH, which are superantigens in that they can cause non-specific clonal proliferation of T-cells via interaction with the major histocompatibility complex (MHC) class II. The amplified immunological response caused by the superantigens can result in toxic shock. SpeB is a cysteine protease, which contributes to virulence through cleavage of host connective tissues, and by the release of other virulence factors which are cell-bound (Berge, 1997). Other secreted proteins that contribute to virulence include the streptolysins O and S, which form pores in host cells causing cytolysis; hyaluronidase, which cleaves the connective tissue hyaluronan to assist spreading; streptokinase, which activates plasminogen assisting penetration of biological barriers; and various hydrolytic enzymes such as DNAase and esterase which breakdown host macromolecules [(Berge, 1997; Mims et al., 1998).

1.3.2 Phagocytosis and the complement system

To understand more fully the functions of the virulence factors of *S. pyogenes,* a basic knowledge of phagocytosis and the complement system is required.

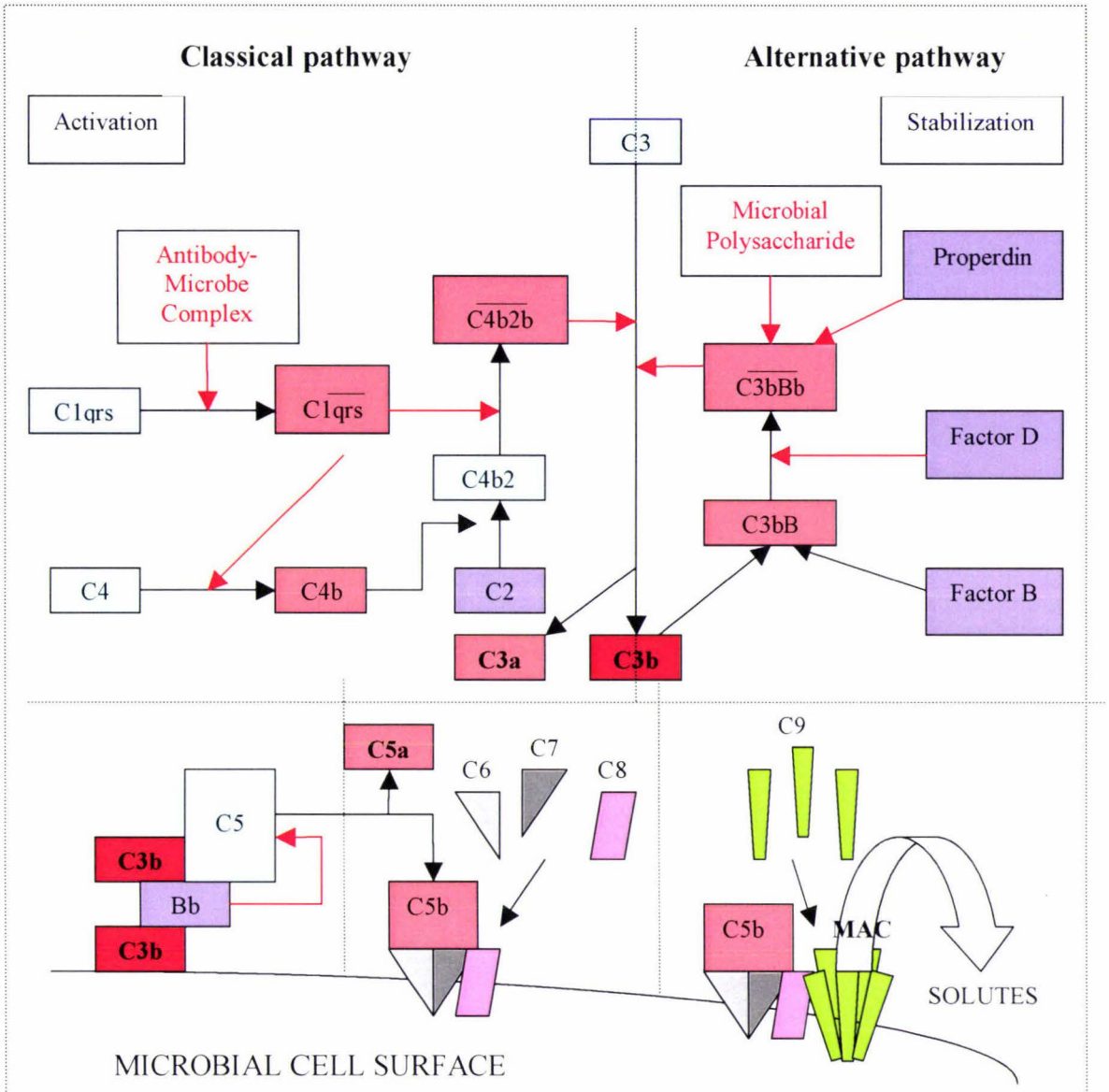
I Phagocytosis

Phagocytosis is performed by macrophages and polymorphonuclear neutrophils (PMNs). The phagocyte must first adhere to the microbe, which may be mediated by recognition of carbohydrate moieties, or through binding of C3b and iC3b deposited on the microbial surface for which the phagocyte has receptors. The factors C3b and iC3b are opsonins, which by definition are substances that enhance phagocytosis by promoting adhesion of an antigen to a phagocyte (immunoglobulins are also opsonins). Attachment triggers the engulfment phase, which proceeds by an actin-myosin contractile system that pulls the plasma membrane around the microorganism to form a phagosome. Cytoplasmic granules fuse with the phagosome; these granules together with reactive oxygen intermediates formed in a respiratory burst triggered during engulfment, and reactive nitrogen intermediates, are responsible for killing the captive microbe. However, before this process can proceed,

phagocytic cells must be recruited from the bloodstream to the site of infection. Recruitment has two requirements: increased permeability of the capillary endothelium and a chemical gradient for chemotaxis, both of which are mediated by the activities of complement factor fragments C3a and C5a. These fragments are released during the complement cascade (Roitt, 1997).

II The complement cascade

The complement cascade consists of classical and alternative pathways illustrated in figure 1.1. The following is an overview of the alternative complement cascade (the classical complement cascade follows a different pathway to the generation of a C3 convertase, but proceeds in parallel from this point): The complement factor C3 undergoes slow spontaneous cleavage of an internal thiolester bond to form C3b. In the presence of Mg^{2+} C3b can bind to complement factor B, which is then cleaved by the enzyme factor D to generate the C3 convertase: C3bBb. This complex is unstable as C3b also has affinity for an inhibitor of the complement cascade (factor H), but can be stabilized by adherence to a microbial surface. The C3 convertase splits C3 into C3a and C3b. Recruitment of a further molecule of C3b forms a C5 convertase, which splits C5 into C5a (a 74 residue fragment (Fernandez and Hugli, 1978)) and C5b. The C5b is incorporated into the enzymic complex and sequential attachment of C6 and C7 to C5b forms a complex with a membrane-binding site to which factor C8 is recruited. Factor C8 produces conformational changes in factor C9 that transform it into an amphipathic molecule, which inserts itself into the microbial membrane. The C9 molecules polymerise to form a transmembrane channel which is permeable to electrolytes and water. Factors C5b-9 constitute the membrane attack complex (MAC). The microbe is lysed by a sudden influx of Na^{+} and water through the MAC channel (Berge, 1997; Roitt, 1997).



The complement cascade can be activated by both the classical or alternative pathways. Activation by the classical pathway is initiated by the binding of antibodies specifically to microbial antigens. The presence of the antibody-antigen complex leads to the conversion of C1qrs to *C1qrs*. *C1qrs* acts on the components C4 and C2 to generate many molecules of *C4b2b, which* is a C3 splitting enzyme. Activation by the alternative pathway is initiated by stabilizing spontaneously formed C3b on the polysaccharide surface of the microbe. C3b complexes with factor B, which then undergoes cleavage by factor D to generate *C3bBb*. This complex is stabilized by properdin. *C3bBb* is a C3 convertase which splits C3 to give C3a and C3b. Recruitment of another C3b molecule into the *C3bBb* complex generates C5 convertase which activates C5 by proteolytic cleavage generating C5a, which is released and C5b. Factors C6 and C7 attach to C5b. This complex binds to factor C8 which directs conformation chanes in C9 to transform it into an amphipathic molecule. C9 polymerizes to form the membrane attack complex which forms a transmembrane channel fully permeable to electrolytes and water.

Figure 1.1 The complement cascade

III The complement factors C3a and C5a

The complement factors C3a and C5a are anaphylotoxins and can therefore trigger the release of chemical mediators from mast cells. These mediators, as well as C3a and C5a, have chemotactic properties and phagocytic cells are recruited by moving up the concentration gradient of these molecules emanating from the focus of infection. These molecules also stimulate capillary dilation and exudation of plasma proteins and fluid, which together with the accumulation of the phagocytes is collectively known as the acute inflammatory response. The factor C5a in particular is a particularly potent phagocyte chemotactic agent and vasodilator of the capillary endothelium. Both C3a and C5a can also stimulate the respiratory burst and the expression of receptors for C3b and iC3b in phagocytes (Roitt, 1997).

1.3.3 Antiphagocytic factors

Resistance to phagocytosis has been described as the hallmark of virulent *S. pyogenes* and these organisms have evolved overlapping mechanisms to meet this end (Ji et al., 1998). The cell surface of virulent strains of *S. pyogenes* is covered in hair-like projections, called the M protein, which is arranged as an alpha-helical coiled-coil dimer (Philips et al., 1981). The M proteins are remarkably heterogeneous and immunologically diverse. They are used as a typing scheme for epidemiological purposes, for which greater than 90 serotypes have been identified (Stollerman, 1998a). Another cell surface protein used to classify *S. pyogenes* is the serum opacity factor (OF), which is believed to contribute to virulence through binding to fibronectin. All isolates of a certain M serotype are either OF^+ or OF^- (Berge, 1997). M proteins have been suggested to contribute to virulence of the bacterium through either binding of fibrinogen (Whitnack and Beachey, 1995), which covers antigenic epitopes and reduces deposition of the complement factor C3b, or by binding of the complement factor H (Hortsmann et al., 1992), which is an inhibitor of the alternative complement pathway. Serotype-specific neutralizing antibodies react with surface displayed hyper-variable amino acids at the N-terminus of the M proteins. This counteracts the antiphagocytic properties of the M proteins and opsonises the bacterium. *S. pyogenes* also express M-like proteins that have homology to the M proteins and also have antiphagocytic properties (Hollingshead et al., 1993). The M-like proteins are encoded on the *mga* regulon (see section 1.5). Another virulence factor of *S. pyogenes* that has antiphagocytic properties is the hyaluronic acid

capsule. This capsule is chemically equivalent to the hyaluronate of human connective tissue so antibodies to it are difficult to raise. The bacteria are shielded by the capsule from phagocytic cells, which are hence unable to bind to C3b deposited on the bacterial surface (Wessels et al., 1991). The multi-domain cell surface protein C5a peptidase (C5P) is another member of the phagocytic defence (Ji et al., 1998). Whereas the M protein and the hyaluronic acid capsule limit the uptake of streptococci, the C5P protein reduces recruitment and activation of phagocytes.

1.4 The contribution of C5P to S. pyogenes virulence

Virulent strains of Group A, Group B, and Group G streptococci produce a cell surface bound C5a peptidase (Cleary et al., 1991; Hill et al., 1988; O'Connor and Cleary, 1986b). C5P cleaves C5a within the PMN binding site (Cleary et al., 1992; Wexler et al., 1985), which destroys phagocyte activation, as well as the chemotactic and anaphylotoxin potential of C5a. C5P also functions as an invasin (a factor which assists microbial invasion of a host cell) (Beckmann et al., 2002; Cheng et al., 2002).

1.4.1 The target cleavage site of C5P

Sequencing of the 74 residue C5a after incubation with C5P, extracted by a limited trypsin digest of *S. pyogenes* strain M49T14 (135kDa), initially suggested a six-residue peptide was lost from the carboxyl terminal of C5a (Wexler et al., 1985). These authors later amended this when the scissile bond was found to be between His67 and Lys68, rather than between Lys68 and Asp69, which would give rise to a seven-residue peptide (Cleary et al., 1992). In this experiment extraction of the C5P was performed with muralytic enzyme phage C lysin, and a synthetic 20-residue peptide of the C-terminal of C5a was used as the substrate. The location of the cleavage site is controversial, as both the method of C5P extraction, and the C5a substrate were different in these two experiments. C5P derived from the tryptic digest ranged in molecular weight from 103,000 to 114,000 Da, whereas native C5P extracted with detergent from the cell wall is 135,000 to 137,000 Da (Wexler and Cleary, 1985). The cell wall spanning and anchoring domains of C5P have 94 and 41 residues, respectively (Siezen, 1999). This suggests trypsin cleaves within the A domain, which is a putative modifier of specificity, of C5P (see section 1.9). It has been proposed that C5P is a highly specific endopeptidase, and that its function is to eliminate the chemotactic and macrophage-activating anaphylatoxin C5a from the foci of infection (Cleary et al., 1992), although evidence in support of specificity is not very substantial.

1.4.2 Identification of C5P and elucidation of its C5a-ase function

Group A streptococci were found to inhibit the human PMN chemotactic response via interference with complement-derived factors in underagarose migration assays (Wexler et al., 1983). This 'Streptococcal Chemotactic Factor Inactivator' (SCFI, later renamed C5P) was extracted from the surface of *S. pyogenes* strain M49 T14 and purified (Wexler and Cleary, 1985). Ligand-receptor binding studies demonstrated that C5P could reduce the ability of C5a to bind to PMN receptors, which was attributed to the cleavage of a small peptide from the carboxyl-terminus (Wexler et al., 1985) (see above). *In vivo* assays of mouse infection models of the peritoneum, comparing wild-type *S. pyogenes* with C5P mutants induced by transposons and nitrosguanidine (O'Connor and Cleary, 1987), suggested C5P contributes to the virulence of *S. pyogenes* by retarding the influx of inflammatory cells and the clearance of streptococci from the foci of infection. This was confirmed with the use of defined C5P mutant strains (generated with recombinant plasmids) in mouse infection models of the air sac (Ji et al., 1996), the nasopharynx (Ji et al., 1997), and the oral mucosa (Ji et al., 1998).

1.4.3 Identification of C5P as an invasin

Some strains of GBS express C5P with little or no C5a-ase activity attributable to a genetic polymorphism (Bohnsack et al., 2000). The expression of C5a-ase irrespective of C5a-ase functional activity suggested that C5P may have a second important function. Antibody against C5P was found to be opsonic (Cheng et al., 2001) which suggested that the cell surface protease may also function as an adhesin or invasin (Cheng et al., 2002). C5P was identified from a phagemid clone, selected on immobilized fibronectin, from a shotgun phage display library constructed from GBS chromosomal DNA (Beckmann et al., 2002). GBS C5P was demonstrated to bind Hep2 and A549 human epithelial cells, and both GAS and GBS C5P were demonstrated to bind fibronectin, in enzyme-linked immunosorbent assays. Rabbit anti-C5P blocked invasion of A549 cells by GBS, but did not inhibit adherence, indicating that C5P can function as an invasin, independent of adherence (Cheng et al., 2002).

1.4.4 Localization of C5P

C5P was localized to the cell surface by immunofluorescent staining using hyperimmune serum raised against purified C5P as the source of primary antibodies. Cell-bound C5P

accounted for 90% of the antigen detected; the other 10% was cell-free C5P (O'Connor and Cleary, 1986a). Streptococcal cysteine proteinase (SpeB) has demonstrated the ability *in vitro* assays to release a biologically active 116-kDa fragment of C5P, which may account for the 10% cell-free C5P detected by O'Connor et al (Berge and L., 1995; O'Connor and Cleary, 1986a).

1.4.5 Antigenicity of C5P

Polyclonal antiserum raised against purified C5P from *S. pyogenes* was shown to neutralize purified C5P anti-chemotactic activity (Wexler and Cleary, 1985). Furthermore, antiserum raised against C5P extracted from strain CS101 was capable of neutralizing cell-bound C5P activity in six different strains of *S. pyogenes*. This suggests that unlike the M proteins, which are serotype specific, C5P may have only one antigenic type (O'Connor and Cleary, 1986a). Intranasal immunization with a recombinant protein fragment of C5P from M49 streptococci stimulated significant levels of IgA and Ig G antibody production, which were shown to inhibit nasapharyngeal colonization of mice by strains representing the major subdivisions of *S. pyogenes* (Ji et al., 1997). Antibodies raised by this M49 C5P fragment were able to neutralize the C5P activity of M1, M6, M11, as well as M49 streptococci, which supports the suggestion that anti-C5P antibodies lack serotype specificity. In Group B streptococci cell-free C5P can be effectively neutralized, however, IgG does not neutralize cell-bound C5P of the encapsulated type III bacterium (Bohnsack et al., 1992). Examination by enzyme linked immunosorbent assay (ELISA) of human sera and saliva found measurable concentrations of anti-C5P IgG and IgA, in samples from healthy adults and convalescent patients of Group A streptococcal pharyngitis, whilst young children that were less likely to have been exposed to *S. pyogenes* lacked anti-C5P immunoglobulins. It was suggested from the results of these assays that antibody response to C5P may account for the partial non-M type-specific immunity that is observed in adults (O'Connor and Darip, 1991).

1.5 The mga regulon

A partial *Sau*3A digest of the *Streptococcus pyogenes* CS24 genome was used to construct a lambda genomic expression library, which was screened for expression of C5P with hyperimmune rabbit serum. The C5P gene was localized to a 4.3-kb fragment from isolated clones by restriction endonuclease and BAL 31 nuclease deletion analysis. Probes designed from the coding sequences of the type 12 M and C5P proteins both hybridised to the same

genomic fragment demonstrating linkage of the two genes (Chen and Cleary, 1989). Expression of C5P and M proteins was shown to be co-regulated at the level of transcription, under the control of the regulatory locus *virR* (later renamed as *mga*), which is located upstream of the M protein coding sequence as illustrated in figure 1.2. It was proposed that these genetic elements constitute a virulence regulon (Simpson et al., 1990). The regulatory protein of these virulence factors, formerly known as Mry or VirR, was found also to activate the expression of other putative virulence factors, and was accordingly renamed Multigene Regulator in GAS (Mga). A 45-bp binding site for Mga (*mga*) overlapping the –35 region in the promoters of the genes encoding both M protein and C5P was located in DNase I protection assays (McIver et al., 1995). The gene encoding another virulence factor, SIC, was located to the *mga* regulon in M1 strains of *S. pyogenes*. SIC is a secreted protein that interferes with the formation of the membrane attack complex of complement (Akesson et al., 1996). An insertion element, IS*1548* was identified in GBS associated with endocarditis, and in ten GAS strains of five different serotypes. In all of the GBS strains harbouring IS*1548*, and some of the GAS strains, the insertion sequence was found to be located between the genes coding for the M protein and C5P (Granlund et al., 1997). Genes of the *mga* regulon exhibit growth phase dependent expression, with maximal expression during the exponential phase, which is shut off as the bacteria enters stationary phase (McIver and Scott, 1997).

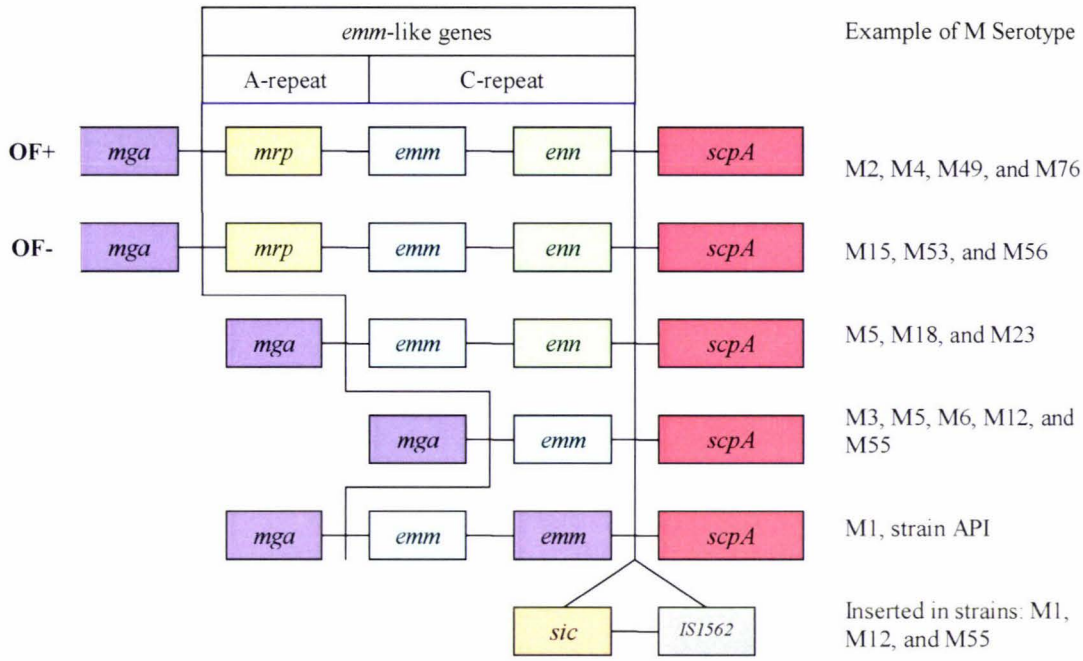


Figure 1-2 The *mga* regulon. **Organization of the *mga* regulons in various strains of *S. pyogenes*(Berge, 1997).**

The genes of the *mga* regulon constitute a pathogenicity island which is a gene locus containing multiple genes involved in virulence. Opacity factor + strains (OF+) harbour three *emm*-like genes, while OF- strains are heterogeneous.

1.6 Serine Proteases

The major classes of proteolytic enzymes are the serine, zinc, cysteine and aspartyl proteases. As C5a peptidase is a serine protease it is important to consider the catalytic mechanism of this family of proteases. All serine proteases have a characteristic serine, histidine, and aspartic catalytic triad (Steitz and Shulman, 1982). The two best studied serine protease families are the (chymo)trypsin and subtilisin classes, which share the catalytic triad but have vastly different protein scaffolds. The catalytic triad is arranged in a scaffold constructed mainly of anti-parallel β-strands in (chymo)trypsin, whereas in subtilisin the scaffold is made up of both α-helices and β-sheets (Dodson and Wlodawer, 1998; Perona and Craik, 1995). This suggests the catalytic triad shared by these serine protease families has arisen by a convergent evolutionary mechanism (Perona and Craik, 1995).

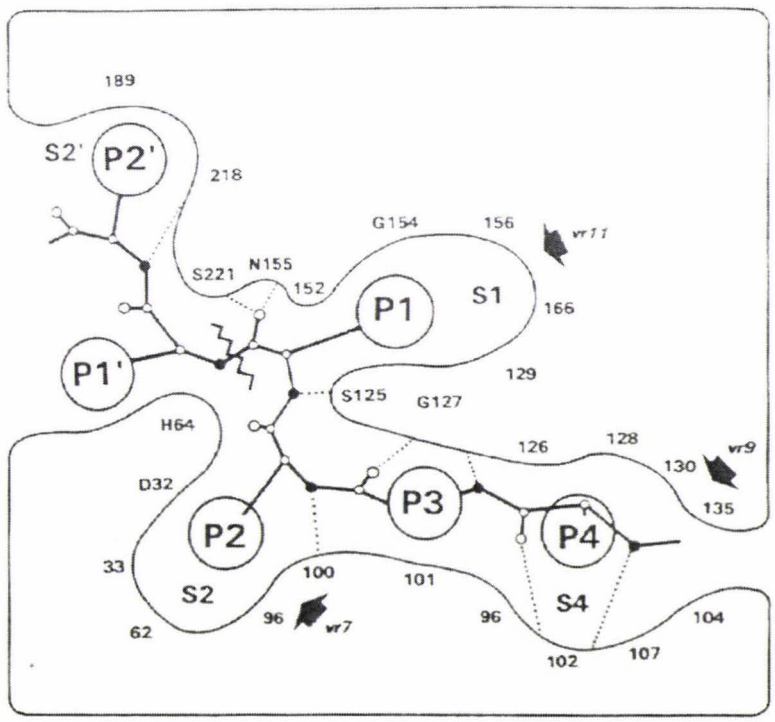
The mechanism of proteolysis occurs in two stages: acylation and deacylation. Acylation proceeds with cleavage of the amide bond of the substrate by nucleophilic attack of the unusually reactive serine protease's seryl Oγ. The nucleophilic potential of the seryl Oγ is generated in the following manner: The carboxylate group of the triads Asp residue forms a hydrogen bond with the adjacent imidazole group of the triads His residue. This hydrogen bonding orientates the His residue in a particular tautomeric form, which together with the increased pKa and basicity of the hydrogen bonded imidazole group, facilitate the acceptance of the seryl Oγ proton. Substrate binding will bury the active centres and cause structural adjustments that favour the protons transfer (Dodson and Wlodawer, 1998). The protonated imidazole group has a positive charge, which is stabilized by electrostatic interaction with the negatively charged carboxyl group of the Asp residue. The negatively charged Oγ performs nucleophilic attack on the carbonyl group C of the susceptible peptide bond. Hydrolysis proceeds through a transient tetrahedral intermediate, which is stabilised by hydrogen bonding between the carbonyl oxygen (which assumes the negative charge and is hence an oxyanion) and two amino groups from the main chain of the protease (or the amine group of an Asn residue) in the oxyanion-hole. Cleavage occurs when the oxyanion reforms a double bond with the carbonyl C, whilst the protonated imidazole ring donates the proton originally from the Ser residue to the N atom of the peptide bond. The carbonyl component of the substrate is hence acylated to the Ser residue, whereas the amine component can diffuse away. The deacylation stage involves deprotonation of a water molecule by the imidazole ring, followed by nucleophilic attack of the carbonyl C atom by the resultant

hydroxyl group. This results in a second tetrahedral transition state, which is resolved when the protonated imidazole ring donates a proton to the seryl Oγ, with concomitant release of the carbonyl component of the substrate. The active site is now reset and ready to hydrolyse another substrate molecule (Kossiakoff and Spencer, 1981).

1.7 The subtilases

C5a peptidase is a subtilase, which is a subtilisin-like serine protease. Subtilases have been found in organisms from the three major kingdoms: the Eukarya, the Archaea, and the Bacteria, which indicates they evolved a long time ago and suggests they may be ubiquitous. Subtilases are all putative endopeptidases or tripeptidases and are mostly extracellular proteins (Siezen and Leunissen, 1997). A structurally conserved subtilase core, consisting of 191 residues, has been defined through multiple alignment of amino acid sequences (including C5P) to identify conserved residues, followed by superposition of known 3D structures to identify conserved conformations or motifs (Siezen et al., 1991).

Figure1.3 The substrate binding site of subtilisin-like serine proteases



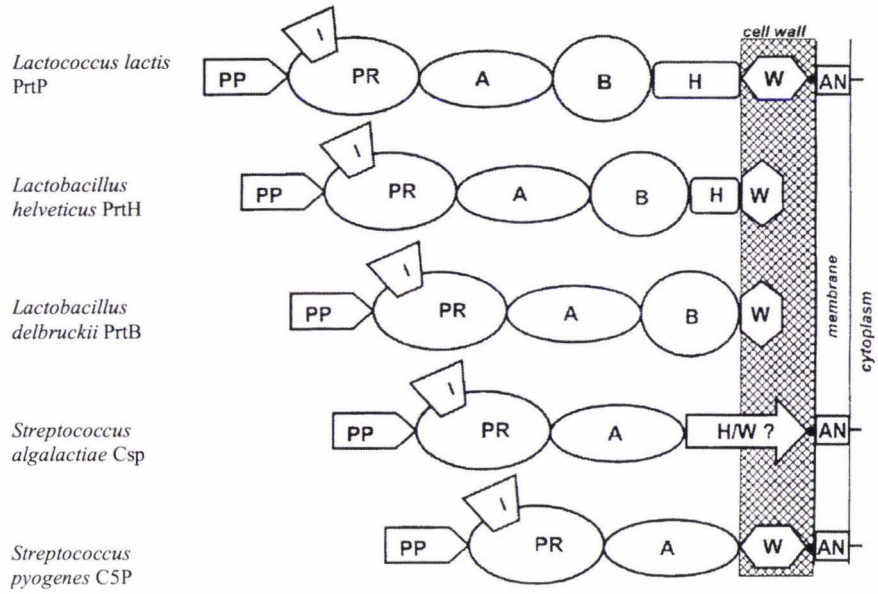
Schematic diagram of substrate (bold lines) in the substrate binding site of a subtilase (smooth surfaces). Side chains of subtilase residues that may interact with P4-P2' side chains are shown with numbering corresponding to subtilisin BPN for which a structure has been solved. The catalytic residues indicated are aspartic acid (D) 32, histidine (H) 64, and serine (S) 221, and the oxyanion-hole residue asparagine (N) 155. Hydrogen bonds are shown as dotted lines. The site of substrate cleavage is represented by a zig-zag line between the scissile bond P1'-P1. Positions of variable regions (*vr*), which are large inserts of sequence within a structurally conserved core, are indicated by large arrows. (Seizen and Leunissen, 1997)

1.7.1 General features of subtilases

Connections between structurally conserved regions (SCRs), which differ considerably between members of the subtilase family, have been termed variable regions (VRs). The VRs generally occur in connecting loops between helices and β-strands at the external surface of the protein (Siezen et al., 1991). Most subtilases are synthesized as pre-pro-enzymes. The 'pre' refers to an amino-terminal signal peptide, which is cleaved during translocation through a cell membrane. Cleavage of the pro-segment is generally required for activation of the subtilase, however this has not been established for C5P. The catalytic domain is located at the amino-terminal end, and is the domain with sequence homology to subtilisin. The substrate/inhibitor-binding region is a surface channel, capable of accommodating at least six amino acid residues, in which both main-chain and side-chain interactions between the enzyme and the substrate/inhibitor contribute to binding as illustrate in figure 1.3. A three-stranded anti-parallel β-sheet is formed between the backbones of the enzyme and the substrate/inhibitor residues P4-P1, with the substrate/inhibitor as the central strand. Hydrogen bonding also occurs between the protease and the leaving residues P1'-P2' but these interactions are weaker than those with residues P4-P1. Interactions between the P4-P1 residues with their S1-S4 binding sites, respectively, mostly determines the specificity of the subtilases (Perona and Craik, 1995; Siezen and Leunissen, 1997). Subtle variations in these substrate-binding regions engineered by formation of hybrid proteinases (Vos et al., 1991), and site-directed mutagenesis (Siezen et al., 1993) have been demonstrated to produce major changes in specificity.

1.7.2 Subtilase families

The subtilises have been subdivided into six families based on characteristic sequence patterns in both SCRs and VRs. These are the subtilisin, thermitase, proteinase K, lantibiotic peptidase, kexin, and pyrolysin families and are designated by the letters A-F respectively. C5P is of the pyrolysin family (Siezen and Leunissen, 1997).



PP=pre-pro domain; PR=protease domain; I=insert domain; A=A-domain; B=B-domain; H=helical domain; W=cell-wall domain; AN=anchor domain.

Most subtilases are synthesized as pre-pro (PP) enzymes. The pre corresponds to a signal sequence that targets the protease for Sec-dependant translocation over the cell membrane, after which the enzyme is activated by cleavage of the pro-peptide. The catalytic or protease domain (Pr) is the segment with sequence homology to the subtilisins within which the subtilases have a few large inserts (I) outside of a structurally conserved core. Substate specificity of the subtilases is held attributable to changes in these variable regions within the structurally conserved core and to the presence of the A-domain (A). Common to many surface proteins of G+ve bacteria are the cell-wall domain (W) that spans the peptidoglycan layer and the cell wall anchor domain (AN).

Figure 1.4 Predicted domains in CEPs of different strains of lactic acid bacteria (Siezen, 1999).

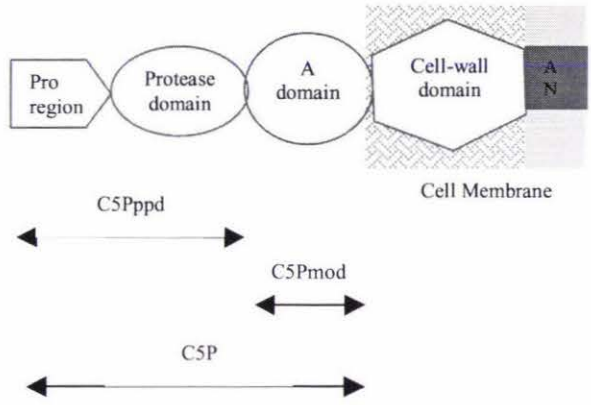
1.8 Cell envelope proteinases of lactic acid bacteria

A dendrogram depicting the homology between catalytic domains of subtilases has shown *S pyogenes* C5P to be closely related to two other cell envelope proteinases (CEPs): *Lactococcus lactis* subspecies *cremoris* SK11 PrtP, and *Lactobacillus delbrueckii* PrtB (Siezen and Leunissen, 1997). Comparative sequence analysis of these CEPs along with *S. agalactiae* Csp, and *Lactobacillus helveticus* PrtH, has suggested the presence of several domains, as illustrated in figure 1.4. In common with other subtalises are the amino-terminal pre-pro-domain (PP) for secretion and activation (101-192 residues), the serine protease domain (PR) with a few large VRs outside of the SCRs (482-512 residues), and the carboxyl-terminal cell-wall attachment (W) (131-214 residues) and cell-wall anchor domains (AN) (36-41 residues). Four residues within the protease domain, Ser^{512}, Asp^{130}, His^{193}, and Asn^{295} were predicted to compose the charge transfer structure of the enzyme (Siezen et al., 1991). Single amino acid substitutions to alanine at any of these sites eliminated at least 99.9% of the proteolytic activity (Stafslien and Cleary, 2000). Novel domains almost exclusive to the CEPs were also identified. Within the protease domain of all five CEPs there is an insert of 134-156 residues, which is classified as the 'I'-domain. Sequence

identity between the CEPs is considerably lower within the I-domain in comparison to the rest of the protease domain (Siezen, 1999). Significantly, the I-domain of *L. lactis* SKII PrtP, whilst demonstrated to be dispensable for protease activity of deletion mutants, was found to influence substrate specificity (Bruinenberg et al., 1994; Siezen, 1999). Located on the carboxyl side of the PR-domain in all five CEPs is a novel domain designated the A-domain, of approximately 400 residues, the sequence of which is well conserved amongst the CEPs. This domain has been speculated to be involved in regulating proteolytic activity and/or specificity of the PR-domain (Siezen, 1999). In support of this, two mutations made in the putative A-domain of *L. lactis* SKII and Wg2 PrtP proteinases were found to affect cleavage specificity (Vos et al., 1991). Furthermore, a highly conserved domain located C-terminally to the catalytic region in all mammalian subtilisin-related pro-protein convertases (SPCs), the P domain, (analogous to the A-domain of the CEPs,) has been deemed essential for the correct folding and the stability of these proteinases (Bergeron et al., 2000; Siezen, 1999). The CEPs PrtB, PrtH, and PrtP also have a novel B-domain of approximately 500 residues, which may play a role in stabilizing the PR-domain. An α-helical domain is predicted near the carboxyl terminal in PrtP and PrtH, which may serve as a spacer to situate the amino-terminal domains away from the cell surface. *S. agalactiae* Csp has a large C-terminal domain for which no function has been ascribed (Siezen, 1999). Whereas the subtilisins in general show broad substrate specificity profiles (Perona and Craik, 1995), the CEPs of lactic acid bacteria demonstrate remarkable substrate specificity. The greater specificity of the CEPs is held attributable to changes within the VRs and the presence of the A-domain (Siezen, 1999). which have been demonstrated to modify specificity in PrtP proteinases (Vos et al., 1991).

1.10 Structural and biochemical characterization of a C5a peptidase from S. pyogenes

The work described in this thesis is based on the hypothesis that both the variable region insertions and deletions, and the A domain influence the specificity of streptococcal C5a peptidase.



Streptococcus pyogenes C5a peptidase (C5P) cloning strategy: Three regions of C5P were cloned, expressed, and purified. These regions were the pro region plus protease domain (C5Pppd); the modifier or A domain (C5Pmod); and the full length of C5a peptidase with the exception of the cell wall spanning domains (C5P).

Figure 1.5 C5P cloning strategy

Aims

- To clone C5P, C5P protease domain plus proregion (C5Pppd), and C5P A domain (C5Pmod) as illustrated in figure 1.5, and synthetic C5a genes.
- To express and purify recombinant C5P, C5Pppd, C5P A domain, and synthetic C5a.polypeptide
- To biochemically characterize the activities of C5P and its derivatives.
- To screen for crystallization conditions for the various C5P fragments.

Cloning of C5P and derivatives

C5P and derivatives, as outlined above, were cloned as Glutathione-S-Transferase (GST) fusion proteins, expressed and purified. The activity of these cloned C5P constructs were determined using under-agarose lymphocyte migration assays and *in vitro* digestion assays of commercially acquired synthetic human C5a and cloned His- C5a. No activity was found in the absence of the A-domain. Biologically active C5P was purified from a mutanolysin digest of the cell wall of *S. pyogenes*. The C5P obtained by this method, however, was not pure enough to be used as a positive control in activity assays of the cloned C5P constructs.

Structural studies

The CEPs are large multi-domain proteins, which can prove extremely difficult to crystallize (Siezen, 1999). C5P is of the pyrolysin family of subtilases for which no 3D

structures are known. Large insertions within the subtilisin structurally conserved core would make homology modelling and protein engineering based on 3D structures of members from other families problematic (Siezen and Leunissen, 1997). Therefore, attempts were made to crystallize the purified domains of C5P individually. Crystals were obtained for C5P, with the exception of the cell wall spanning domains. Crystallization conditions were refined to obtain diffraction quality crystals, which were used in preliminary X-ray diffraction studies. Resolution of the data was beyond the scope of this project. However, if the hypothesis holds true, the structure of C5P should reveal potential interactions of VRs and the A-domain with substrate binding sites and/or the substrate.

Chapter two: Materials and Methods

2.1 Materials

Oligonucleotides were from Sigma-Aldrich, St. Louis, MO, USA or Invitrogen Life Technologies™ Incorporated, MD, USA.

Restriction endonucleases and digestion buffers were from Life Technologies Incorporated, MD, USA and Roche Molecular Biochemicals Mannheim, Germany.

Pwo and Taq polymerase and Expand High Fidelity PCR system were purchased from Roche Molecular Biochemicals, Mannheim, Germany.

T4 DNA ligase and ligation buffer were purchased from Roche Molecular Biochemicals, Mannheim, Germany.

High Pure PCR Product Purification Kit (Concert) was purchased from Invitrogen Life Technologies Incorporated, MD, USA.

QIAGEN Midi-Prep DNA Purification Kit was purchased from QIAGEN Pty Ltd, Australia.

$pP_{RO}EX^{TM}$ HTb vector was purchased from Invitrogen Life Technologies Incorporated, MD, USA.

pGEX-6P-3 vector was purchased from Pharmacia Biotech, Piscataway, NJ 08855, USA.

Ampicillin was from Sigma-Aldrich, St. Louis, MO, USA.

PreScission™ Protease and Glutathione Sepharose™ were from Amersham Pharmacia Biotech, Piscataway, NJ 08855, USA.

Polypeptide SDS-PAGE Molecular Weight Standards and acrylamide were purchased from Bio-Rad laboratories, CA 94547.

HiTrap Q ion exchange sepharose and columns and Bio-Prep SE100/17 columns were from Pharmacia Biotech, Piscataway, NJ 08855, USA.

Minimal Essential Medium was from Technologies Incorporated, Grand Island, NY 14072, USA.

Falcon 3002 tissue culture dishes were from Becton Dickinson Labware, NJ 07035, USA.

Centriprep Centrifugal Filter Units were purchased from Millipore Australia Pty. Ltd. NSW 2113, Australia.

Hampton Research Crystal Screen's 1 and 2 were purchased from Hampton Research, CA, USA.

2.2 Methods

2.2.1 PCR

Primers included 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning. The amplified regions were inserted between the *Bam*HI and *Eco*RI sites of plasmid pGEX6P-3. This produced an in-frame transcriptional fusion of the protein of interest to Glutathione S-Transferase (GST), an affinity tag used for purification. The resultant plasmids were named pGEXC5P (residues 32-1032 of C5P), pGEXC5Pppd (residues 32-583 of C5P), pGEXC5Pmod (residues 584-1032 of C5P), and pGEXC5a (C5a cDNA, 253 residues).

2.2.1.1 *PCR of recombinant C5P fragments*

The coding sequences for three different fragments of the *S. pyogenes* B220 C5P gene were amplified by PCR. The proofreading *Pwo* polymerase or the Expand High Fidelity PCR System, by Roche (composed of an enzyme mix containing both *Taq* DNA polymerase and a polymerase with 3'-5' exonuclease proofreading activity) were used for amplification. Conditions for amplification were: C5Pmod denaturation 94°30", annealing 54°30", extension 72°1 min 20", 8 mM $MgSO_4$; C5Pppd denaturation 94°30", annealing 50°30", extension 72°1 min40" 6 or 8 mM $MgSO_4$; C5P denaturation 94°30", annealing 50°30", extension 72°3 min, 4 or 6 mM MgCl2. All PCR reactions were repeated for 30 cycles. The primers included 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning. Streptococci have an AT rich genomic sequence. This necessitates the use of high amounts of salt in PCR reaction mixtures.

2.2.1.2 *PCR of synthetic human C5a*

The synthetic human C5a coding sequence of 253 bp was generated from the ligation of thirteen phosphorylated oligonucleotides: six in the forward direction and seven in the reverse (supplied by Dr Jakki Cooney). The phosphorylation reaction was performed separately for the forward and reverse oligonucleotides: 1 μL of each oligonucleotide at 50 pmol/μL were added to the phosphorylation reaction mixture (10 μL 5X T4 kinase buffer; 1.5 μL 10 mM ATP). 1 μL T4 kinase and milli-Q H_2O, to adjust the total reaction mixture to 50 μL, were added. The mixture was vortexed gently to mix and incubated for 30 minutes at 37°C. The phosphorylation reactions were heated to 95°C for 5 minutes to stop the reaction. 10 μL of the forward and reverse phosphorylated oligonucleotides were mixed. The oligonucleotides were denatured by heating briefly to 95°C and then cooled to 45°C for 10 minutes to allow for annealing. The mixture was then cooled to room temperature. 1 μL of DNA ligase was added and the ligation mixture was left at room temperature overnight. The ligation product was amplified by PCR with *Pwo* polymerase. This process was repeated several times before a substantial PCR product was obtained. Factors affecting the stringency, including the temperature of annealing and the salt concentration, were adjusted in each PCR reaction. The conditions that gave the cleanest PCR product were denaturation 94°30", annealing 48°30", extension 70°30", repeated for 30 cycles with a salt concentration of 2 mM $MgSO_4$. The primers included 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning.

2.2.2 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA fragments by size for identification, purification, and quantification. At or near physiological pH a DNA molecule carries a net negative charge, conferred by the phosphate groups of its backbone. During electrophoresis an electric field is applied to a solution, within

which the negatively charged DNA molecules migrate towards the positively charged anode. Migration is dependent on the size of the DNA fragments and their conformation. Smaller and compacted DNA molecules have greater electrophoretic mobility as the sieving effect exerted by the agarose matrix is less than for larger and non-compacted molecules.

Agarose gel electrophoresis was performed using 0.8-1.2 % agarose in 1X TAE buffer (40 mM Tris-acetate; 2 mM EDTA pH 8.5), with the percentage of agarose used being dependent on the size and conformation of the DNA molecules being separated. The agarose was melted in a microwave and allowed to cool to approximately 55°C before being poured into a minigel casting apparatus. DNA samples were prepared for electrophoresis with the addition of DNA loading dye (10% glycerol, 62.5 mM Tris pH 8.0; 0.1% bromophenol blue) to 10% of the sample volume. The gel was electrophoresed at approximately 100 V for one hour. Migration of DNA was assessed during electrophoresis by the bromophenol dye front, which migrates with DNA fragments of 500 bp in 1% agarose. The gel was soaked in ethidium bromide solution (0.5 µg/mL) for approximately 20 minutes, and then soaked briefly in water before visualizing by UV transillumination. The size of DNA molecules was determined by comparing mobility with an appropriate molecular size standard electrophoresed on the same gel. The concentration of DNA molecules was estimated by visual comparison of fluorescence intensity relative to standards that contained known quantities of DNA.

2.2.3 Restriction endonuclease digests

Before digestion PCR products were purified with the High Pure PCR Product Purification Kit. DNA is adsorped onto the glass fibre fleece filter of the purification kit column. The binding process is specific for DNA and impurities (salts, free nucleotides, and proteins) can be removed by washing the column. The purified DNA

was eluted from the glass fiber fleece with water. Digestion reactions were performed in a total volume of 50-150 µL. 1-20 µg of DNA was digested using approximately 2 units of restriction enzyme per µg of DNA. The restriction endonuclease buffer provided with the restriction enzyme (diluted to 1x) was used for the digest. When two restriction enzymes were used in a digestion, the buffer that would provide optimal DNA cleavage by both enzymes was used. Digestions were performed for $1^1/_2$ hours at 37°C. The plasmid or PCR digestion products were purified with the High Pure PCR Product Purification Kit.

2.2.4 Ligation of inserts into vectors

Vector to insert ratios of 1:1 and 1:4 were ligated at room temperature overnight with 1 unit of T4 DNA ligase per µg of DNA and 1x ligation buffer (66 mM Tris-HCl; 10 mM dithioerythritol; 50 mM $MgCl_2$; 10 mM ATP; pH 7.5) in a total volume of 10-20 µL The amplified regions were inserted between the *Bam*HI and *Eco*RI sites of plasmid pGEX6P-3 or $pP_{RO}EX^{TM}$ HTb. This produced an in-frame transcriptional fusion of the protein of interest to an affinity tag used for purification.

2.2.5 Preparation of competent cells

E. coli XL1 cells were stored in glycerol at -70°C. Cells were taken from the glycerol with a flamed loop and streaked for single colonies on LB agarose plates. The plates were incubated overnight at 37°C. 2 mL of LB broth was inoculated overnight at 37°C with an isolated *E. coli* XL1 colony. 100 mL of LB broth was inoculated with the overnight culture until the OD 600 nm was 0.6. The cells were harvested by centrifugation for five minutes at 5000 rpm in a SS34 rotor at 4°C. The cells were resuspended in 20 mL of CM1 (10mM NaOAc, pH 5.6; 50 mM $MnCl_2$; 5 mM NaCl) and kept on ice for 20 minutes. The cells were harvested by centrifugation for five minutes at 5000 rpm in a SS34 rotor at 4°C. The cells were resuspended in 2 mL of CM2 (10 mM NaOAc, pH 5.6; 5 mM $MnCl_2$; 5% glycerol; 70 mM $CaCl_2$).

100 uL aliquots of the culture were pipetted into sterile microcentrifuge tubes and stored at -70°C.

2.2.6 Transformation of E. coli

0.5 μL-1 μL of a ligation mix containing 10-50 ng of DNA was added to 50 μL of pre-competent *E.coli* XL-1 cells. The cells were put on ice for 60 minutes, then put into a 42°C water bath for exactly 1 minute, and on ice again for 1-2 minutes. The cells were then put into 2 mL of LB (no antibiotic) and shaken at 37° for 45 minutes. 100 μL of the transformation mix was plated on LBA200 plates (4 g tryptone powder, 4 g NaCl, 2 g yeast extract, and 6.7 g agar per 200 mL plus 200 μg/mL ampicillin) and incubated at 37° overnight. The pGEX6P-3 and pProEXHTb vectors encode an ampicillin resistance gene, which enables positive selection of transformed cells.

2.2.7 Plasmid preparation by alkalyne lysis

Plasmids were isolated for analysis from transformed cells using the alkaline lysis method, described in Sambrook *et al*, 1999. This method is based on cell lysis in alkaline solution, followed by neutralization to precipitate genomic DNA and protein. The plasmid DNA is then precipitated from solution using ethanol.

2.2.8 Long-term storage of *E. coli* cells

Isolated colonies were used to inoculate 2 mL of LB broth (containing antibiotic 200 μg/mL ampicillin if the cells had been transformed with a pGEX6P3 or $pP_{RO}EX^{TM}$ HTb plasmid) and shaken overnight at 37°C. 0.5 mL of sterile filtered glycerol was mixed by inversion with 1.0 mL of transformed cells in a cryogen tube. The cells were frozen at -70°C.

2.2.9 Analysis of proteins by SDS-PAGE

Proteins can be separated by electrophoresis under denaturing conditions on the basis of mass. The anionic detergent sodium dodecyl sulfate (SDS) disrupts nearly all noncovalent interactions, and mercaptoethanol or dithiothreitol (DTT) is used to reduce disulfide bonds. Anions of SDS surround and coat the polypeptides of the protein, forming a micelle, causing the protein to unfold. The negative charges carried by the SDS molecules bound to the protein confer a negative charge on the protein proportional to the length of the chain. During electrophoresis an electric field is applied to a solution, within which the negatively charged micelles migrate towards the positively charged anode.

2.2.9.1 *Preparation and electrophoresis of 8-15% polyacrylamide gels*

The separating gel was made (refer to table 2.1), with the addition of ammonium persulphate (APS) and N,N',N',N' tetraethyl ethylene diamine (TEMED) last, and poured immediately into a pre-stacked gel casting apparatus. A thin layer of butanol saturated H_2O was pipetted onto the surface of the poured gel and the gel was left to set for an hour. The stacking gel was made by mixing 12.2 mL H_2O, 5 mL 0.5 M Tris/HCl pH 6.8, 0.2 mL 10% SDS, 2.6 mL 30% T, 2.7% C N-Methylene bis acrylamide, 0.1 mL 10% APS, and 0.02 mL TEMED. The butanol saturated H_2O was washed off the top of the set separating gel using water. The stacking gel was poured onto the top of the separating gel and was left to set for half an hour. The separated gels were stored at 4°C for up to a week.

The protein sample (0.5-10 μg) was diluted at least 1:4 parts with final sample buffer (3.8 mL milli-Q H_2O; 0.5 M Tris/HCl pH 6.8; 0.8 mL glycerol; 1.6 mL 10% (w/v) SDS; 0.4 mL 2-mercaptoethanol; 0.4 mL 1% (w/v) bromophenol blue). The protein

sample was then heated at 95°C for 5 minutes. The sample was left to cool before loading onto a gel.

Table 2-1 Separating gel SDS-PAGE: recipe to make 10 mini gels

	8%	10%	12%	15%
H2O /mL	28.2	24.2	20.1	14.0
1.5 M Tris/HCl pH /8.8 mL	15.0	15.0	15.0	15.0
10% SDS /mL	0.6	0.6	0.6	0.6
30% T, 2.7% C N-Methylene bis acrylamide /mL	16.0	20.0	24.0	30.0
10% APS /mL	0.3	0.3	0.3	0.3
TEMED /mL	0.03	0.03	0.03	0.03

The gels were electrophoresed on a minigel apparatus in SDS-PAGE running buffer (15.15 g Tris; 72.05 g glycine; 5.0 g SDS; dissolved in milli-Q H_2O to a total volume of 5 L) at 20 mA for approximately one hour (the amperage was doubled if two gels were run on the same powersource). The bromophenol blue present in the final sample buffer provided a visual measure of the progression of electrophoresis.

2.2.9.2 *Preparation and electrophoresis of 16.5% tricine gels*

The separating gel (10 mL 49.5% T, 3% C N-Methylene bis acrylamide; 10 mL gel buffer: 3 M Tris/HCl pH 8.45, 0.3% w/v SDS, 36.3 g Tris, 3 mL 10% SDS buffer /100 mL; 4 mL 80% glycerol; 6 mL H_2O; 150 µL 10% APS; 15 µL TEMED) was prepared and poured immediately into a pre-stacked gel casting apparatus. A thin layer of butanol saturated H_2O was pipetted onto the surface of the poured gel and the gel was left to set for an hour. The stacking gel was made by mixing 1.0 mL 49.5% T, 3.0% C N-Methylene bis acrylamide, 3.1 mL gel buffer, 8.4 mL H_2O, 100 µL 10%

APS, and 10 µL TEMED. The butanol saturated H_2O was washed off the top of the set separating gel using water. The stacking gel was poured onto the top of the separating gel and was left to set for half an hour. The separated gels were stored at 4°C for up to a week.

The protein sample (0.01-1 µg) was diluted at least 1:4 parts with Tris-Tricine sample buffer (4.0 mL milli-Q H2O; 2.0 mL 0.5 M Tris/HCl pH 6.8; 2.4 mL glycerol; 1.0 mL 10% (w/v) SDS; 0.2 mL 2-mercaptoethanol; 0.4 mL 0.5% Coomassie blue G-250). The protein sample was then heated at 95°C for 5 minutes and then left to cool before loading onto a gel.

The gels were electrophoresed on a minigel apparatus with cathode buffer (0.1 M Tris; 0.1 M Tricine, self adjusting to pH 8.25; 0.1 M SDS) and anode buffer (0.2 M Tris/HCl pH 8.9) at 30 mA for approximately two hours (the amperage was doubled if two gels were run on the same powersource). The Coomassie blue present in the final sample buffer provided a visual measure of the progression of electrophoresis.

2.2.9.3 *Staining of polyacrylamide gels*

Gels were stained either with Coomassie blue stain or Coomassie blue stain followed by silver stain dependent on the amount of protein in the samples. Silver staining detection (sensitive to 1-10 ng of protein) is up to 100 times more sensitive than Coomassie blue stain and is one of the most sensitive visualization methods. Staining initially with Coomassie blue is reported to increase the sensitivity of silver staining (Ross and Peters, 1990).

Coomassie blue stain

Gels were covered with Coomassie blue stain (0.1% Coomassie blue powder; 40% methanol; 10% acetic acid) for 30 minutes and shaken gently. The Coomassie blue

stain was poured off the gel and the gel was covered with destain (40% methanol; 10% acetic acid) for 30 minutes and shaken gently. The destaining solution was replaced 2-3 times until background staining was removed.

Silver stain

The gel was soaked in 50% methanol for 10 minutes followed by 5% methanol for at least 10 minutes. The gel was soaked in DTT (3.2 uL 1 M DTT/ 100 mL milli-Q H_2O) for 10 minutes followed by 0.1% $AgNO_3$ (0.1 g/ 100mL milli-Q H_2O) for 10 minutes. After rinsing with milli-Q H_2O 3 times, the gel was rinsed in developer (7.5g Na_2CO_3; 250 mL milli-Q H_2O; 125 µL 100% formaldehyde) and then soaked in 150 mL of the same solution until protein bands were visible. The gel was transferred into a 2% acetic acid solution for 10 minutes to stop development and then soaked in milli-Q H_2O for 10 minutes. All steps of the staining process were carried out on an orbital shaker.

2.2.10 Screening for optimum expression conditions for fusion proteins in *E. coli*

Isolated colonies were used to inoculate 2 mL of LB broth containing 200 µg/mL ampicillin (A200) and shaken overnight at 37°C. 2 mL of LBA200 inoculated with 100 µL of the overnight culture was shaken at 37°C for 3 hours. Isopropyl β-D-thiogalactoside (IPTG) was added to the 3-hour culture to induce expression of the GST-C5P fusion proteins. The induced culture was shaken at 37°C for 2 hours. Cells from 100 µL of culture were harvested by centrifugation. The pelleted cells were analyzed by sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) to demonstrate protein expression.

E. coli XL1 transformed cells were grown in LBA200 at 25°, 30°, or 37°C for GST-fusion proteins and 30° or 37°C for His-fusion proteins for periods of two to four

hours. Expression of the GST tagged fusion proteins was induced with 0.1 or 1 mM of IPTG while expression of the His-C5a fusion protein was induced with 0.1, 1, or 2 mM of IPTG. The cells were induced for 2 or 3 hours. Expression of a novel fusion protein was examined by comparative analysis of a constitutively expressed *E. coli* protein and the fusion protein bands on SDS-PAGE gels stained with Coomassie blue.

2.2.11 Purification of proteins from *E. coli*

Sonication

Cells resuspended in 1 mL or less were lysed, using a thin probe with the sonicator set at power level 2, for 3 rounds of 10 second sonication. For cells resuspended in 10 mL or greater cells were a fat probe was used with the sonicator set at power level 5, for 3 rounds of 20 second sonication. All samples were kept on ice during the sonication procedure.

GST-fusion proteins

When the conditions for optimal expression had been determined, 0.5-2.5 L cultures were grown for the various fusions. The induced cells were harvested by centrifugation for five minutes at 5000 rpm in a SS34 rotor at 4°C. The cell pellets were resuspended in 10 mL PBS buffer (1.4 M NaCl, 27 mM KCl, 0.1 M Na_2HPO_4, 18 mM KH_2PO_4, pH 7.3) per litre of initial culture. To inhibit proteolysis of the cloned GST-C5P fusion protein 20 μg/mL of the serine protease inhibitor PMSF was added. The addition of PMSF was omitted for protein prepared for use in activity assays. The resuspended cells were sonicated to release the cytoplasmic proteins. Soluble proteins were separated from the cell debris by centrifugation for 5 minutes at 5000 rpm in a SS34 rotor. The sonicate supernatant was shaken at 4°C for 2 hours with approximately 1 mL of washed Glutathione Sepharose 4B per mg of C5P fusion protein. The GST binds to the glutathione resin and therefore can be used to separate

GST fusion proteins away from other soluble proteins. After binding, the sonicate supernatant and glutathione resin was centrifuged for 5 minutes at 1000 rpm in a SS34 rotor at 4°C. The pellet containing the resin and bound GST-C5P fusion protein was washed 5 times with PBS and finally resuspended in 5 mL PBS/mL of resin. To release the C5P protein from its GST tag the resuspended resin was adjusted to 1 mM DTT and the protease PreScission (10 units/5 mL) was added. The samples were shaken at 4°C overnight. The progress of the cleavage of GST-C5P fusion protein was monitored by analysis of samples by SDS-PAGE. Released C5P protein was harvested from the resin by centrifugation for 5 minutes at 1000 rpm in a SS34 rotor at 4°C. When analysis showed that the cleavage was incomplete, the resin was resuspended in fresh PBS/DTT and additional PreScission (5 units/5 mL) was added. This process was repeated until most of the C5P protein had been released.

Expression of His-C5a in E. coli

20 mL of LB broth containing 200 μg/mL ampicillin (A200) was inoculated with an isolated colony and shaken overnight at 37°C. 500 mL of LBA200 was inoculated with 10 mL of the overnight culture and shaken at 37°C for 3 hours. 1 mM of isopropyl β-D-thiogalactoside (IPTG) was added to the 3-hour culture to induce expression of the His-C5a fusion protein. The induced culture was shaken at 37°C for 2 hours. Cells from 100 μL of culture were harvested by centrifugation. Expression was demonstrated by running a sample on a sodium dodecyl sufate-polyacrlamide gel (SDS-PAGE) stained with Coomassie blue.

Purification of His-C5a from inclusion bodies

The induced cells from 0.5 L culture were harvested by centrifugation for five minutes at 5000 rpm in a SS34 rotor at 4°C. The cell pellets were washed in Tris/HCl buffer, pH 7.5. The cells were resuspended in 25 mL of denaturation buffer (6 M guanidium-HCl). The resuspended cells were sonicated to release the cytoplasmic

proteins. The lysed cells were incubated at room temperature for 1 hour. After the incubation period 250 mL of renaturation buffer (1 mM reduced glutathione, 1 mM oxidized glutathione, 50 mM Tris/HCl pH 7.5) was added and the cells were incubated at room temperature overnight. Soluble proteins were separated from the cell debris by centrifugation for 5 minutes at 10,000 rpm in a SS34 rotor. The supernatant was dialysed into chelating sepharose start buffer (25 mM imidazole, PBS 500: 1.4 M NaCl, 27 mM KCl, 0.1 M Na_2HPO_4, 18 mM KH_2PO_4, pH 7.3, 500 mM NaCl)

Dialysis

Proteins were dialyzed in Amicon YM regenerated cellulose membrane with a molecular weight cut off of 10 kDa. The exchange buffer was changed three times during dialysis. The dialysis procedure was carried out from several hours to overnight at 4°C. The exchange volume was greater than 20 times that of the dialyzed sample.

Centrifugal concentration of C5P fragments

Recombinant C5P protein was concentrated by centrifugation at 3000 rpm in a Centriprep Centrifugal Filter Device with a molecular weight (Mw) cut-off of 10,000 Da (YM-10).

2.2.11.1 Fast protein liquid chromatography of recombinant proteins

Separation of C5P protein from lower molecular weight contaminants co-eluting off the glutathione resin was attempted with size exclusion fast protein liquid chromatography (SE FPLC) on a Bio-Prep Size Exclusion Column SE100/17. The column was eluted with a flow rate of 0.5 mL/min or 1.0 mL/min using PBS, pH7.4 containing 0.2 M NaCl as the eluting buffer. The volume of collected fractions was

0.5 mL. Samples of 0.25-5 μL from fractions corresponding to peaks on the chromatogram were assessed by SDS-PAGE.

Peak fractions eluted from the SEFPLC were pooled, dialysed into IE start buffer (20 mM Tris/HCl pH 7.5), and further purified by IE chromatography. The column was eluted with a flow rate of 1 mL/min with a gradient from 20 mM Tris/HCl pH7.5 to 1 M Tris/HCl pH 7.5. Samples of 0.25-5 μL from fractions corresponding to peaks on the chromatogram were assessed by SDS-PAGE. Peak fractions eluted from the IEFPLC were pooled and concentrated by centrifugation through a Centriprep Centrifugal Filter Device for use in screening for crystallization conditions and in activity assays.

2.2.11.2 Chelating sepharose purification of His-C5a

1 mL of chelating sepharose was used per 5 mg of His-C5a (amount of protein was assessed visually by SDS-PAGE). The chelating sepharose was prepared as follows. The sepharose was washed in 5 volumes of H_2O for 5 minutes, and then one half volume of 0.1 M $NiSO_4$ was added and the mixture was rotated for 5 minutes. The sepharose was washed with 5 volumes of H_2O and rotated for 5 minutes. This washing step was repeated 3 times. The sepharose was resuspended in 1 volume of PBS 500 (1.4 M NaCl, 27 mM KCl, 0.1 M Na_2HPO_4, 18 mM KH_2PO_4, pH 7.3, 500 mM NaCl). The His-C5a in start buffer was added to the sepharose and incubated with rotation for 30 minutes at room temperature. The sepharose was washed with 5 volumes of start buffer and rotated for 5 minutes. This washing step was repeated 3 times. The sepharose was packed into a column and a step-wise gradient of start buffer to elution buffer (500 mM imidazole, PBS 500) was applied to the column at a flow rate of 1 mL/min on an Econosystem™ at 4°C. Aliquots of the His-C5a eluted in fractions with 200 mM of imidazole were stored at -70° for use in activity assays.

2.2.12 Sequencing of C5P fragments and synthetic C5a

Plasmids were purified for sequencing from transformed strains using the QIAGEN midi kit system, which works by a process of gravity-flow anion-exchange, following the manufacturer's instructions. The cell wall is hydrolysed by alkaline lysis caused by the detergent SDS in the presence of NaOH. Potassium ions present in the lysis buffer neutralize the high pH and precipitate the SDS. The insoluble SDS traps the larger genomic DNA and removes it from the supernatant, leaving plasmid DNA and RNA in solution. The RNA is removed by digestion with RNaseA. The negatively charged nucleic acids are purified from proteins, carbohydrates and RNA nucleoside monomers in solution using anion-exchange chromatography. QIAGEN resin is a macroporous silica-based resin coated with positively charged diethylaminoethyl groups which interact with the negatively charged phosphates of the nucleic acid backbone. The resin is washed to remove impurities and then the DNA is eluted. The DNA is precipitated with isopropanol and washed with ethanol, further purifying it. The purified DNA was dissolved in 10 mM Tris-Cl, pH 8.5.

The purified plasmids were sent to the Massey University DNA Analysis Service (MUseq). The inserts were sequenced from primers (refer to appendix B) designed to the 5' and 3' regions of the cloning site by the dideoxy-chain termination method using an ABI Prism 377-64 DNA Sequencer. The primer is extended in four DNA polymerase catalyzed reactions. The reactions are performed with fluorescently labelled 2', 3'-dideoxyribonucleoside triphosphates (ddATP, ddCTP, ddGTP, or ddTTP), which serve as terminators of chain extension because they lack 3' hydroxyl termini. A series of fragments of varying lengths are produced, with each fragment identifying the insertion of a 2', 3'-dideoxyribonucleoside triphosphate. Gel electrophoresis produces a sequencing ladder of products labeled with fluorescent dyes corresponding to the dideoxy terminators which are detected using a laser. The sequencing results were analyzed using the Wisconson Package v. 9.1, Genetics Computer Group, Madison, WI.

2.2.13 Quantification of Protein by UV Spectrometry

The concentrations of purified recombinant proteins were determined by UV spectroscopy using the wavescan programme on a Pharmacia Biotech Ultraspec 3000 UV/Visible spectrophotometer. The wavelengths scanned were from 200-300 nm. The aromatic amino acids in proteins exhibit strong absorption of light in the near-ultraviolet region of the spectrum. Tryptophan and tryrosine produce most of the UV absorbance by proteins in the region around 280 nm. The extinction coefficients for the proteins were determined using the Prot Param tool at the ExPASy web site, http://us.expasy.org/. The measured absorption at 280 nm of the protein, and the extinction coefficients, were used to calculate the concentrations of the purified recombinant proteins.

2.2.14 Screening for Crystallization conditions for C5P

Crystallization conditions for recombinant C5P protein were screened for using the Hampton Crystal Screens one and two employing the hanging drop vapour diffusion method of crystallization. The primary screen variables were salt, pH, and precipitant (salts, polymers, volatile organics, and non-volatile organics). The recombinant proteins had been purified in PBS buffer. However, the phosphate in PBS readily crystallizes, and this would interfere with protein crystallization assays. Therefore, the recombinant proteins were exchanged into 50 mM HEPES pH 7.5 by multiple washes in a Centricon YM10. 0.1 M NaCl was added to improve the solubility of the protein. 1 μL of protein was mixed with 1 μL of the precipitant solution and suspended as a drop from a cover slip above 500 μL of the precipitant solution in a well of a 24 well Linbro plate, forming a sealed chamber. The plates were stored in the dark at room temperature and examined weekly by light microscopy.

The ammonium sulfate screen ranged over pH 4-pH 9 in 1 pH unit steps, and through 0.8 M-3.2 M ammonium sulfate in 0.8 M steps, resulting in 24 conditions. The

buffers used were: pH 9, 0.2 M Boric Acid; pH 8, 0.2 M Bis-Tris phosphate; pH 7, 0.2 M Hepes; pH 6, 0.2 M Citric Acid; pH 5, 0.2 M Citric Acid; pH 4, 0.2 M Formic Acid. The pH of all buffers were adjusted with KOH and/or HCl.

The ammonium sulfate/Hepes screen ranged over pH 6.5-pH 7.5 in 0.2 pH unit steps, and through 1.6 M-2.7 M ammonium sulfate in 0.2 M steps, resulting in 42 conditions.

2.2.15 Amino-terminal Sequencing

A sample of the C5P prepared for crystallization was separated by SDS-PAGE, 10% acrylamide. An Immobilon-P transfer membrane (PVDF) was prepared for electroblotting: The PVDF membrane was immersed for 15 seconds in 100% methanol, rinsed for 2 minutes in milli-Q purified water, and equilibrated in CAPS transfer buffer (10 mM CAPS, 10% methanol, pH 11.0) for 15 minutes. After electrophoresis, the gel was equilibrated in CAPS transfer buffer for 15 minutes. The gel and PVDF membrane were placed between Whatman 3mm paper and assembled into the blotting cassette. The gel was blotted in CAPS transfer buffer at 100 V for 1 hour. After the transfer was complete, the membrane was rinsed several times in Milli-Q water. Proteins were visualised by staining with Coomassie blue R-250 (0.1% in 50% methanol), and destained with several changes of destain solution (50% methanol, 10% acetic acid). The membrane was rinsed with several changes of water and air dried overnight. The protein of interest was cut out with a scalpel. The excised band was taken to Protein Sequencing Services, Institute of Molecular Biosciences, Massey University for amino-terminal sequencing by the Edman degradation method. The compound phenylisothiocyanate reacts with the terminal amino group to form a phenylcarbamoyl derivative. This derivative is then treated with a strong anhydrous acid resulting in cleavage of the peptide bond and liberating the N-terminal residue. The released N-terminal residue rearranges into a

phenylthiohydantoin derivative of the amino acid, which can be identified by high-pressure liquid chromatography. The sequence of degradation reactions can be repeated to identify up to 50 residues.

2.2.16 Purification of C5P from the cell wall of S. pyogenes

Initially, a method described by Fernandez-Espla *et al,* 2000 to harvest a Lactococcal cell envelope proteinase was attempted. 50 mL of Todd-Hewitt broth (selective media for streptococci: 3.1 g beef heart infusion; 20 g peptone; 10 g yeast extract; 2 g dextrose; 2 g NaCl; 0.4 g Na_2PO_4; 2.5 g $NaCO_3$; 8.0 mg gentamicin; 15 mg nalidixic acid) was inoculated with an isolated colony and incubated overnight at 37°C. The overnight culture was used to inoculate 800 mL of Todd Hewitt broth plus 20 mM glycine and incubated at 37°C for 16 hours. When the cell density was 0.4 at A_{600} nm 0.03% sodium tetrathionate was added to inhibit Spe B from releasing C5P from the cell wall. The cells were harvested by centrifugation at 8500 xg at 4°C. The cells were washed with 50 mM β-glycerophosphate buffer pH 7.0. One of two methods was used for the disruption of the cells: The cells were resuspended in 25 mL 50 mM Bis-Tris buffer pH 6.5. 2.5 x the pellet weight of alumina powder was added to the washed cells and they were crushed manually with a pestle for 5 minutes. The crushed cells were centrifuged at 500 xg to remove the alumina. Separating the streptococcal cells from the alumina powder by centrifugation proved problematic as the cells were found in the pellet with the powder, perhaps due to matting of the streptococcal chains. Therefore, the cells were resuspended in 10 mL 50 mM Bis-Tris buffer pH 6.5 and they were disrupted by sonication. The cell walls were separated from soluble cellular proteins by centrifugation at 20,000 xg for 15 minutes at 4°C. The cell walls were resuspended in 25 mL of 50 mM Bis-Tris buffer, pH 6.5. 400 units of mutanolysin were added to release the proteinase from the cell wall debris and the digest was incubated for 30 minutes at 37°C. The digest was centrifuged at

8500 xg for 15 minutes at 4°C. Supernatant harvested after digestion with mutanolysin was analyzed by SDSPAGE.

The second method followed to purify C5P from the cell wall was described by Bohnsack *et al* 1991. 50 mL of Todd-Hewitt broth was inoculated with an isolated colony and incubated overnight at 37°C. 800 mL of Todd Hewitt broth was inoculated with the overnight culture and incubated 37°C for 16 hours. When the cell density was 0.4 at A_{600} nm 0.03% sodium tetrathionate was added to inhibit Spe B from releasing C5P from the cell wall. The cells were harvested by centrifugation at 8500 xg at 4°C. The cells were washed with PBS buffer pH 7.4. The cells were resuspended in PBS buffer pH 7.4. 400 units of mutanolysin were added to release the proteinase from the cell wall debris and the digest was incubated for 30 minutes at 37°C. The digest was centrifuged at 8500xg for 15 minutes at 4°C. Supernatant harvested after digestion with mutanolysin was analyzed by SDSPAGE. The supernatant of this digest, which contained released cell surface proteins, was concentrated by centrifugation through a 100 kDa size exclusion filter. The supernatant was then dialyzed into start buffer (0.01 M phosphate buffer pH 7.0) and purified on a hydroxyapatite column. The column was operated with a step-wise gradient of start buffer to elution buffer (0.2 M phosphate buffer pH 7.0) applied to the column at a flow rate of 1 mL/min on an Econosystem™ at 4°C. Fractions containing the 130 kDa protein released at 0.15 M-0.2 M phosphate buffer were pooled and assayed for activity on His-C5a.

2.2.16 *In vitro* Activity assays

Activity assays were performed in 50 mM Tris/HCl ph 7.5 at 37°C for two hours. Activity was assessed by SDS-PAGE, as described in section 2.2.9, and mass spectrometry.

A 50 µL sample of the His tagged C5a digest mix was diluted to 10 pmol/µL in 50% acetonitrile, 50% water, and 1% acetic acid and purified by RPHPLC. The hydrophobicity of His-C5a was determined using the ProtParam tool at the ExPASy web site, http://us.expasy.org/, which was used to predict when the protein would be eluted. A 10 µL sample from the major peak fraction eluted from the RPHPLC column was analyzed by Electrospray Ionisation Mass Spectrometry (ESI-MS) at Protein Sequencing Services, Massey University. ESI-MS is a technique used on water-soluble biomolecules. The low volatility of proteins was a barrier to using mass spectrometry and is circumvented by using an acidic volatile solvent. The protein sample is sprayed into a vacuum chamber of the mass spectrometer in which the solvent evaporates leaving protein molecules carrying multiple positive charges. The charged protein molecules are accelerated by an electric field and then deflected by a magnetic field. This separates the ionized protein molecules according to the ratio of their mass to their charge (*m/z*), which is represented in a spectrum whose major peaks consist of the molecular ion with a different number of charges attached. The molecular ion peak gives the molecular weight of the compound, and is usually the peak of highest mass number (except for the isotope peaks which have a relatively low abundance).

2.2.17 Polymorphonuclear neutrophil purification

10 mL of blood was collected from a healthy volunteer into heparinized tubes. 2 mL of balanced salt solution was added to 2 mL of blood. (The balanced salt solution was prepared from two stock solutions, A and B. Solution A: 0.1% anhydrous D-glucose, 5.0×10^{-5} M $CaCl_2 \times 2H_2O$, 9.8×10^{-4} M $MgCl_2 \times 6H_2O$, 5.4×10^{-3} M KCl, 0.145 M Tris pH 7.6. Solution B: 0.14 M NaCl. To prepare the balanced salt solution 1 volume of solution A was mixed with 9 volumes of solution B.) The diluted blood sample was layered carefully onto 3 mL of Ficoll-Paque (Pharmacia Biotech) in a centrifuge tube and centrifuged in a Heraeus Megafuge 1.0 at 200 rcf for 30 minutes

at room temperature. The neutrophils were removed from the Ficoll-Paque and red blood cell interface with a Pasteur pipette. Contaminating red blood cells were removed by hypotonic lysis with 0.87% NH_4Cl followed by centrifugation in a Heraeus Megafuge 1.0 at 800 rcf for 10 minutes. The neutrophils were washed with 3 x volume of balanced salt solution and centrifuged in a Heraeus Megafuge 1.0 at 800 rcf for 10 minutes. Neutrophil concentration and purity was measured by a Coulter Cell Counter. The cells from 2 mL of blood were resuspended in 1 mL of Eagles minimal essential media (MEM).

2.2.18 Under-agarose migration assay

6 mL agarose solution (1.0% agarose, 0.5% gelatin, 50 mM Tris pH7.5, in Eagles MEM) was poured into 60 x 15 mm petri dishes. Wells of 2.4 mm diameter and 2.4 mm apart were cut into the cooled agarose. 10 uL of the chemotaxins and neutrophils (2×10^5) were loaded into the wells. The plates were incubated for 2 hours at 37°C in 5% CO_2-air atmosphere with high humidity. The plates were fixed with methanol for 30 minutes followed by 37% formaldehyde for 30 minutes. The agarose gel was removed after fixing and the cells were stained with May Grunwald Stain. Cell migration was viewed at 100 x magnification using a light microscope.

2.2.19 Polymorphonuclear neutrophil staining

The cells were covered with water free methanol for 10 minutes. The May Grunwald stain was prepared by diluting a stock solution 50:50 with Sorenson's buffered water. The methanol was poured out of the Petri dishes. May Grunwald stain was pipetted into the Petri dishes to cover the cells on the bottom with and left for 10 minutes. The Giemsa stain was prepared by mixing one volume of stock stain with 9 volumes of Sorenson's buffered water. The May Grunwald stain was poured out of the Petri dishes. Giemsa stain was pipetted into the Petri dishes to cover the cells and left for

10 minutes. The Giemsa stain was poured out of the Petri dishes. Sorenson's buffered water was pipetted into the Petri dishes to cover the cells and left for 10 minutes. The buffered water was poured out of the Petri dishes, and the dishes were left to dry completely before examining with the microscope.

2.2.19.1 *May Grunwald stain*

0.6 g of May Grunwald stain powder was added to 200 mL of methanol. The suspension was mixed, warmed to 50°C, and then cooled to room temperature. The mixture was left at room temperature for 24 hours with occasional shaking, and then filtered.

2.2.19.2 *Giemsa stain*

2 g of Giemsa stain powder was added to 200 mL of methanol. The suspension was mixed, and warmed at 50°C for 30 minutes, with occasional shaking. The mixture was left at room temperature overnight, and then filtered.

2.2.19.3 *Sorenson's buffer*

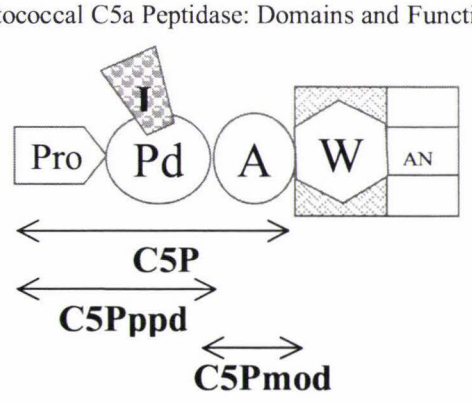
To make buffer pH 6.8 51 mL of solution A (1.82 g KH_2PO_4 dissolved in 200 mL of milli-Q H_2O) was mixed with 49 mL of solution B (2.38g $Na_2HPO_42H_2O$ dissolved in 200 mL of milli-Q H_2O). To make buffered water 5 mL of buffer was added to 95 mL of milli-Q H_2O. The pH was measured, if the pH was higher than 6.8 more solution A was added, if the pH was lower than 6.8 more solution B was added.

Chapter 3 Cloning, Expression, and Purification of domains of Streptococcal C5P

3.1 Introduction

Specific fragments of Streptococcal C5a Peptidase (C5P) were cloned, expressed and purified as a first step to study the specificity of this protease. These specific fragments included C5P (pro region plus protease domain plus modifier or A domain), C5Pmod (modifier domain), and C5Pppd (pro region plus protease domain), refer to figure 3.1. Most subtilisins are synthesized as pre-pro-enzymes. The 'pre' refers to an amino-terminal signal peptide, which is cleaved during translocation through the cell membrane. Cleavage of the pro-segment is generally required for activation of the subtilase, however this has not been established for C5P (Siezen, 1999). Removal of the predicted propeptide of C5P has been reported to be unnecessary for activation of the protease (Chen and Cleary, 1990). The recombinant fragments were purified and screened for crystallization conditions and also used in activity assays.

Figure 3-1 Streptococcal C5a Peptidase Domains

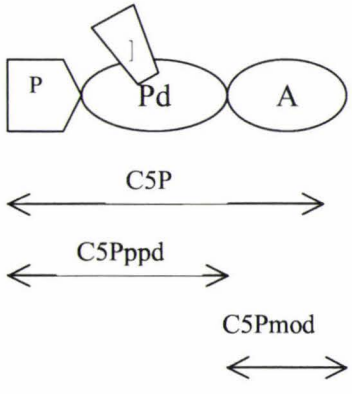


Pro: Resembles a domain of related proteases which require removal of this region for activation.
Pd: Protease Domain Shares sequence homology to a group of well characterized proteases: the subtilisins.
I: Insertion Large insertions exist within the conserved core of the subtilisins, which may modify activity.
A Domain: May modify activity.

3.1.1 Cloning Strategy

The coding sequence for three different fragments of the *S. pyogenes* B220 C5P gene, comprising the full length C5P (C5P: residues 32-1032), the protease plus proregion (C5Pppd: residues 32-583), and the putative modifier 'A' domain (C5Pmod: residues 584-1032), were amplified from genomic DNA by PCR, refer to figure 3.2. C5Pppd and C5Pmod were amplified using the proofreading *Pwo* polymerase. The full length C5P could not be amplified using *Pwo* DNA polymerase, whereas PCR using *Taq* DNA polymerase was successful. The *Pwo* polymerase has lower processivity than *Taq* polymerase, however, the greater error rate of *Taq* polymerase makes it unsuitable for use in cloning. Therefore, C5P was amplified using the Expand High Fidelity PCR System (Roche) which is composed of an enzyme mix containing both *Taq* DNA polymerase and a polymerase with 3'-5' exonuclease proofreading activity. The primers included 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning. The amplified regions were inserted between the *Bam*HI and *Eco*RI sites of plasmid pGEX6P-3. This produced an in-frame transcriptional fusion of the protein of interest to Glutathione S-Transferase (GST), an affinity tag used for purification. The resultant plasmids were named pGEXC5P (residues 32-1032 of C5P), pGEXC5Pppd (residues 32-583 of C5P), and pGEXC5Pmod (residues 584-1032 of C5P).

Figure 3-2 Constructs of C5P and design of the primers used in cloning the respective encoding regions of genomic DNA

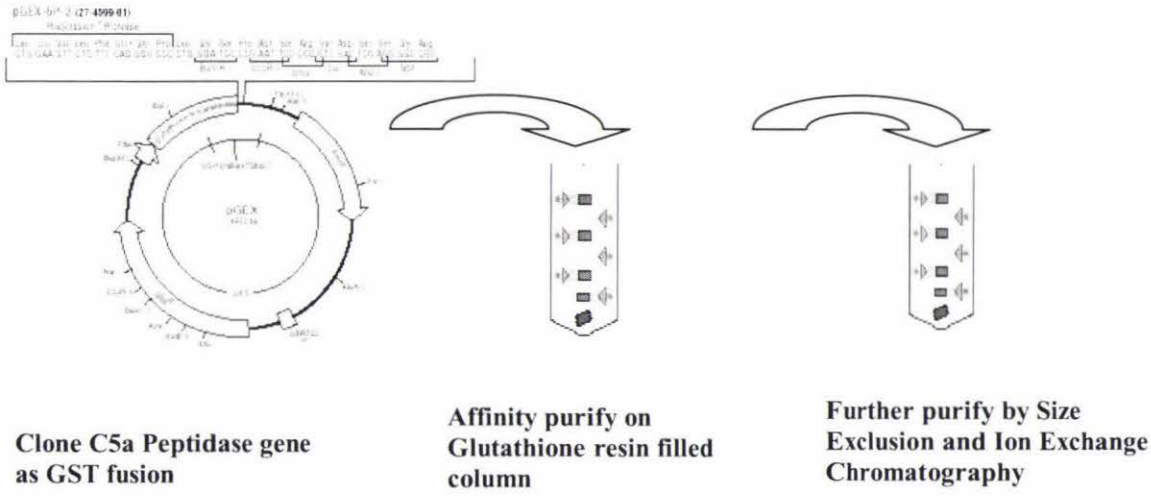


C5P construct	Residues	Primer pair
C5P	32-1032	For: CAGCAGGGATCCAATAC TGTGACAGAAGACACTCCT
		Rev:CAGCAGGAATTCTTAAG AGTGGCCCTCCAATAGCTT
C5Pppd	32-583	For: CAGCAGGGATCCAATAC TGTGACAGAAGACACTCCT
		Rev: CAGCAGGAATTCTTATG CTGCTGAAGCTTTTTTAGC
C5Pmod	584-1032	For:CAGCAGGGATCCACGAT GTATGTGACAGATAAGGAT
		Rev:CAGCAGGAATTCTTAAG

3.2 Synthesis, expression, and purification of specific fragments of C5P

All specific fragments of C5P were synthesized by PCR, expressed in *E. coli* as GST fusion proteins, and purified from cellular soluble proteins on glutathione resin. Further purification was achieved using size exclusion fast protein liquid chromatography (SEFPLC) and ion exchange fast protein liquid chromatography (IEFPLC), refer to figure 3.3.

Figure 3-3 Cloning, Expression, and Purification of recombinant C5P proteins



C5P, C5Pppd, and C5Pmod recombinant fragments of C5P were synthesized by PCR. The primers included 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning. The amplified regions were inserted between the *Bam*HI and *Eco*RI sites of plasmid pGEX6P-3. The recombinant proteins were expressed in *E. coli* as GST fusion proteins, and purified from cellular soluble proteins on glutathione resin. Further purification was achieved using size exclusion fast protein liquid chromatography (SEFPLC) and ion exchange fast protein liquid chromatography (IEFPLC).

3.2.1 Synthesis, expression and purification of C5Pppd

Primers were designed to amplify the pro plus protease domain (residues 32-583) of *Streptococcus pyogenes* C5P based on sequence information (accession number J05229) available in the GenBank database. Regions of C5P were amplified by PCR from genomic DNA prepared from *Streptococcus pyogenes* strain B220. A PCR product of 1.65 kb was expected, but as the forward primer sequence is almost completely repeated several times within the C5P gene, several products were amplified during PCR as illustrated in figure 3.4a. Stringency of PCR was increased

by increasing the annealing temperature and decreasing the $MgSO_4$ concentration, but this resulted in decreased product overall. The PCR product of the desired size was purified from the agarose gel shown in figure 3.4b using the HighPure Purification kit (Roche) after digestion with the restriction enzymes *Bam*HI and *Eco*RI.

Figure 3-4 Synthesis of C5Pppd from *S. pyogenes* B220 genome

Figure 3.4a
PCR of C5Pppd from *S. pyogenes* B220 genome

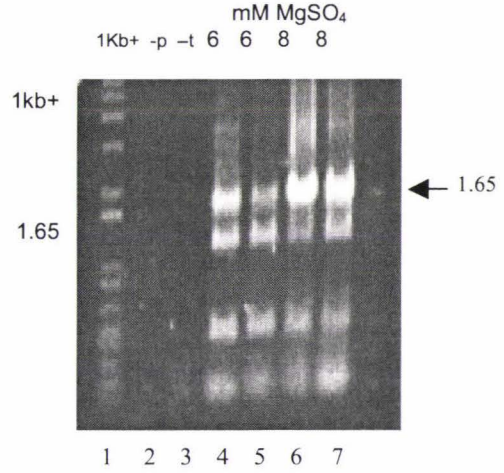


Figure 3.4a Lane 1 1Kb+ markers, Lane 2 PCR control with no primers, Lane 3 PCR control with no template, Lanes 4-7 $MgSO_4$ titration. 10 µL of each PCR reaction was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination. Conditions for amplification were: C5Pppd denaturation 94°30", annealing 50°30", extension 72°1min40" 6 or 8 mM $MgSO_4$; repeated for 30 cycles.

Figure 3.4b
C5Pppd for gel purification

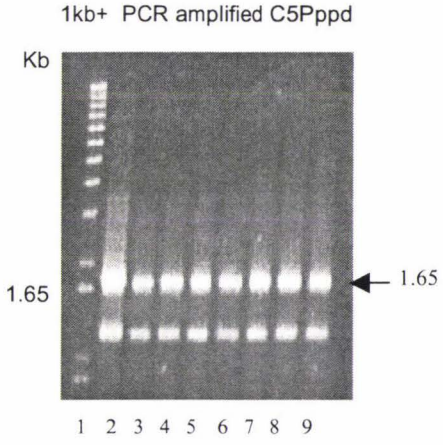


Figure 3.4b Lane 1 1Kb+ markers, Lanes 2-9 10 µL aliquots of pooled PCR products from reactions using 6 and 8 mM $MgSO_4$ were electrophoresed on 1% agarose gel at 100 V for 1 hour in 1x TAE buffer, pH 8. The gel was stained with ethiduium bromide (0.5 µg/mL) and visualized by UV transillumination. The 1.65 kb band that was excised from the gel is indicated by an arrow.

The gel purified fragments were ligated into a pGEX6P3 vector which had also been digested, within its polylinker cloning site, with *Bam*HI and *Eco*RI. 1 µL aliquots of the ligation mixture were added to 50 µL of competent *E. coli* XL1 cells. The cells were transformed by heat shock treatment. Transformed cells were selected by plating on LB-ampicillin plates. Plasmid preparations were made using an alkaline lysis method from overnight cultures grown in LB-ampicillin broth. A plasmid of 6.65 kb

was expected: pGEX6P3 vector (5 kb) plus C5Pppd insert (1.65 kb). Several colonies yielded plasmids of this size as shown in figure 3.5a. The plasmids were digested with *Bam*HI and *Eco*RI and analyzed by agarose gel electrophoresis to confirm the insert was of the correct size. An insert of 1.65 kb was released from the vector by digestion as illustrated in figure 3.5b.

Figure 3-5 Plasmid preparation of C5Pppd

Figure 3.5a
Plasmid Preparation C5Pppd

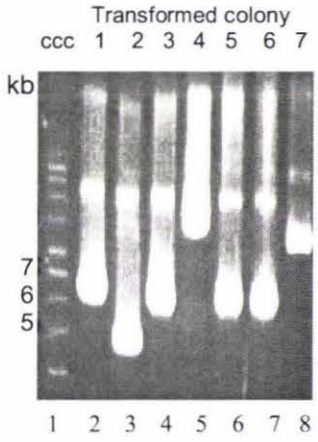


Figure 3.5a Lane 1 ccc markers, Lanes 2-8 samples from pGEXC5Pppd transformed cells prepared by alkaline lysis. 10 µL of each plasmid preparation was electrophoresed on a 0.8% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

Figure 3.5b
Bam*HI* and Eco*RI* digest of
Plasmid Preparation C5Pppd

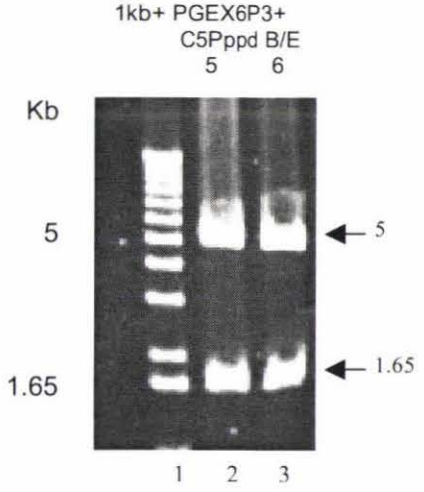


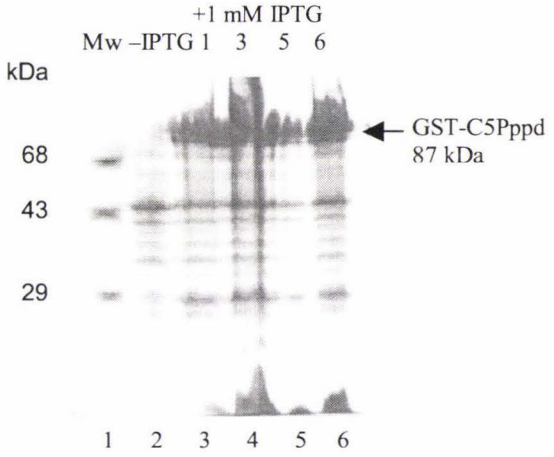
Figure 3.5b Lane 1 1Kb+ markers, Lanes 2 and 3 *Bam*HI and *Eco*RI digest of plasmid preparations. 10 µL of each digest was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

Cultures of putative transformed cells were induced with 1 mM IPTG, lysed, and the crude cellular extract was analyzed by SDSPAGE to check for protein expression. A protein of 86.6 kDa was expected consisting of the pro plus protease domains of C5P (60.6 kDa) and the amino terminal GST tag (26 kDa). Expression of a protein of this size was induced with addition of IPTG by four transformed colonies, as illustrated in figure 3.6a. The 87 kDa protein from two colonies was assayed for the ability to bind to glutathione resin and to be released from this resin with the site-specific protease PreScission, refer to figure 3.6b. The GST-fusion protein was soluble and when cleaved a significant portion of the C5Pppd was present in the supernatant. The

relevant protein bands are indicated in figure 3.6b. The C5Pppd recombinant protein had an apparent molecular weight of 68 kDa. This can be considered close to the expected molecular weight of 60.6 kDa, within the resolving power of the gel.

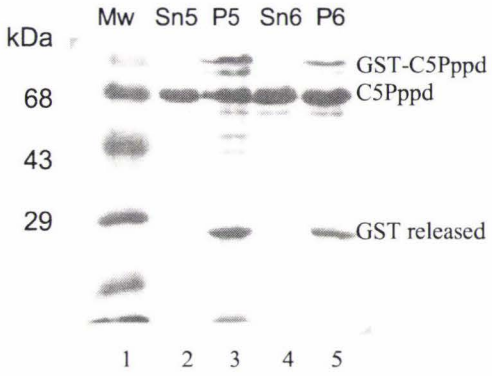
Figure 3-6 Induction of C5Pppd transformed cells and GST binding assay

Figure 3.6a
Induction of C5Pppd transformed cells



Whole cell lysate of cells induced with 1 mM IPTG. Lane 1 molecular weight markers; lane 2 uninduced cells; lanes 3-6 cells induced with IPTG. 10 µL of whole cell lysates prepared by sonication were analyzed by SDS-PAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

Figure 3.6b
GST Tagged C5Pppd

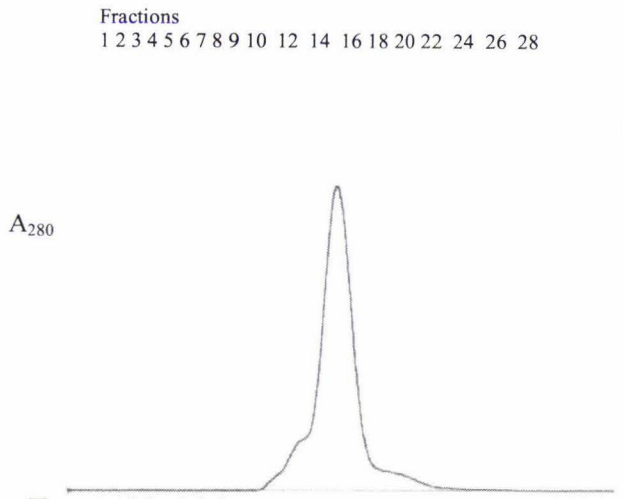


Cleavage of GST-C5Pppd with PreScission™ protease. Lane 1 molecular weight markers; lanes 2 and 4 C5Pppd released from GST tag in the supernatant (Sn) of a 20 µL sample; lanes 3 and 5 GST tagged C5Pppd, released C5Pppd and GST in the pellet (P) of a 20 µL sample. 10 µL samples were analyzed by SDS-PAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

GST tagged C5Pppd was then purified on a larger scale (2.5 L of culture) from the other cellular soluble proteins by binding to GST resin, and released from the resin with PreScission™ protease. The protein released was contaminated with lower molecular weight proteins, possibly breakdown products of C5Pppd, refer to lanes 2 and 4 in figure 3.6b. While concentrating the protein by centrifugation, through a 10 kDa size exclusion filter, C5Pppd was further degraded. Purification was attempted using SEFPLC with limited success as multiple bands were visualized when analyzed by SDS-PAGE as illustrated in figure 3.7b.

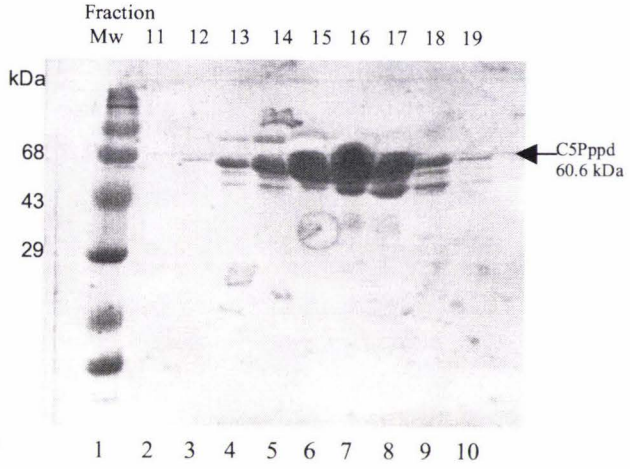
Figure 3-7 SEFPLC of C5Ppd

Figure 3.7a Chromatogram of SEFPLC of C5Ppd



Flow rate 0.5 mL/min
Buffer: PBS-GST/0.2 M NaCl
The A_{280} absorbance is represented with a bold line.

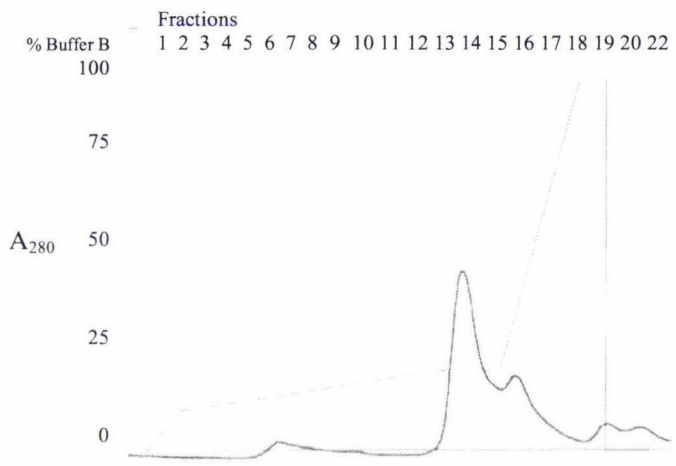
Figure 3.7b SDSPAGE of fractions eluted from SEFPLC of C5Ppd



10μL samples of fractions eluted off the SEFPLC column were analyzed by SDSPAGE on a 12% acrylamide gel. The gel was stained with Coomassie Blue. Lane one molecular weight markers; lanes 2-10 samples from fractions covering peak of chromatogram from a 500 μL injection of recombinant C5Pppd sample.

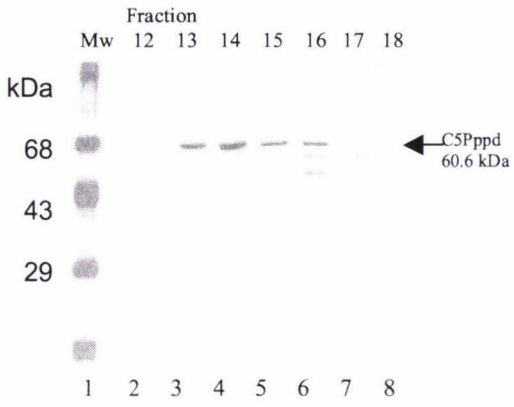
Figure 3-8 IEFPLC of C5Pppd

Figure 3.8a Chromatogram of IEFPLC of C5Pppd



Flow rate: 1 mL/min
Buffer A: 20 mM Tris/HCl pH7.5
Buffer B: 1 M Tris/HCl pH 7.5
The A_{280} absorbance is represented with a bold line. The % buffer B is represented with a thin line

Figure 3.8b IEFPLC eluted fractions run on 15% SDS-PAGE



10 μL samples of fractions eluted off the IEFPLC column were analyzed by SDS-PAGE on a 12% acrylamide gel. The gel was stained with Coomassie Blue. Lane one molecular weight markers; lanes 2-8 samples from fractions covering peak of chromatogram from a 500 μL injection of recombinant C5Pppd sample.

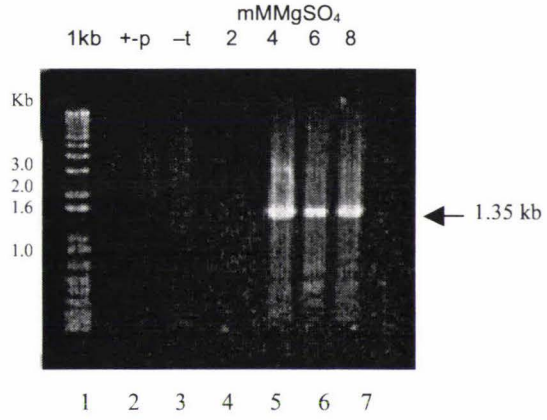
The fractions containing the majority of recombinant C5Pppd protein (fractions 15 and 16) were pooled, dialyzed into IE start buffer (20 mM Tris/HCl pH 7.5), and purified further by IEFPLC, as illustrated in figure 3.8. Aliquots from fractions 13, 14, and 15 eluted from the IE column, which represented homogenous protein, were frozen at -70°C for use in activity assays.

3.3 Synthesis, expression, and purification of C5Pmod

Primers were designed to amplify the modifier or 'A' domain of *Streptococcus pyogenes* C5P (residues 584-1032) based on sequence information (accession number J05229)available in the GenBank database. The region of C5P encoding the modifier region was amplified by PCR from genomic DNA from *Streptococcus pyogenes* strain B220. A PCR product of 1.5 kb was expected and was produced with 4, 6, and 8 mM $MgSO_4$ in the PCR reaction mixture (lanes 5-7 in figure 3.9a). There were no major contaminating PCR products so a gel purification step was not necessary.

Figure 3.9 Synthesis of C5Pmod

Figure 3.9a C5Pmod PCR

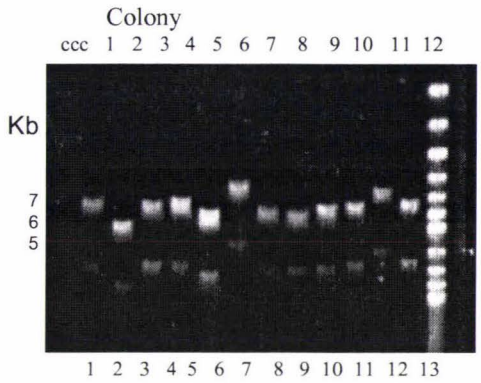


Lane 1 1Kb+ markers, Lane 2 PCR control with no primers, Lane 3 PCR control with no template, Lanes 4-7 $MgSO_4$ titration. 10μL of each PCR reaction was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with EtBr (0.5 μg/mL) and visualized by UV transillumination. Conditions for amplification were: C5Pmod denaturation 94°30", annealing 54°30", extension 72°1min20", 4, 6, or 8mM $MgSO_4$, repeated for 30 cycles

The product was digested with the restriction enzymes *Bam*HI and *Eco*RI, and was ligated into a pGEX6P3 vector prepared by digestion with the same restriction enzymes. 1 μL aliquots of the ligation mixture were added to 50 μL of *E. coli* XL1 competent cells. The cells were transformed by heat shock treatment. Transformed cells were selected by plating on LB-ampicillin plates. Plasmid preparations were made by alkaline lysis from overnight cultures of resultant colonies. A plasmid of 6.35 kb was expected: pGEX6P3 vector (5 kb) plus C5Pmod insert (1.35 kb). Several plasmids of this size were produced as shown in figure 3.10a. The plasmids were digested with *Bam*HI and *Eco*RI and analyzed by agarose gel electrophoresis to confirm the insert was of the correct size. Three colonies yielded inserts of the expected size (1.35 kb) as shown in figure 3.10b.

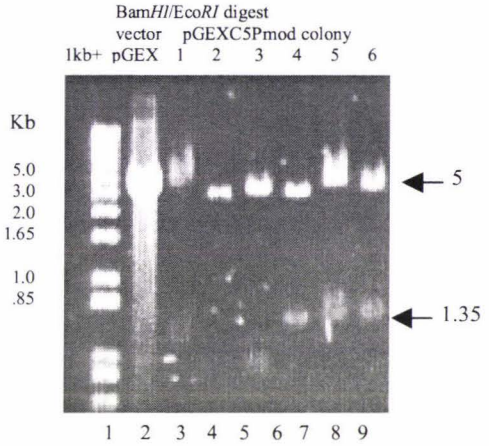
Figure 3.10 Plasmid preparation of C5Pmod

Figure 3.10a
Plasmid preparation of pGEXC5Pmod transformed cells



Lane 1, ccc markers; Lanes 2-13, samples from pGEXC5Pmod transformed cells prepared by alkaline lysis. 10 µL of each plasmid preparation was electrophoresed on 0.8% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

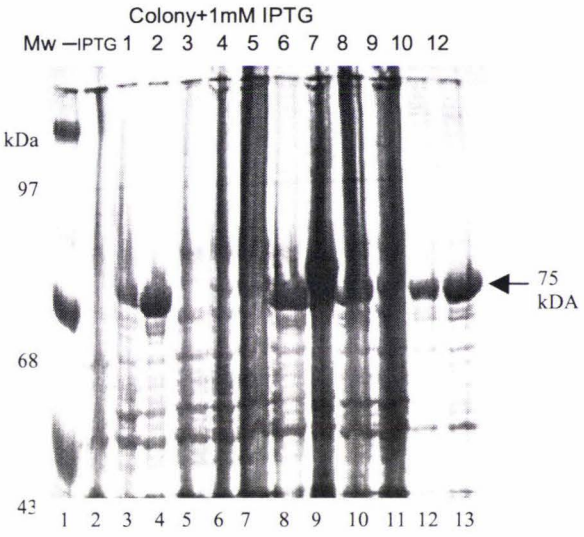
Figure 3.10b
Bam*HI* and Eco*RI* digest of plasmid preparation from pGEXC5Pmod transformed cells



Lane 1 1Kb+ markers, Lane 2 *Bam*HI and *Eco*RI digest of pGEX6P3, Lanes 3-9 *Bam*HI and *Eco*RI digest of plasmid preparations. 10 µL of each digest was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

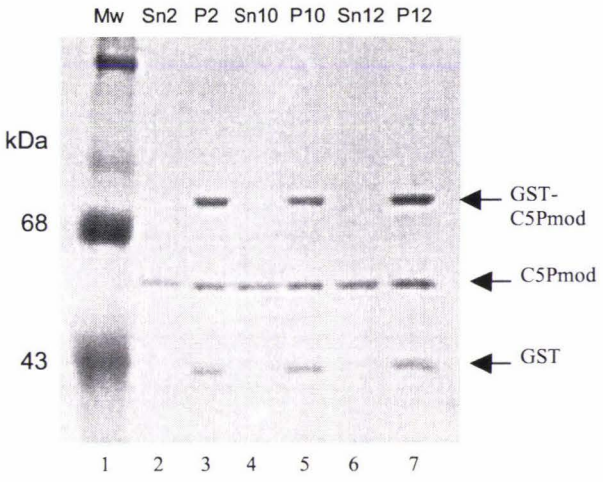
Figure 3.11 Expression of C5Pmod and GST binding assay

Figure 3.11a Transformation of C5Pmod



Whole cell lysate of cells induced with 1 mM IPTG. Lane 1 molecular weight markers; lane 2 uninduced cells; lanes 3-13 cells induced with IPTG. 10uL of whole cell lysates prepared by sonnication were analyzed by SDSPAGE on a 12% acrylamide gel. The gel was stained with Coomassie Blue.

Figure 3.11b GST purification of C5Pmod



Cleavage of GST-C5Pmod with PreScission™ protease. Lane 1 molecular weight markers; lanes 2, 4, and 6 C5Pmod released from GST tag in the supernatant (Sn) of a 20uL spun sample; lanes 3, 5, and 7 GST tagged C5Pmod, released C5Pmod and GST in the pellet (P) of a 20uL spun sample. 10uL samples were analyzed by SDSPAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

Colonies were induced with 1 mM IPTG, lysed by sonication, and the crude cellular extract was analyzed by SDSPAGE to check for protein expression. A protein of 75.4 kDa was expected consisting of the A domain of C5P (49.4 kDa) and the amino terminal GST tag (26 kDa). Expression of a protein of this size was induced with addition of IPTG by several transformed colonies as illustrated in figure 3.11a.

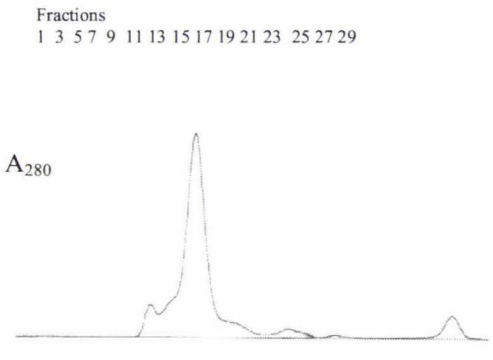
C5Pmod produced by three transformed colonies (2, 10, and 12) was assayed for binding to GST resin and release with PreScission™ protease, refer to figure 3.11b. A significant amount of soluble C5Pmod was released from the GST tag as illustrated in lanes 2, 4, and 6 of figure 3.11b. Colony 2 was used for the large scale preparation (1.5 L) of C5Pmod.

C5Pmod released by PreScission™ protease was concentrated by centrifugation through a 10 kDa size exclusion filter and purified by SEFPLC as illustrated in figure 3.12. Peak fractions eluted from the SEFPLC containing recombinant C5Pmod protein were pooled, dialysed into IE start buffer (20 mM Tris/HCl pH 7.5), and purified by IE chromatography on an Econosystem, refer to figure 3.12. Protein

concentration of each fraction was determined using the Bradford assay (Biorad reagent), and analyzed by SDSPAGE to determine protein content. A homogeneous protein of the appropriate molecular weight for C5Pmod was recovered with no apparent degradation. The peak fractions eluted from the IE column were pooled and concentrated by centrifugation, through a 10 kDa size exclusion filter, and used for screening of crystallization conditions.

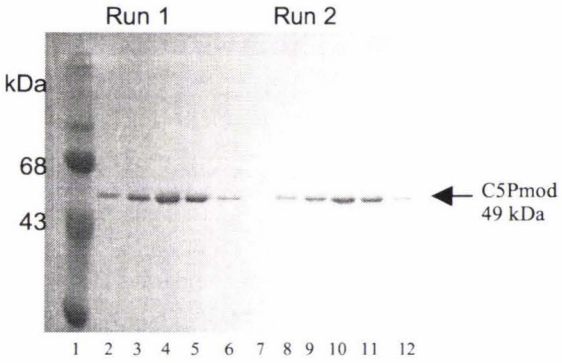
Figure 3.12 SEFPLC of C5Pmod

Figure 3.12a
SEFPLC of C5Pmod



Flow rate: 0.5 mL/min Buffer: PBS-GST/0.2M NaCl
The bold line represents absorbance A_{280}

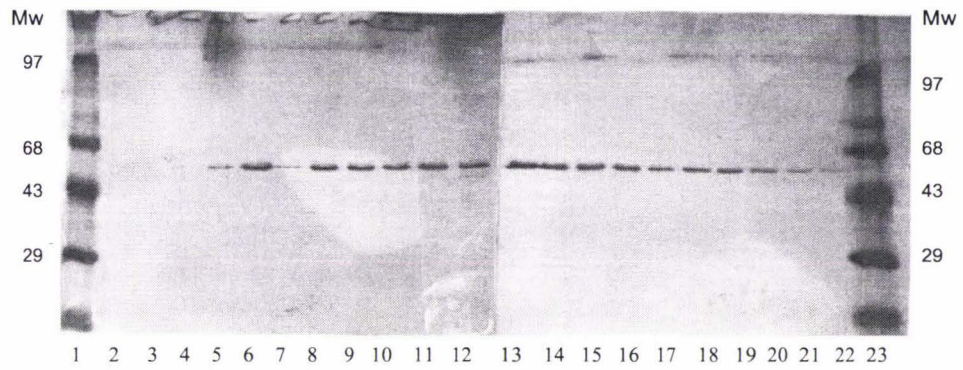
Figure 3.12b
SDS-PAGE of peak fractions eluted from C5Pmod SEFPLC



10 µL samples of fractions eluted off the SEFPLC column were analyzed by SDSPAGE on a 12% acrylamide gel. The gel was stained with Coomassie Blue. Lane 1 molecular weight markers; lanes 2-6 and 7-12 samples from fractions covering the peak of chromatogram in two separate injections of recombinant C5Pmod sample.

Figure 3.13 IE Chromatography of C5Pmod

SDSPAGE of C5Pmod fractions eluted off the IE column



Flow rate: 1 mL/min
Buffer A: 20mMTris/HCl pH7.5
Buffer B: 1M Tris/HCl pH7.5
A_{280}

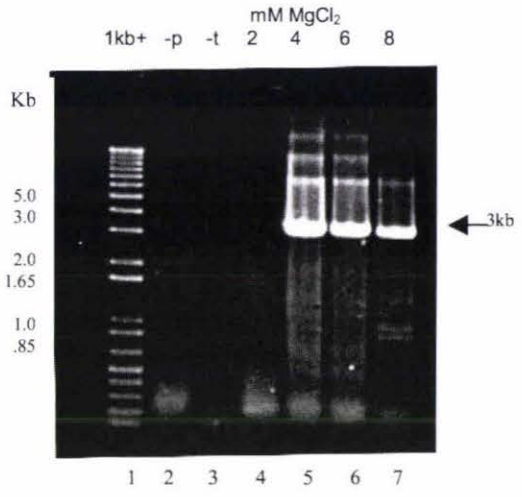
10 µL samples of fractions eluted off an IE source 15Q column were analyzed by SDSPAGE on a 12% acrylamide gel. The gel was stained with Coomassie Blue. Lanes one and 23 molecular weight markers; lanes 2-22 samples from fractions covering peak of chromatogram from a 10 mL sample of recombinant C5Pmod.

3.4 Synthesis, expression, and purification of C5P

Primers were designed to amplify, by PCR from genomic DNA prepared from *Streptococcus pyogenes* strain B220, the complete coding region of *Streptococcus pyogenes* C5P with the exception of the cell wall spanning domains (residues 32-1032). The sequence of the primers was based on sequence information (accession number J05229) available in the GenBank database. A PCR product of 3 kb was expected. This region would not amplify with *Pwo* polymerase (the polymerase used to amplify the other regions of C5P). A product was achieved with *Taq* polymerase; however, the lower fidelity of the polymerase was a concern. Finally, the C5P gene was synthesized with Expand polymerase, a mixture of both *Pwo* and *Taq* polymerases. A product of the appropriate size (3 kb) was generated using 4, 6, and 8 mM $MgCl_2$ as illustrated in figure 3.14a.

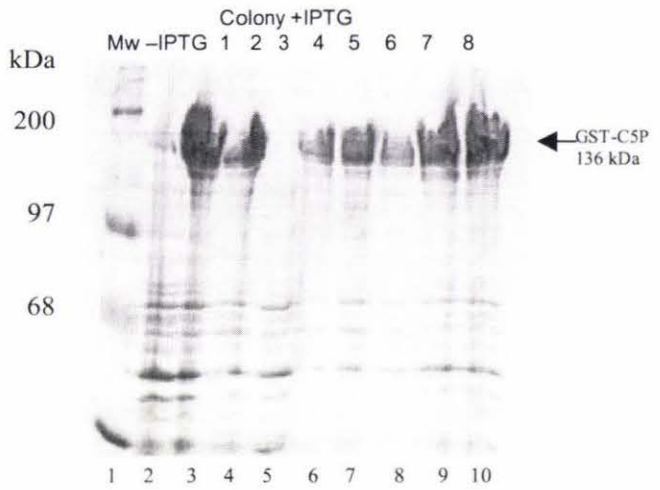
Figure 3.14 Synthesis of C5P

Figure 3.14a PCR of C5P



Lane 1 1Kb+ markers, Lane 2 PCR control with no primers, Lane 3 PCR control with no template, Lanes 4-7 $MgCl_2$ titration. 10 μL samples were electrophoresed on 1% agarose at 100V for 45 minutes in 1x TAE buffer, pH8, stained with ethidium bromide and visualized by UV transillumination. Conditions for amplification were: C5P denaturation 94°30", annealing 50°30", extension 72°3min, 4, 6 or 8 mM $MgCl_2$; repeated for 30 cycles

Figure 3.14b Transformation of C5P



Whole cell lysate of cells induced with 1 mM IPTG. Lane 1 molecular weight markers; lane 2 uninduced cells; lanes 3-10 cells induced with IPTG. 10 μL of whole cell lysates prepared by sonication were analyzed by SDSPAGE on a 10% acrylamide gel. The gel was stained with Coomassie Blue.

The product was digested with the restriction enzymes *Bam*HI and *Eco*RI, and was ligated into a pGEX6P3 vector digested with the same restriction enzymes. 1 μL aliquots of the ligation mixture were added to 50 μL of *E. coli* XL1 competent cells, which were then transformed by heat shock treatment. Transformed cells were selected by plating on LB-ampicillin plates. Plasmid preparations were made by alkaline lysis from resultant colonies. A plasmid of 8 kb was expected: pGEX6P3 vector (5 kb) plus C5Pppd insert (3 kb). The plasmids were digested with *Bam*HI and *Eco*RI and analyzed by agarose gel electrophoresis to confirm the insert was of the correct size.

Colonies containing plasmids with the correct sized inserts were induced with IPTG, lysed by sonication, and the crude cellular extract analyzed by SDSPAGE to check for

protein expression, refer to figure 3.14b. A protein of 136 kDa was expected consisting of C5P (110kDa) and the amino terminal GST tag (26kDa). Expression of a protein of this size was induced with addition of IPTG by several transformed colonies as illustrated in figure 3.14b.

Colony 1 was used for the large scale preparation (2.5 L culture) of C5P. Recombinant C5P released from the amino terminal GST tag with PreScission™ protease was concentrated by centrifugation through a 10 kDa size exclusion filter and purified by SEFPLC, as illustrated in figure 3.15. Samples of each fraction across the A peak from the SEFPLC column were analyzed by SDS-PAGE (figure 3.15b). This demonstrated that a 110 kDa protein was purified but some lower molecular weight contaminants were present; these were possibly degradation products.

Figure 3.15 SEFPLC of C5P

Figure 3.15a Chromatogram of C5P SE Chromatography

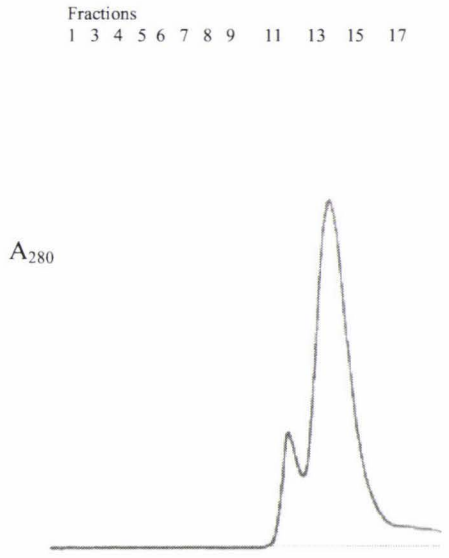


Figure 3.12a
Flow rate: 1 mL/min
Buffer: PBS-GST/0.2M NaCl
The A_{280} absorbance is represented with a bold line.

Figure 3.15b SDSPAGE of C5P fractions eluted off SEFPLC

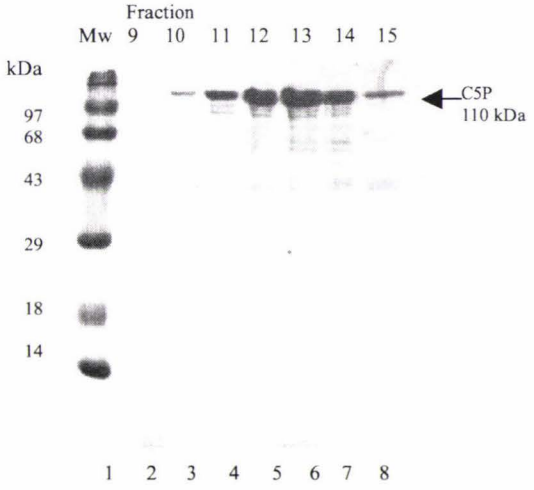


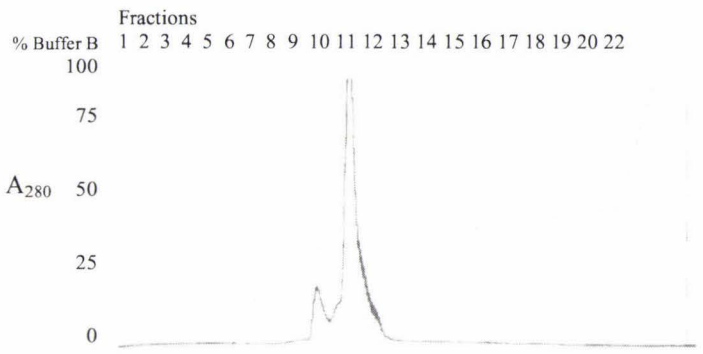
Figure 3.12b
10µL samples of fractions eluted off the SEFPLC column were analyzed by SDSPAGE on a 10% acrylamide gel. The gel was stained with Coomassie Blue.Lane one molecular weight markers; lanes 2-8 samples from fractions covering peak of chromatogram from 500 µL injection of recombinant C5P sample.

Peak fractions eluted from the SEFPLC were pooled, dialysed into IE start buffer (20 mM Tris/HCl pH 7.5), and further purified by IE chromatography as illustrated in

figure 3.15. Protein eluting from the IE column in fraction number 11 (lane 7 in figure 3.15b) appeared to be homogeneous and of the appropriate molecular weight for C5P. Several samples of recombinant C5P were purified by IEFPLC. Fractions 11 eluted from the IE column, from several 500 μL injections of sample, were pooled and concentrated by centrifugation through a 10 kDa size exclusion filter for use in screening for crystallization conditions and in activity assays.

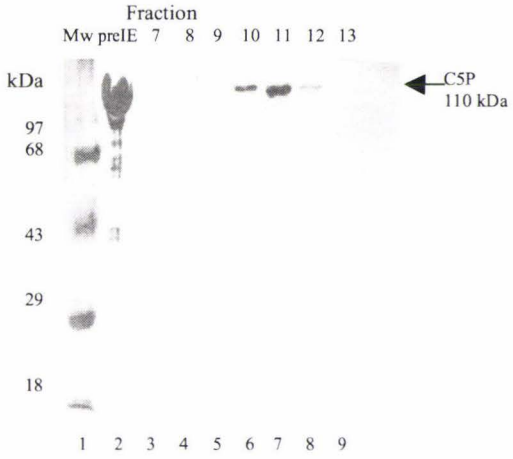
Figure 3.16 IEFPLC of C5P

Figure 3.16a IEFPLC of C5P



Flow rate: 1 mL/min
Buffer A: 20mMTris/HCl pH7.5
Buffer B: 1M Tris/HCl pH7.5
The A_{280} absorbance is represented with a bold line. The % buffer B is represented with a thin line

Figure 3.16b
IE Chromatography of C5P



10 μL samples of fractions eluted off the IEFPLC column were analyzed by SDSPAGE on a 10% acrylamide gel. The gel was stained with Coomassie Blue. Lane one molecular weight markers; lane 2 sample pre-IE; lanes 3-9 samples from fractions covering peak of chromatogram from a 500 μL injection of recombinant C5P sample.

Sequencing

Sequencing of each of the C5P constructs confirmed that the C5P fragments had been successfully cloned. The query sequences did not share 100% identity with the *S. pyogenes* C5P sequences found in a BLAST search, which may reflect variation between strains or errors introduced during PCR and cloning. For all fragments only one clone was sequenced. Sequencing of a second clone could confirm if variations were PCR errors or variation between strains. The fragments were cloned from *S. pyogenes* strain B220. The significant alignments found in a BLAST search were

with the C5a peptidase gene in *S. pyogenes* strain cs24; *S.agalactiae* (GBS) strains I30 and I25.

The BLAST search also identified homology between C5a peptidase and the *S. pyogenes* M protein, M49. The M and M-like proteins of *S. pyogenes* are cell surface proteins like C5P, which assist in adherence to host cells. The M proteins also have antiphagocytic activity. The genes encoding the M and M-like proteins and C5a peptidase are organised into a pathogenicity island, called the *mga* regulon, as transcription is regulated by Mga.

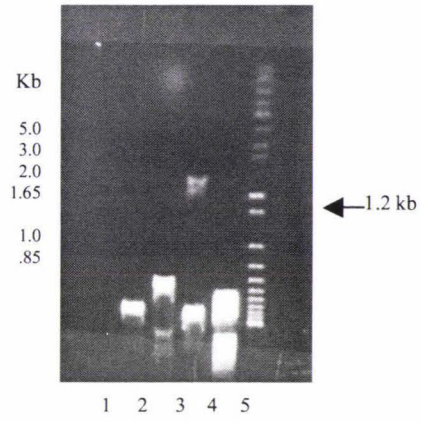
3.6 Preparation of C5P with an Active Site Ser to Ala Mutation

The synthesis of $C5P^{S512toA512}$, a knock-out mutation of the active site serine of the protease, was attempted in two ways. The first method used a $C5Ppd^{S512toA512}$ clone that had previously been engineered by Dr Jakki Cooney. The site of mutation and several flanking residues were restriction endonuclease digested and ligated into a pGEX6P3 vector containing a C5P insert cut with the same restriction enzymes. The second method involved PCR with forward and reverse oligonucleotides encoding the serine to alanine mutation. The C5P Ser to Ala protein expressed from this mutated gene could then be used as a negative control in activity assays.

The $C5Ppd^{S\rightarrow A}$ clone was provided in a pGEX6P3 vector. This vector was named $pGEXpd^{S2A}$. The sequence of pdS2A was analyzed for single cutting restriction endonuclease sites by the DNA analysis program Gene Jockey, Biosoft. A *Bgl*II endonuclease restriction site 5' of Ser 512 and *Sac*I endonuclease restriction site 3' of Ser 512 were found. These restriction endonuclease sites were found to be absent in the pGEX6P3 vector. The *Bgl*II endonuclease restriction site starts at residue 462 and the *Sac*I endonuclease restriction site starts at residue 1649, whilst Ser 512 is encoded by residues 1532-1534, of the C5P coding sequence. The digestion of pGEXC5P and $pGEXpd^{S2A}$ with *Bgl*II and *Sac*I should have released a fragment of approximately 1.2 kb. A fragment of this size was released from digested pGEXC5P, however $pGEXpd^{S2A}$ was cut by only one of the restriction enzymes, as illustrated in figure 3.17a. $pGEXpd^{S2A}$ was digested individually by the restriction enzymes and the digests were analyzed by electrophoresis, refer to figure 3.17b. This revealed that $pGEXpd^{S2A}$ was not cut by *Sac*I, possibly due to a PCR error within this restriction endonuclease site.

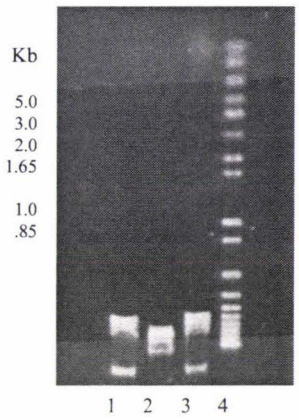
Figure 3.17 *Bgl*II and *Sac*I digest of pGEXC5P and pGEXpdS2A

Figure 3.17a *Bgl*II and *Sac*I digest of pGEXC5P and pGEXpd^{S2A}



Lane 1 1Kb+ markers, Lane 2 pGEXC5P uncut (0.5 uL); Lane 3 *Bgl*II and *Sac*I digest of pGEXC5P, Lane 4 pGEXpd^{S2A} uncut; Lane 5 *Bgl*II and *Sac*I digest of pGEXpd^{S2A}. 10 µL of each digest was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

Figure 3.17b *Bgl*II and *Sac*I digest of pGEXpd^{S2A}

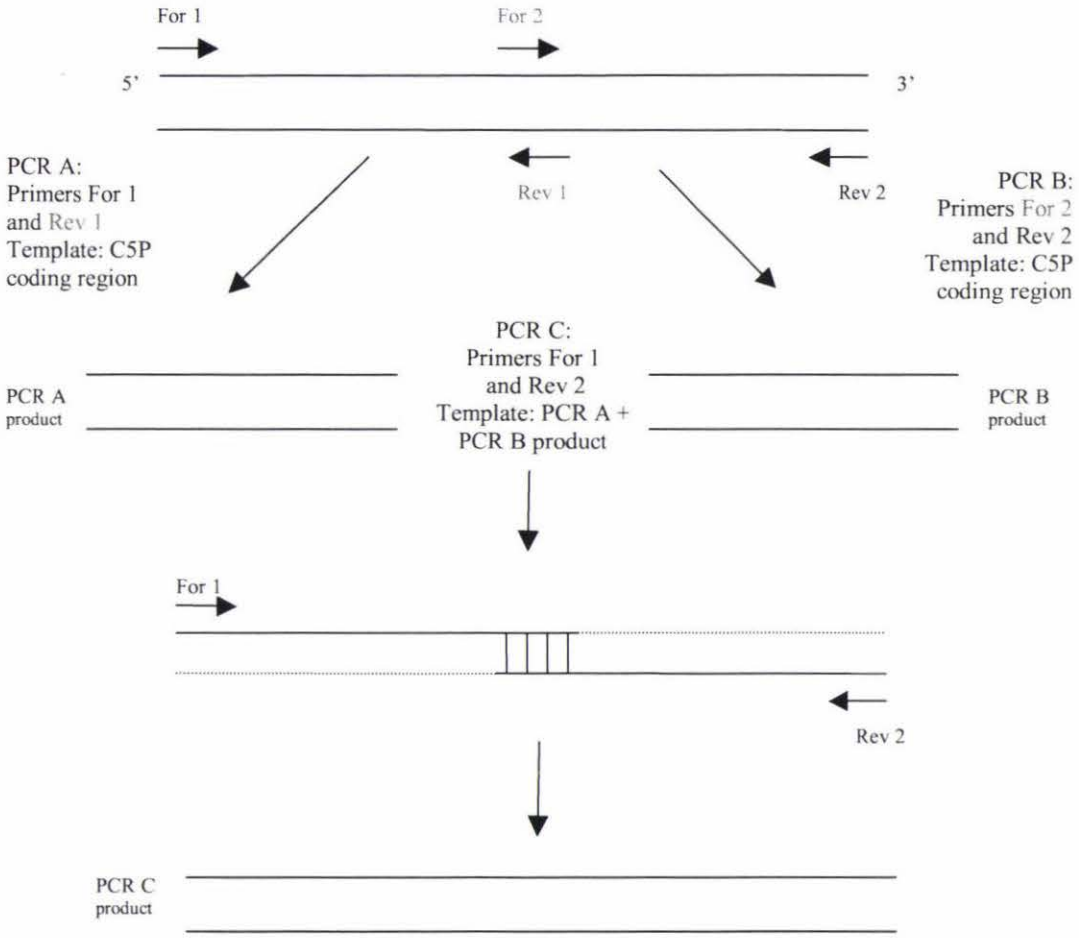


Lane 1 1Kb+ markers, Lane 2 pGEXpd^{S2A} uncut; Lane 3 *Bgl*II digest of pGEXpd^{S2A}, Lane 4 *Sac*I digest of pGEXpd^{S2A}. 10 µL of each digest was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

The mutation was introduced by PCR using *Taq* polymerase over the Ser 512 codon using a plasmid preparation of C5P in the pGEX6P3 vector as template and forward (PCRA) and reverse (PCRB) primers designed with a Ser 512 to Ala 512 mutation, as illustrated in figure 3.18. A product of the expected size of 3 kb was amplified with Expand polymerase using the products generated by PCR A and PCR B as template, as illustrated in figure 3.19. The presence of lower molecular weight bands suggests there is non-specific annealing of the primers. This had been demonstrated previously in the amplification of the C5P coding region and is most likely to be due to repetition of sequence within the gene.

Figure 3.18 Synthesis of C5P with an active site Ser to Ala Mutation

A C5P Ser to Ala mutation was introduced by PCR using *Pwo* polymerase over the Ser 512 from a plasmid preparation of C5P in the pGEX6P3 vector as template and forward (For 2) and reverse (Rev 1) primers coding for a Ser 512 to Ala 512 mutation. The products generated by PCR A and PCR B purified from an agarose gel were used as template to generate the C5P coding region with an active site Ser to Ala Mutation, amplified with Expand polymerase.

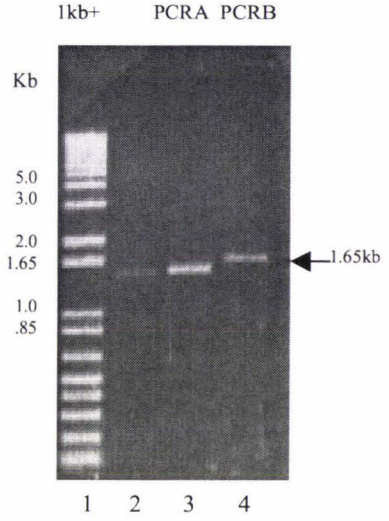


C5P coding region with an active site Ser to Ala Mutation

Forward mutagenic oligo (For 2):
GCC AAA CTT TCT GGA ACT GCT ATG TCT GCG CCA TTA GTA GCG GG
Reverse mutagenic oligo (Rev 1):
CC CGC TAC TAA TGG CGC AGA CAT AGC AGT TCC AGA AAG TTT GGC
The mutated codon is shown in red.

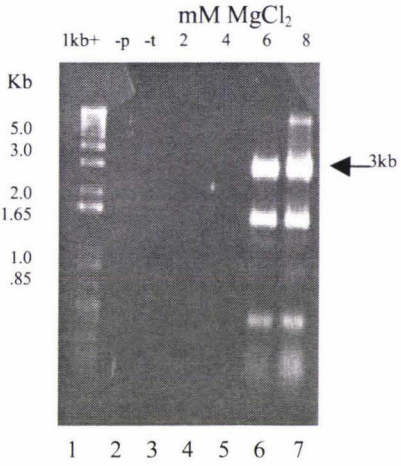
Figure 3.19 Synthesis of C5P Ser to Ala mutation

Figure 3.19a
C5P Ser to Ala PCR A + B



Lane 1 1Kb+ markers, Lane 2 and 3 Ser to Ala forward mutation (PCRA); Lane 4 Ser to Ala reverse mutation (PCRB). 10 μL of each PCR reaction was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 μg/mL) and visualized by UV transillumination. Conditions for amplifiction were: denaturation 94°30", annealing 45°30", extension 72°30", 4, 6, or 8 mM $MgSO_4$; repeated for 30 cycles. Positive and negative controls were used in both PCR reactions (not shown).

Figure 3.19b
C5P Ser to Ala PCR C



Lane 1 1Kb+ markers, Lane 2 PCR control with no primers, Lane 3 PCR control with no template, Lanes 4-7 $MgCl_2$ titration. 10 μL of each PCR reaction was electrophoresed on 1% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with EtBr (0.5 μg/mL) and visualized by UV transillumination. Conditions for amplifiction were: C5P Ser to Ala denaturation 94°30", annealing 54°30", extension 72°1min20", 4, 6, or 8 mM $MgCl_2$; repeated for 30 cycles

The C5P Ser to Ala PCR product was digested with *Bam*HI and *Eco*RI restriction enzymes. The product was ligated into a pGEX6P3 vector cut with the same restriction enzymes. 1 μL aliquots of the ligation mixture were added to 50 μL of *E. coli* XL1 competent cells. The cells were transformed by heat shock treatment and transformants selected by plating on LB-ampicillin plates. Plasmid preparations were made from transformed cells by alkaline lysis (figure 3.20a) Colonies were induced with 1 mM IPTG, lysed by sonication, and the crude cellular extract was analyzed by SDSPAGE to check for protein expression. Only a few colonies were selected, and only one expressed a protein upon induction with 1 mM IPTG corresponding to the size of C5P Ser to Ala (lane 5, figure 3.20b). A protein of 136 kDa was expected consisting of C5P (110 kDa) and the amino terminal GST tag (26 kDa).

Figure -3.20 C5P Ser to Ala mutation

Figure 3.20a plasmid preparation
C5P Ser to Ala transformed cells

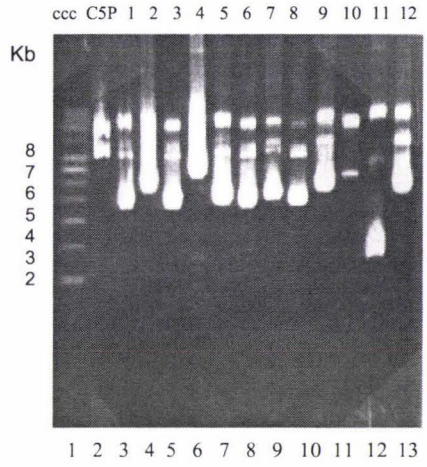


Figure 3.20b C5P Ser to Ala
transformed cells induction

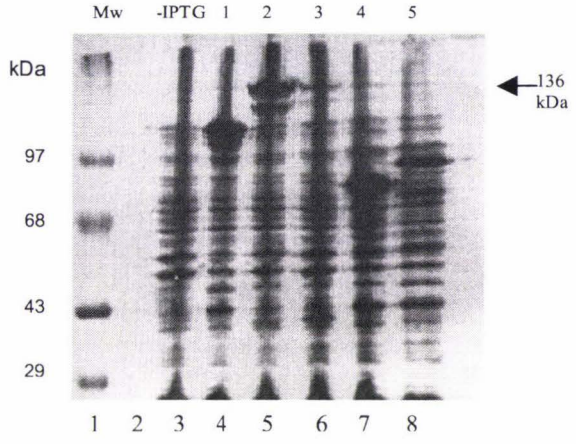


Figure 3.20c Comparison of C5P Ser to
Ala expressed protein to recombinant C5P

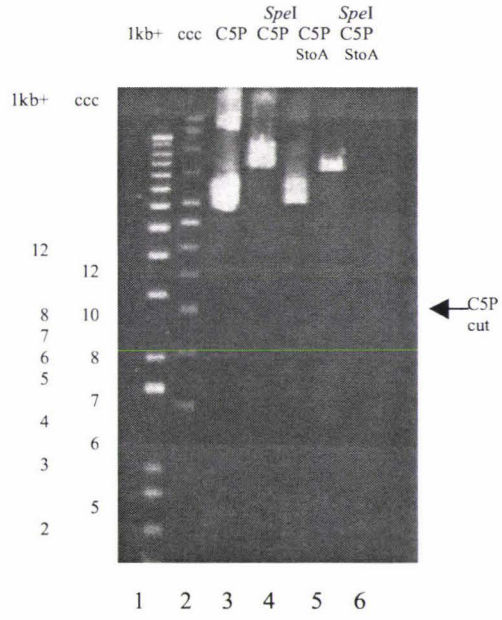


Figure 3.20a
Lane 1 ccc markers, Lane 2 pGEXC5P, Lanes 3-13 samples from pGEXC5PStoA transformed cells prepared by alkaline lysis. 10 µL of each plasmid preparation was electrophoresed on 0.8% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

Figure 3.20b
Whole cell lysate of cells induced with 1 mM IPTG. Lane 1 molecular weight markers; lane 3 uninduced cells; lanes 4-8 cells induced with IPTG. 10 µL of whole cell lysates prepared by sonication were analyzed by SDSPAGE on a 10% acrylamide gel. The gel was stained with Coomassie Blue.

Figure 3.20c
*Spe*I digest of C5P and putative C5P Ser to Ala. Lane 1 1kb+ ladder; Lane 2 ccc ladder; Lane 3 pGEXC5P uncut; Lane 4 *Spe*I digest of pGEXC5P; Lane 5 pGEXC5PStoA uncut; Lane 6 *Spe*I digest of pGEXC5PStoA. 10 µL of each digest was electrophoresed on 0.8% agarose gel at 100 V for 45 minutes in 1x TAE buffer, pH 8. The gel was stained with ethidium bromide (0.5 µg/mL) and visualized by UV transillumination.

Colony 2 expressed two proteins upon induction, one protein appeared to be of the correct molecular weight (136 kDa) while the other was a somewhat lower molecular weight. The expression of two protein suggested that colony 2 was either a mixed culture or was harboring two plasmds. Colonies from a single colony streak of colony 2 were induced and the expressed proteins were analyzed by SDSPAGE. This demonstrated that colony two had been a mixed culture.

The introduction of the Ser512 to Ala512 removes an *Spe*I restriction endonuclease site within the C5P coding region. *Spe*I cuts within the six nucleotide sequence ACTAGT, which is replaced by the sequence ACTGCT with the introduction of the Ser to Ala mutation. To check for a C5P Ser to Ala mutation, a plasmid preparation of the putative transformant was digested with *Spe*I. The plasmid was cleaved to give products identical to pGEXC5P demonstrating that the restriction site was present and therefore did not encode a C5P Ser to Ala mutation. The low transformation frequency suggested the transformed cells could have arisen from rare PCR events. There would have been residual pGEXC5P template in the PCR A and PCR B products. This could have been amplified in PCR C which would account for the absence of a C5P Ser to Ala mutation. This experiment was repeated with gel purification of the PCR A, PCR B, and PCR C products, however, no transformed cells were obtained.

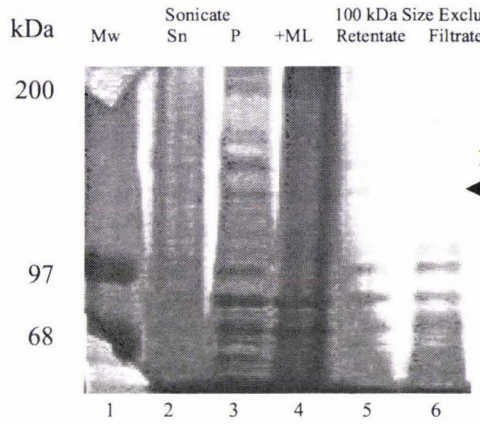
3.7 Purification of C5P off the Cell Wall of Streptococcus pyogenes

Attempts were made to purify the full length C5P, minus the cell wall spanning domains, from *Streptococcus pyogenes* strain B220 to be used as a positive control in activity assays. Initially a method described by Fernandez-Espla *et al*, 2000 to harvest a Lactococcal cell envelope proteinase was attempted. This method involved growth of the cells in nutrient broth with the addition of 20 mM glycine, which would be incorporated into the cell wall during growth and thus weaken the cell wall. Disruption of the cell wall was achieved by grinding with alumina powder, which was later removed by centrifugation, and the proteinases were released with lysozyme. Separating the streptococcal cells from the alumina powder by centrifugation proved problematic as the cells were found in the pellet with the powder, perhaps due to matting of the streptococcal chains. The cell wall was therefore disrupted by sonication. Mutanolysin was substituted for lysozyme to release the proteinase from the cell wall debris. Supernatant harvested after digestion with mutanolysin was analyzed by SDSPAGE. No protein bands corresponding to the expected size of 130 kDa were visible with silver staining.

The second method followed to purify C5P from the cell wall was described by Bohnsack *et al* 1991. This involved the digestion with mutanolysin of an overnight culture of streptococcal cells washed with PBS buffer pH 7.4. The supernatant of this digest, which contained released cell surface proteins, was concentrated by centrifugation through a 100 kDa size exclusion filter. The supernatant was then dialyzed into 0.01 M phosphate buffer pH 7.0 and purified on a hydroxyapatite column. The full length C5P (residues 1-1167) is 130 kDa. A protein corresponding to this size was eluted off the hydroxyapatite column with 0.15 M phosphate buffer, as illustrated in lanes 3 and 4 of figure 3.21b, however the protein sample was contaminated with a higher molecular weight protein, and the dark staining at the bottom of the acrylamide gel demonstrates the presence of unresolved lower molecular weight proteins also. Fractions containing the 130 kDa protein released at 0.15-0.2 M phosphate buffer were pooled and assayed for activity on His-C5a. His-C5a was digested, demonstrating the C5P extracted from the cell wall of *S. pyogenes* has proteolytic activity as illustrated in figure 3.22. The sensitivity of silver staining is 1-10 ng/ band. The final amount of protein dispensed between five 1 mL fractions eluted off the hydroxyapatite column is approximately 2.5 μg, estimated on the basis of its density in SDS-PAGE after silver staining.

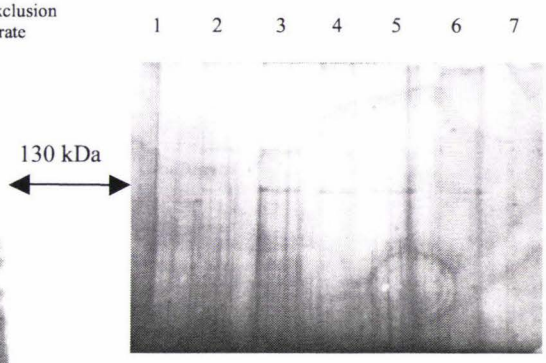
Figure 3.21 Purification of C5P from the cell wall of *S. pyogenes*

Figure 3.21a
Mutanolysin extract of C5P



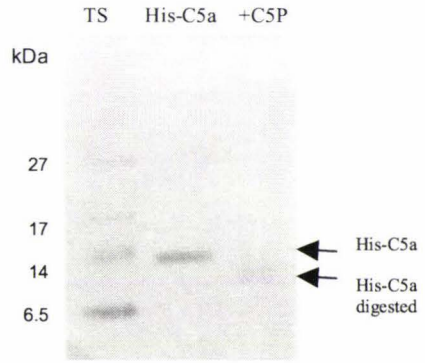
Lane 1 Molecular weight markers; Lane 2 supernatant from sonicated *S pyogenes* B220 cells; Lane 3 pellet from resuspended lysed cells digested with mutanolysin (ML); Lane 4 mutanolysin digest supernatant; Lane 5 mutanolysin digest supernatant concentrated on a 100kDa size exclusion filter (retentate); Lane 6 mutanolysin digest supernatant 100 kDa size exclusion filter filtrate. 10 µL samples were analyzed by SDSPAGE on an 8% acrylamide gel. The gel was stained with Coomassie Blue followed by silver stain.

Figure 3.21b
Mutanolysin extract of C5P run on a hydroxyappetite column



Lanes 1-7 fractions eluted off a hydroxyapetite column. Lane 1 0.15 M phosphate buffer; Lane 4 0.2 M phosphate buffer.
The column was run with a linear flow from 0.01 M phosphate buffer pH 7.0 to 0.2 M phosphate buffer pH 7.0 at a flow rate of 1 mL/min. 1 mL fractions were collected. 10 µL samples were analyzed by SDSPAGE on an 8% acrylamide gel. The gel was stained with Coomassie Blue followed by silver stain.

Figure 3.22 His-C5a digestion assay with C5P purified off *S. pyogenes* cell wall



Lane 1 Tricine Standards; Lane 2 His-C5a control; Lane 3 His-C5a plus C5P from *S. pyogenes* cell wall purified on hydroxyapatite column. 10 µL samples were analyzed by SDSPAGE on an 16% tricine gel. The gel was stained with Coomassie Blue followed by silver stain.

3.8 Chapter Summary

The GST-C5P fusion proteins were cloned and expressed in *E. coli* XL1. The fusion proteins were purified from sonicates of *E.* coli XL-1 transformed with pGEXC5P, pGEXC5Pppd, and pGEXC5Pmod by affinity absorption to glutathione resin. The site-specific protease, PreScission^{TM}, was used to release the recombinant C5P proteins from the GST tag, which remained bound to the glutathione resin. However, lower molecular weight contaminants co-eluted off the glutathione resin. The contaminating proteins would most likely be either breakdown or premature termination products of the recombinant C5P proteins. SEFPLC failed to separate the C5P proteins from contaminants. The recombinant proteins were further purified by IEFPLC. When analyzed by SDS-PAGE the recombinant proteins appeared to be purified to homogeneity. However, the purified recombinant proteins continued to degrade.

Most subtilases are synthesized as pre-pro-enzymes. The 'pre' refers to an amino-terminal signal peptide, which is cleaved during translocation through a cell membrane. Cleavage of the pro-segment is generally required for activation of the subtilase, however this has not been established for C5P. A sample of the C5P prepared for crystallization was electroblotted onto a PVDF membrane for amino-terminal sequencing. This is described in chapter five. The N-terminal sequence was TIRDLN, which corresponds to residues 101 to 106 of C5P. Comparison sequence analysis predicted residues 32-101 of C5P to constitute a putative pro-sequence, Siezen *et al* 1999. The pro-sequence of the closely related subtilases PrtP and PrtB are removed by autocatalytic processing, (Siezen *et al* 1997), which suggests that the pro region of the protease is being processed by one residue towards the amino terminal from the predicted processing site. Both the recombinant C5P and C5Pppd breakdown to a product that may correspond to the loss of the seventy residue pro region (7.7kDa). Cleary *et al* 1992 state that there is no indication that C5P is activated by further processing subsequent to removal of the signal peptide. However, the amino terminal sequencing of C5P released the cell surface by Streptococcal Cysteine Proteinase (SCP) starts at residue 90, (Berge and Bjorck 1995). C5a-mediated granulocyte migration was inhibited by the SCP released C5P,

suggesting that amino acids 32-89 are not required for proteolytic activity, (Berge and Bjorck 1995). This supports the possibility of pro region processing.

Both methods attempted to prepare C5P with a Ser to Ala active site mutation were unsuccessful. The first method involved the digestion of C5Ppd^{S512toA512} clone in a pGEX6P3 vector, that had previously been engineered, and ligation into a pGEX6P3 vector containing a C5P insert cut with the same restriction enzymes. The restriction enzymes *Bgl*II and *Sac*I were used as they had recognition sites that flanked the site of mutation by several residues. Whilst both restriction endonuclease sites were cut within the C5P insert, *Sac*I did not cut within the C5Ppd ^{S512toA512} insert. This was possibly due to a PCR error within this endonuclease restriction site. The second method involved introduction of the mutation by PCR over the Ser 512 codon. This method was unsuccessful, probably due to problems within the PCR reactions, which may explain the low transformation efficiency.

A protein of the approximate molecular weight of C5P was harvested off the cell wall of S. *pyogenes*. This protease demonstrated proteolytic activity on His-C5a. However, the protease was not purified to homogeneity and the protein yield was low; only approximately 2.5 μg of C5P was harvested from an 800 mL culture of *S. pyogenes*.

Chapter 4 Human C5a Synthesis, Expression, and Purification

4.1 Introduction

The substrate of Streptococcal C5a Peptidase is human C5a. C5a is a 74 residue fragment of the protein C5. Synthetic human C5a is available commercially, but it is very expensive. Some synthetic human C5a was purchased for use in activity assays, and as a more economical source of substrate, synthetic human C5a was also cloned.

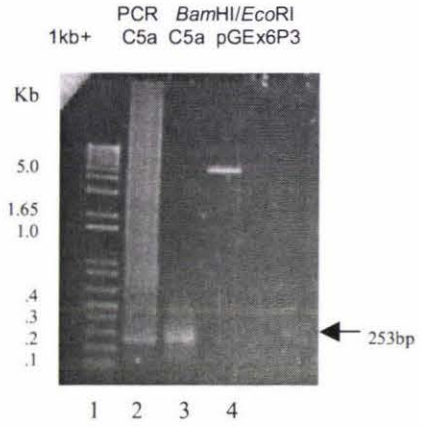
4.2 Synthesis and Expression of Synthetic Human C5a

The synthetic human C5a coding sequence of 253 bp was generated from the ligation of thirteen oligonucleotides that had been 5' phosphorylated using T4 polynucleotide kinase: six in the forward direction, and seven in the reverse (supplied by Dr Jakki Cooney). The overnight ligation was then amplified by PCR with *Pwo* polymerase using primers designed with 5' *Bam*HI and 3' *Eco*RI restriction enzyme sites to facilitate cloning. This process was repeated several times before a substantial PCR product was obtained. Factors affecting the stringency, including the temperature of annealing and the salt concentration, were adjusted in each PCR reaction. The conditions that gave the most homogeneous PCR product were denaturation 94°30", annealing 48°30", extension 70°30", repeated for 30 cycles with a salt concentration of 2 mM $MgSO_4$. A PCR product of the expected 253 bp was synthesized with 2 mM $MgSO_4$ as illustrated in figure 4.1a. The cloning of C5a cDNA was performed following a method previously described (Mandecki et al., 1985), with some variations. The PCR product was digested with *Bam*HI and *Eco*RI and ligated into a pGEX6P3 vector cut with the same restriction enzymes. 1 µL aliquots of the ligation mixture were added to 50 µL of competent *E. coli* XL1 cells. Transformed cells were selected by plating on LB-ampicillin plates. Colonies were induced with 1 mM IPTG and whole cell lysates analysed by SDS-PAGE. A protein of 35 kDa expressed as an amino terminal GST fusion protein was expected (9.1 kDa C5a plus 26 kDa amino

terminal GST tag). Colonies 2, 6, 8, and 9 expressed a protein which ran above the 29 kDa marker and was likely to be GST-C5a fusion protein as illustrated in figure 4.1b.

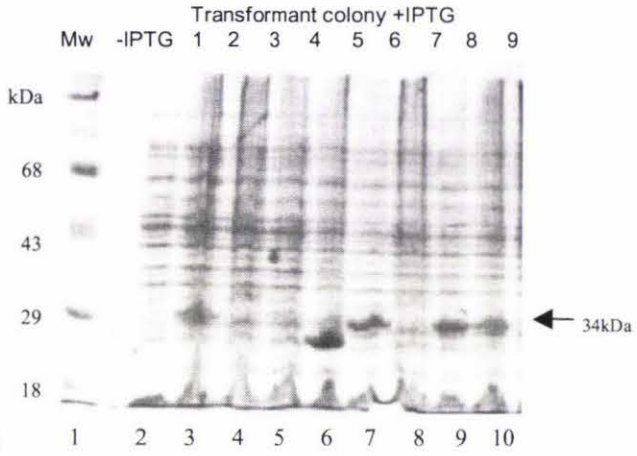
Figure 4.1 Synthesis of GST-C5a

Figure 4.1a Synthesis of Synthetic C5a



Lane 1 1Kb+ markers, Lane 2 PCR of C5a cDNA; Lane 3 *Bam*HI and *Eco*RI digest of C5a; Lane 4 BamHI and EcoRI digest of pGEX6P3. 10 µL samples were electrophoresed on 1% agarose at 100V for 45 minutes in 1x TAE buffer, pH8, stained with ethidium bromide and visualized by UV transillumination. Conditions for amplification were: denaturation 94°30", annealing 48°30", extension 70°30", 2 mM $MgSO_4$, repeated for 30 cycles.

Figure 4.1b Expression of GST-C5a transformant colonies



Whole cell lysate of cells induced with 1mM IPTG. Lane 1 molecular weight markers; lane 2 cells pre-induction; lanes 3-10 cells expressing GST-C5a. 10 µL samples were analyzed by SDS-PAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

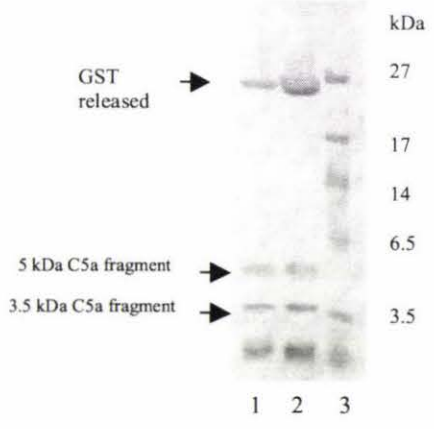


Figure 4.1c
Lanes 1 and 2 PreScission™ cleaved GST-C5a; Lane 3 Tricine Standards. 10 µL samples were analyzed by SDSPAGE on a 16.5% Tricine gel. The gel was stained with Coomassie Blue and Silver stain. (The band present at the bottom of all lanes is glycerol which is a component of the gel running buffer.)

DNA sequencing of the synthetic human C5a cDNA isolated from two transformed *E. coli* revealed a stop codon in the middle of the translated sequence, which can be

tracked back to a mistake in one of the reverse oligos. The error was corrected by PCR across the sequence error using primers with the correct sequence. The corrected synthetic C5a cDNA was ligated into a pGEX6P3 vector and expressed as a GST fusion protein. Protein expression was successful, and the GST-C5a fusion protein could be purified from the soluble cellular proteins on glutathione resin. Release of synthetic C5a off Glutathione resin with PreScission™ protease was analysed by SDS-PAGE on a Tricine gel (16.5%), which was stained with silver stain. Two protein fragments were released off the Glutathione resin one of an apparent molecular weight of 3.5 kDa, the other of approximately 5 kDa, which combined are approximately the expected molecular weight of the synthetic human C5a, illustrated in figure 4.1c.

The corrected synthetic human C5a cDNA was then cloned into a pProExHtb vector as illustrated in figure 4.2a. This plasmid was called pProExC5a. The plasmids were digested with *Bam*HI and *Eco*RI and analyzed by SDS-PAGE to confirm the insert was of the correct size (lane 4, figure 4.1a). The ligation of C5a cDNA into the pProExHtb vector was added to competent cells. Transformed cells were selected by plating on LB-ampicillin plates. Transformed colonies were screened for expression by induction with IPTG and analysis by SDS-PAGE. A protein of 12.5 kDa was expected consisting of C5a (9.1 kDa) plus the His tag (3.4 kDa). Expression of a protein with an apparent molecular weight of 14 kDa was induced with IPTG as illustrated in figure 4.2b. This protein was likely to be the His-C5a fusion protein.

Figure 4.2 Expression of Histidine tagged C5a

Figure 4.2a Cloning of His-C5a

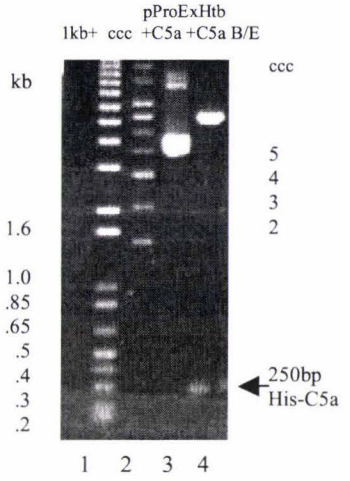


Figure 4.2b
His-C5a Induction

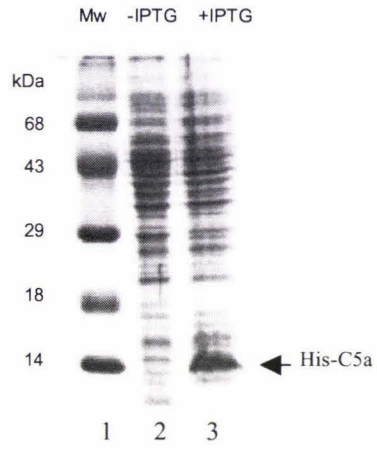


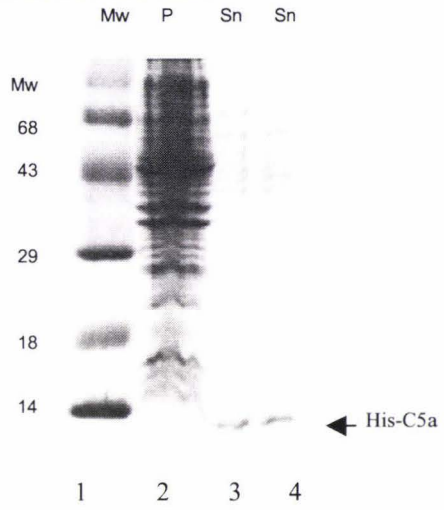
Figure 4.2a
Lane 1 1Kb+ markers; Lane 2 ccc markers; Lane 3 pProExC5a; Lane 4 pProExC5a *Bam*HI and *Eco*RI digest of. 10 µL samles were electrophoresed on a 1% agarose gel at 100V for 45 minutes in 1x TAE buffer, pH8. The gel was stained with ethidium bromide and visualized by UV transillumination.
Figure 4.2b
Whole cell lysate of cells induced with 1mM IPTG. Lane 1 molecular weight markers; lane 2 cells pre-induction; lanes 3 cells expressing His-C5a. 10 µL samples were analyzed by SDSPAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

4.3 Purification of Synthetic Human C5a

Cells expressing His-C5a protein were difficult to lyse by sonication. Analysis by SDS-PAGE revealed that His-C5a was in the insoluble pellet. This suggested that His-C5a forms inclusion bodies within *E. coli* during expression. His-C5a was therefore purified from inclusion bodies pelleted during centrifugation of cellular lysates. The purification of His-C5a from inclusion bodies was performed following a method previously described (Toth et al., 1994). This process involved denaturation with 6 M guanidinium-HCl to solubilise His-C5a followed by renaturation with 1 mM oxidized and 1 mM reduced glutathione. Protein bands of the appropriate molecular weight are visible in lanes 3 and 4 of figure 4.3a. The soluble His-C5a was then purified by Chelating Sepharose Chromatography by elution with imidazole. C5a was eluted with 100 mM and 200 mM imidazole, with the latter fractions containing homogeneous protein, as illustrated in lanes 7 and 8 of figure 4.3b. Aliquots of the His-C5a eluted in fractions with 200 mM of imidazole were stored at -70° for use in activity assays.

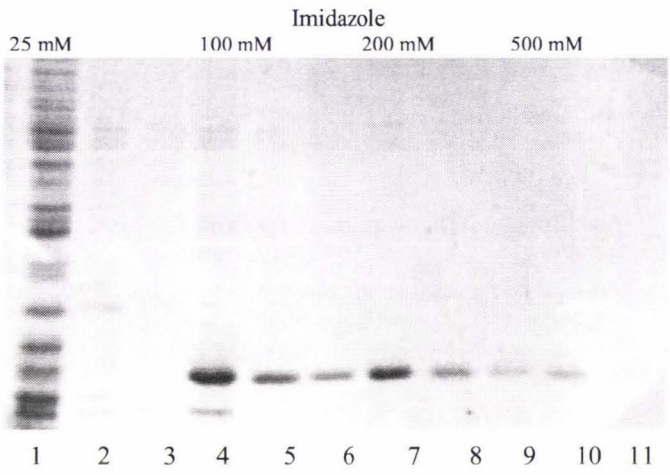
Figure 4.3 His-C5a Chelating Sepharose Chromatography

Figure 4.3a His-C5a Purification from inclusion Bodies



Lane 1 Mw markers; Lane 2 pellet after overnight renaturation of denatured inclusion body preparation; Lanes 3 and 4 supernatant from renatured inclusion body preparation. 10 µL samples were analyzed by SDSPAGE on a 15% acrylamide gel. The gel was stained with Coomassie Blue.

Figure 4.3b His-C5a Chelating Sepharose Chromatography



Chelating sepharose chromatography of His-C5a prepared from inclusion bodies. Lanes 1-3 25 mM imidazole; Lanes 4-6 100 mM imidazole; Lanes 7-9 200 mM imidazole; Lanes 10 and 11 500 mM imidazole. 10 µL samples were analyzed by SDSPAGE on a 15% acrylamide gel. The gel was stained with Coomasie Blue. Start buffer 25 mM imidazole. Final Buffer 1 M imidazole. Flow rate 1mL/min. Chromotography was carried out on a BioRad econosystem fitted with a column packed with chelating sepharose.

4.4 Chapter summary

Synthetic human C5a cDNA was cloned and expressed as a GST fusion protein. Two protein fragments were released off the Glutathione resin one of an apparent molecular weight of 3.5 kDa, the other of approximately 5 kDa, which combined are approximately the expected molecular weight of the synthetic human C5a. The synthetic gene codes for a protein sequence (AARISL/GPR) that resembles the recognition site of PreScission (LEVLPQ/GPL) in that both have a hydrophobic patch followed by Gly in the P1 site, and Pro in the P2 site. The sizes of the synthetic human C5a products suggest PreScission™ may be recognizing this putative site. C5a peptidase cleaves a seven residue fragment from the carboxyl terminus of C5a. Release of this fragment from the 35 kDa GST-C5a fusion protein would have been

difficult to detect by SDS-PAGE, therefore the GST-C5a fusion protein would not have been suitable for C5P activity assays.

Synthetic human C5a was cloned as a His tagged fusion protein. The small size of the His tag (3.4 kDa) would enable the cleavage of His-C5a by C5P to be analyzed by SDS-PAGE. His tagged C5a was purified from inclusion bodies by a process of denaturation with 6 M guanidinium-HCl to solubilise His-C5a followed by renaturation with 1 mM oxidized and 1 mM reduced glutathione. His-C5a was further purified by Chelating Sepharose Chromatography. Aliquots of the His-C5a eluted in fractions with 200 mM of imidazole were stored at -70^{o} for use in activity assays.

Chapter 5 : Characterization of Recombinant C5P

5.1 Introduction

C5P shares homology with the subtilisins, for which structures are available. The catalytic or protease domain of C5P is the segment with sequence homology to the subtilisins within which, lactic acid bacteria cell envelope proteinases (CEPs) all have a few large inserts outside of a conserved core. These inserts are predicted to be surface loops. Whereas the subtilisins in general show broad substrate specificity profiles, the CEPs of lactic acid bacteria demonstrate remarkable substrate specificity. The greater specificity of the CEPs is held attributable to changes in the variable regions within the structurally conserved core and to the presence of the A-domain. The research hypothesis was that the variable region insertions and deletions, and the A domain influence the specificity of streptococcal C5a peptidase. This was addressed by studying the biochemical and structural properties of C5P. The study of C5P commenced with cloning, expression, and purification of recombinant C5P fragments (refer to chapter 3).

C5P is of the pyrolysin family of subtilases (subtilisin like serine proteases), for which no structural information is available. Comparative sequence analysis demonstrated C5P shares sequence homology with the subtilase *Lactococcus lactis* PrtP, (Siezen, 1999; Siezen et al., 1993; Siezen et al., 1991; Siezen and Leunissen, 1997). In contrast to subtilisins, C5P and PrtP are both highly specific proteases. With respect to the conserved core structure of subtilisins and subtilases, both PrtP and C5P have large insertions in variable regions (VRs), particularly the VR13. Siezen *et al* (1999) postulated that variations within the VRs modulate substrate specificity of subtilases. Furthermore, the characteristic cleavage pattern of casein by PrtP appears to be modulated by the A domain. C5P is a highly specific endopeptidase, which cleaves the chemotactic and macrophage-activating anaphylatoxin C5a on the carboxyl side of His67 (Cleary et al., 1992). Homology to PrtP suggests the specificity of C5P may also be modified by the A domain and the VRs (Siezen, 1999).

Structural studies commenced with screening for crystallization conditions of the purified recombinant C5P fragments. Resolution of crystollography data was beyond the scope of this project. However, if the hypothesis holds true, the structure of C5P

could reveal potential interactions of VRs and the A domain with substrate binding sites and\or the substrate.

Activity of the recombinant C5P fragments was assessed using *in vitro* digestion assays. The biological effect of C5a peptidase on the chemotactic potential of C5a was assessed by studying at polymorphonuclear neutrophil migration under agarose. Biological activity would suggest that recombinant C5P had assumed the correct conformation. In absence of the A-domain it was hypothesized that the specificity of C5P would be relaxed. Specificity is still controversial and has only been addressed by one group, who presented two different answers to the question of where the scissile bond lies (reported to be either between His67 and Lys 68 or Lys68 and Asp69 of C5a). The activity and specificity of C5P, and C5P in absence of the A-domain, was therefore studied.

5.2 Screening for Crystallization conditions for C5P

Crystallization conditions for recombinant C5P protein were screened using the Hampton Crystal Screens one and two employing the hanging drop vapour diffusion method. The C5P protein had been purified in PBS buffer, but the phosphate in PBS readily crystallizes, and this would interfere with protein crystallization assays. Therefore, the C5P protein was exchanged into 50 mM HEPES pH 7.5 by multiple washes in a Centricon YM10, which has a molecular weight cut off of 10 kDa. The C5P protein started to precipitate as the protein concentration increased. 0.1 M NaCl was added, which improved the solubility of the protein. 1 µL of C5P protein concentrated to 15.3 mg/mL was mixed with 1 µL of the precipitant solution and suspended as a drop from a cover slip above 500 µL of the precipitant solution in a well of a 24 well Linbro plate, forming a sealed chamber. The plates were stored in the dark at room temperature and examined weekly by light microscopy.

Needles formed in 2.0 M ammonium sulfate at room temperature, which was used as a starting point for further screening. The ammonium sulfate screen ranges over pH 4-pH 9 in steps of 1 pH unit, and through 0.8 M-3.2 M ammonium sulfate in 0.8 M steps, resulting in 24 conditions as illustrated in table 5.1. Hexagonal crystals formed in 2.4 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.0, condition 16 of the ammonium sulfate/pH screen. This condition was used as the starting point for further screening. The ammonium sulfate/Hepes screen ranged over pH 6.5-pH 7.5 in 0.2 pH unit steps, and through 1.6 M-2.7 M ammonium sulfate in 0.2 M steps, resulting in 42 conditions as illustrated in table 5.2. Hexagonal crystals formed in 2.6 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 6.9, condition 33 of the ammonium sulfate/ Hepes screen. Cubic crystals formed in 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1, condition 40 of the ammonium sulfate/ Hepes screen, as illustrated in figure 4.1. The crystal form of C5P that crystallized in condition 40 was p-hexagonal (p=primative) form (3-6x axis of symmetry).

X-ray diffraction of this crystal was performed by Dr Jakki Cooney and Dr Todd Kagawa using a Rigaku R-AxisIIC detector with a RU-200B rotating anode X-ray source at the institute of molecular biosciences, Massey University. Data were processed using HKL and CCCP4 data processing packages. The data was not of sufficient resolution to be used towards providing a solution of the structure. Although C5P shares homology with the subtilisins, for which structures are available. The conserved core of the subtilisins is interrupted with insertions and deletions in variable regions, therefore Molecular Replacement (MR) may not provide a solution of the structure. Dr Jakki Cooney and Dr Todd Kagawa expressed recombinant C5P in a strain of *E. coli* auxotrophic for methionine and grown on minimal media supplemented with amino acids using selenomethionine in place of methionine. Selenium is an anomalous scatterer and can give phase information that can produce high quality electron density maps by Multiwavelength Anomalous Dispersion (MAD). This approach was beyond the scope of this thesis and may be pursued by others at a later date.

Table 5-1 Ammonium sulfate/pH screen

$(NH_4)_2SO_4$/M	PH UNITS					
	4	**5**	**6**	**7**	**8**	**9**
0.8	1	2	3	4	5	6
1.6	7	8	9	10	11	12
2.4	13	14	15	16	17	18
3.2	19	20	21	22	23	24

The ammonium sulfate/pH screen ranges over pH 4-pH 9 in 1 pH unit steps, and through 0.8 M-3.2 M ammonium sulfate in 0.8 M steps, resulting in 24 conditions. Hexagonal crystals formed in 2.4 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.0, condition 16.

Table 5-2 Ammonium sulfate/Hepes screen

$(NH_4)_2SO_4$/M	0.2 M HEPES /PH UNITS					
	6.5	**6.7**	**6.9**	**7.1**	**7.3**	**7.5**
1.6	1	2	3	4	5	6
1.8	7	8	9	10	11	12
2.0	13	14	15	16	17	18
2.2	19	20	21	22	23	24
2.4	25	26	27	28	29	30
2.6	31	32	33	34	35	36
2.7	37	38	39	40	41	42

The ammonium sulfate/Hepes screen ranged over pH 6.5-pH 7.5 in 0.2 pH unit steps, and through 1.6 M-2.7 M ammonium sulfate in 0.2 M steps, resulting in 42 conditions. Hexagonal crystals formed in 2.6 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 6.9, condition 33. Cubic crystals formed in 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1, condition 40.

Figure 5-1 C5P crystallized in 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1

1 µL of C5P protein concentrated to 15.3 mg/mL was mixed with 1 µL of the precipitant solution and suspended as a drop from a cover slip above 500 µL of 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1 in a well of a 24 well Linbro plate, forming a sealed chamber. The plates were stored in the dark at room temperature. The crystal form of C5P crystallized, is p-hexagonal (p=primative) form (3-6x axis of symmetry).

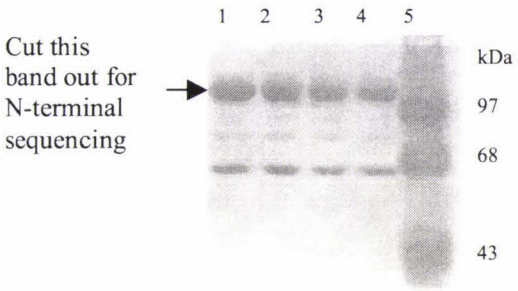
It was possible that crystallization was occurring with a breakdown product of the recombinant C5P protein as the crystals took several weeks to months to form. Considerable breakdown of the unscreened C5P protein was happening over this period (demonstrated by a sample analysed by SDS-PAGE gel stained with Coomassie blue). To determine if it was a breakdown product of C5P that was crystallizing, a crystal was dissolved and analysed by SDS-PAGE gel, and the size of the principal constituent/s were analysed.

A sample of the C5P prepared for crystallization was electroblotted onto a PVDF membrane, illustrated in figure 5.2a for amino-terminal sequencing at Protein Sequencing Services, Massey University. The N-terminal sequence was TIRDLN. This sequence corresponds to residues 101 to 106 of C5P. Comparison sequence analysis predicted residues 32-101 of C5P to constitute a putative pro-sequence. C5P crystallized in 2.4 M $(NH_4)_2SO_4$ and 0.2 M Hepes pH 8 and in 2.4 M $(NH_4)_2SO_4$ and 0.2 M Hepes pH 7.0 were washed and dissolved in crystallization buffer and analyzed by SDS-PAGE as illustrated in figure 4.2b. The C5P crystallized in 2.4 M $(NH_4)_2SO_4$ and 0.2 M Hepes pH 7.0 (Lane 4) was the same molecular weight as the band excised from the crystallization sample for amino terminal sequencing (Lane 2). The lower molecular weight band present in the sample of C5P crystallized in 2.4 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 8 (Lane 3) could be from contamination by the mother liquor. As the majority of the crystallized protein was the same molecular weight as the purified

sample prior to crystallization, it was concluded that the crystals were in fact C5P minus the pro-region.

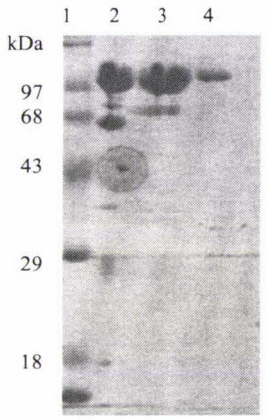
Figure 5-2 Comparison of crystallized C5P and C5P prepared for crystallization

Figure 5.2a C5P crystallization sample on PVDF membrane



C5P electroblotted from SDS-PAGE onto PVDF membrane. Lanes 1-4 0.5uL C5P prepared for crystallization; Lane 5 molecular weight markers.

Figure 5.2b Comparison of crystallized C5P and C5P prepared for crystallization



C5P crystallized. Lane 1 molecular weight markers; Lane 2 C5P sample prepared for crystallization; Lane 3 C5P crystallized in 2.4 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 8; Lane 4 C5P crystallized in 2.4 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.0.

Purified recombinant C5Pmod concentrated to 0.87 mg/mL was also screened for crystallization conditions with the Hampton Crystal Screens one and two employing the hanging drop vapour diffusion method following the method described for recombinant C5P. No crystals were formed in these screens, possibly because the protein was not concentrated enough. It is suggested that the protein be concentrated to 1 mg/mL for screening (Drenth). Recombinant C5Pppd was not screened for crystallization conditions as the protein was not stable after purifying to homogeneity.

5.3 Activity assays

5.3.1 In vitro digestion assays

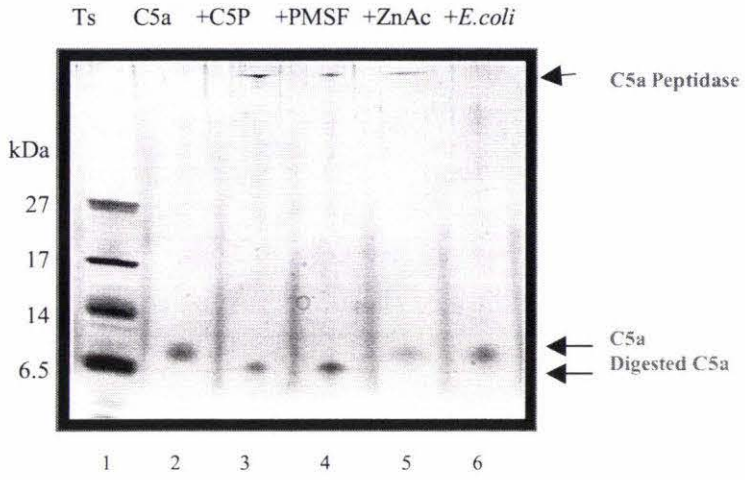
In order to determine the effect on substrate specificity of the A domain of C5P, purified recombinant full length C5P and C5P without the A domain (C5Pppd) were compared for specificity of activity in *in vitro* digestion assays, and in under agarose migration assays. Initially digestion products were analysed on Tricine gels with silver staining. Mass spectrometry was used to confirm the molecular weight of the peptide fragment released. Under agarose migration assays involve the migration of polymorphonuclear neutrophils through agarose towards the chemotactic factor C5a, which can be observed at 40X magnification. The chemotactic potential of C5a is inactivated when cleaved by C5a peptidase. Based on observations of subtilases closely related to C5P, any variation or relaxation found in substrate specificity was expected to correlate with differences in sequence within either variable regions (VR) in the structurally conserved catalytic domain, or the A domain. Synthetic human C5a was both purchased from Sigma and cloned as a His tagged fusion protein for use in these assays.

5.3.2 In vitro Activity Assays

The substrate and protease were incubated in 50 mM Tris pH 7.5 for 2 hours at 37°C. Aliquots of the digest mixture were analyzed by SDS-PAGE on Tricine gels stained with Coomassie blue and Silver stain. Aliquots of the His tagged C5a digest mix were purified by reverse phase high pressure liquid chromatography (RPHPLC) and analyzed by Mass Spectrometry.

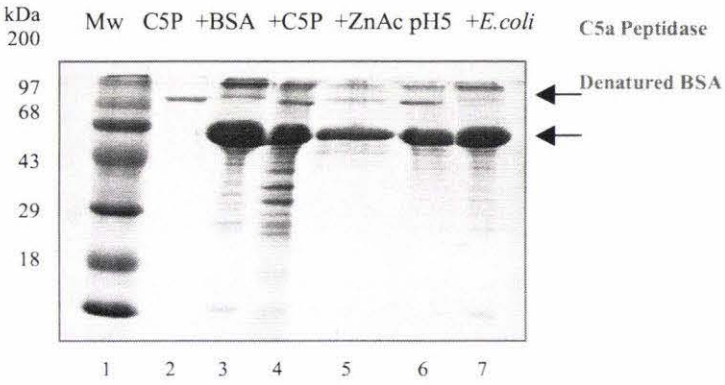
Figure 5-3 *In vitro* digestion activity of recombinant C5P

Figure 5.3a C5a Peptidase Digestion of C5a



Lane 1 Tricine Standards; Lane 2 C5a control; Lane 3 C5a (0.01 mg/mL) + C5P (0.01 mg/mL); Lane 4 C5a (0.01 mg/mL) + C5P (0.01 mg/mL) + PMSF (0.2 mg/mL); Lane 5 C5a (0.01 mg/mL) + C5P (0.01 mg/mL) + ZnAc (0.1M); Lane 6 C5a (0.01 mg/mL) + *E. coli* cell extract (0.27 mg/mL). 10 uL samples were analyzed by SDSPAGE on a 16.5% Tricine gel. The gel was stained with Coomassie Blue and Silver stain.

Figure 5.3bC5a Peptidase Digestion of BSA



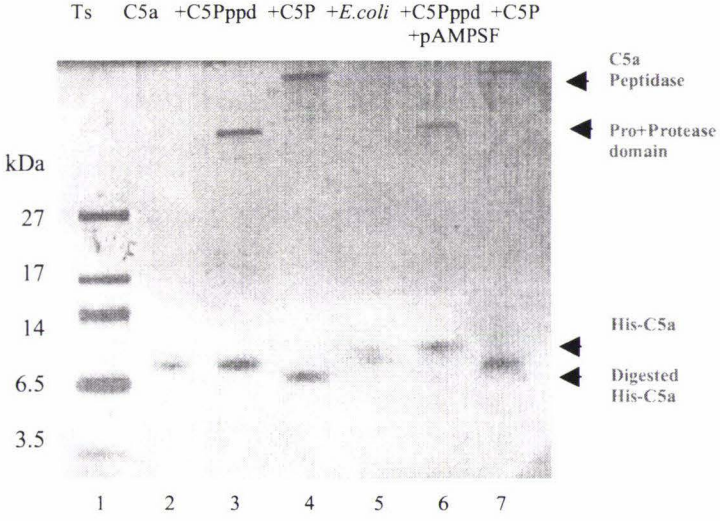
Lane 1 Molecular weight markers; Lane 2 C5P (0.1 mg/mL); Lane 3 BSA (1.25 mg/mL); Lane 4 BSA (1.25 mg/mL) + C5P (0.1 mg/mL); Lane 5 BSA (1.25 mg/mL) + C5P (0.1 mg/mL) + ZnAc (0.1M); Lane 6 BSA (1.25 mg/mL) + C5P (0.1 mg/mL) at pH5; Lane 7 BSA (1.25 mg/mL) + C5P (0.1 mg/mL) + *E. coli* cell extract (0.27 mg/mL). 10 uL samples were analyzed by SDSPAGE on a 10% acrylamide gel. The gel was stained with Coomassie Blue.

Both commercially purchased synthetic C5a and His tagged C5a were digested by recombinant C5P. The scissile bond for C5P is reported to be either between His67 and Lys68 or Lys68 and Asp69 of C5a with the release of a six or seven residue peptide from the carboxyl terminus of C5a. The decrease in gel mobility of the digested products suggests the release of a small peptide fragment as illustrated in figure 5.3a lane 3 and figure 5.3 lanes 4 and 7. Recombinant C5P also demonstrates proteolysis of denatured BSA as illustrated in figure 5.3b lane 4. This activity has been reported for native C5P. C5P activity was inhibited by ZnAc, as illustrated in figure 5.3 lane 5, and ZnCl. The zinc effect has been suggested to be caused by oxidation of sulfhydryl groups on C5a. C5P was also inhibited at pH 5 which is expected as this pH is below the pI of the active site histidine residue, as illustrated in figure 5.3b lane 6. Inhibition is also demonstrated by pAMPSF which is an

irreversible inhibitor of serine proteases with lysine or arginine substrate specificity, as illustrated in figure 5.4a. However, C5P activity was not inhibited by the serine protease inhibitor PMSF, as illustrated in figure 5.3a lane 4. Some non-specific degradation of C5a occurs with the *E. coli* extract, and some C5a appears to be lost during the reaction without the addition of the proteases, as illustrated in figure 5.4a lane 2. The loss of C5a could be due to adherence to the sides of the plastic microcentrifuge tubes during incubation from which the addition of the proteases may offer some protection.

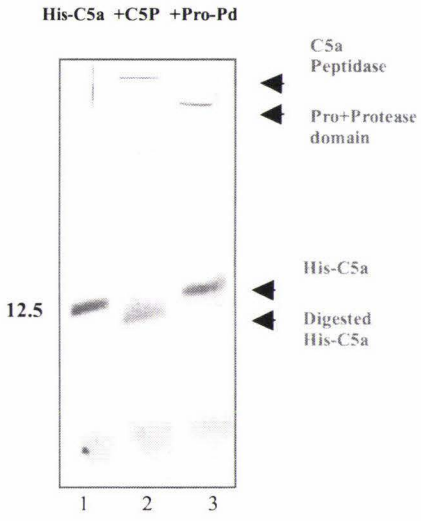
Figure 5-4 *In vitro* digestion activity of recombinant C5P in presence and absence of A-domain

Figure 5.4a



Lane 1 Tricine Standards; Lane 2 C5a control; Lane 3 C5a (0.01 mg/mL) + C5Pppd (0.027 mg/mL); Lane 4 C5a (0.01 mg/mL) + C5P (0.02 mg/mL); Lane 5 C5a (0.01 mg/mL) + *E. coli* cell extract (0.27 mg/mL); Lane 6 C5a (0.01 mg/mL) + C5Pppd (0.014 mg/mL) + pAMPSF (8.3 mM); Lane 7 C5a (0.01 mg/mL) + C5P (0.01 mg/mL) + pAMPSF (8.3 mM). 10 uL samples were analyzed by SDSPAGE on a 16.5% Tricine gel. The gel was stained with Coomassie Blue and Silver stain.

Figure 5.4b



Lane 1 His-C5a control (0.4 mg/mL); Lane 2 His-C5a (0.4 mg/mL) + C5P (0.01 mg/mL); Lane 3 His-C5a (0.4 mg/mL) + C5Pppd (0.014 mg/mL)
10 uL samples were analyzed by SDSPAGE on a 16.5% Tricine gel. The gel was stained with Coomassie Blue and Silver stain.

A synthetic gene for human C5a was synthesized by PCR. This gene was cloned, expressed and the protein purified from inclusion bodies as a His tagged protein. His

tagged C5a was cleaved by recombinant C5a peptidase. However the recombinant pro+protease domains of C5a peptidase demonstrated no activity on commercial synthetic C5a or His-C5a, as illustrated in figure 5.4a lane 3 and figure 5.4b lane 3.

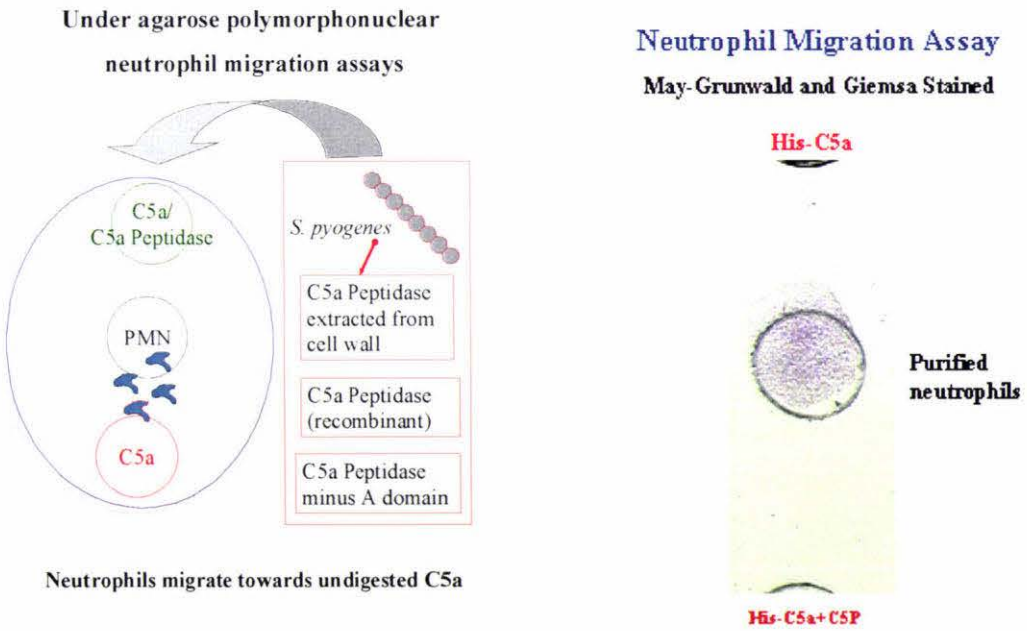
Aliquots of the His tagged C5a digest mix were purified by RPHPLC and analyzed by Electrospray Mass Spectrometry at Protein Sequencing Services, Massey University. The expected molecular weight of His tagged C5a if the scissile bond for C5P is between His67 and Lys 68 was computed by the ProtParam tool provided by ExPASy to be 9381.7 Da. There are seven cysteine residues in C5a, six of which are present as disulfides (Zhang *et al*, 1997). Therefore, the expected molecular weight would be 9375.7, accounting for the loss of six hydrogen atoms. The molecular weight of the released fragment analyzed by Mass Spectrometry was 9391 Da, accurate to 1 Da/10,000 Da. The predicted molecular weight of the digested His tagged C5a is 15.3 Da less than that given by Mass Spectrometry. There are two methionine residues in C5a, one of which may have become oxidized to sulphoxide by the addition of an oxygen atom from hydrogen peroxide. This is a common post-translational modification (reported at the Delta Mass website, a database of protein post translational modifications, http://www.abrf.org/). The His tagged C5a sample was subjected to conditions which may have caused oxidation to occur. If one of the methionine residues was oxidized the expected molecular weight would be 9391, accounting for the addition of an oxygen atom. The discrepancy between molecular weights is therefore within the range of accuracy for Mass Spectrometry. Oxidation of methionine residues can be prevented by the addition of an excess of methionine to out compete the available O_2.

5.3.3 Neutrophil Migration Assays

In order to relate specificity studies of C5a peptidase to biological effect the chemotactic potential of the breakdown products of C5a was investigated. This was assessed by studying neutrophil migration under agarose. Under agarose migration assays involved the migration of polymorphonuclear neutrophils (PMNs) through agarose towards the chemotactic factor C5a.

Samples from C5P digest mixtures, undigested recombinant His tagged C5a, and commercially purchased synthetic human C5a were analyzed for chemotactic potential in under-agarose neutrophil migration assays. Human blood was collected from a healthy volunteer into heparinised tubes. PMNs were purified from fresh blood on a ficoll-paque density gradient (Pharmacia Biotech). Contaminating red blood cells were removed by hypotonic lysis before washing with balanced salt solution. The cells from 2 mL of blood were resuspended in 1 mL of Eagles MEM. PMN concentration and purity was assessed by a Coulter Cell Counter. 10 uL of the C5a samples and PMNs were put into wells cut into agarose in Petri dishes and incubated at 37°C in 5% CO_2 for three hours. Migration was assessed by staining with May-Grunwald and Giemsa stains and examining at 100 x magnification.

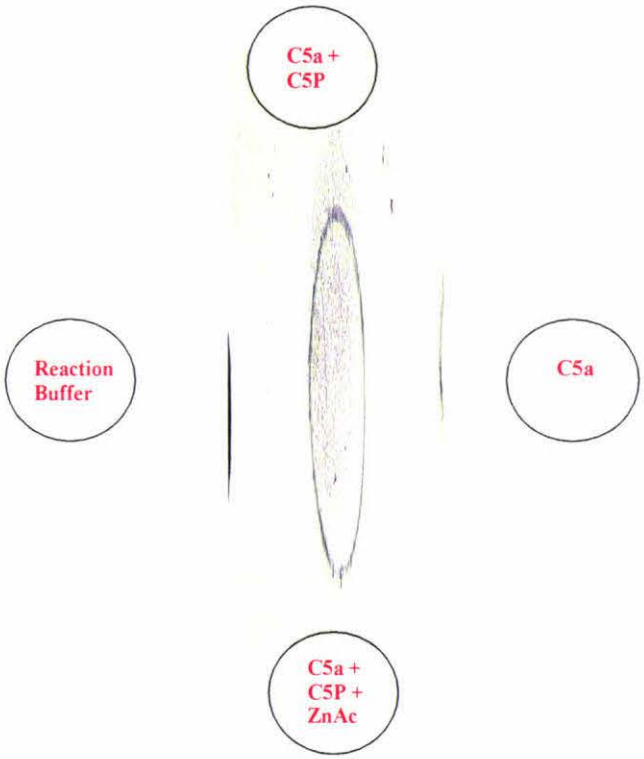
Figure 5-5 Neutrophil migration assays using His-C5a as chemotaxin



Polymorphonuclear neutrophils (PMNs) were purified from fresh blood on a ficoll-paque density gradient. Contaminating red blood cells were removed by hypotonic lyses. 10uL of C5a samples and PMNs ($1.2x10^6$/mL) were put into wells cut into agarose. The Petri dishes were incubated at 37°C in 5% CO_2 for three hours. Migration was assessed by staining with May-Grunwald and Giemsa stains and examining at 100x magnification.

The chemotactic potential of His-C5a and commercial C5a was inactivated when cleaved by recombinant C5a peptidase as illustrated in figures 5.5 and 5.6. The chemotactic potential of C5a was not affected after addition of C5P inactivated by ZnAc as illustrated in figure 5.6. The well with reaction buffer in figure 5.6 was included as a spontaneous migration control. Therefore the recombinant His-C5a was shown to have appropriate biological activity which was destroyed by inactivation by recombinant C5P. Thus the recombinant C5P also had appropriate biological activity.

Figure 5.6 Neutrophil migration assays using commercial C5a as chemotaxin



Polymorphonuclear neutrophils (PMNs) were purified from fresh blood on a ficoll-paque density gradient. Contaminating red blood cells were removed by hypotonic lyses. 10uL of C5a samples and PMNs ($1.0x10^6$/mL) were put into wells cut into agarose. The Petri dishes were incubated at 37°C in 5% CO_2 for three hours. Migration was assessed by staining with May-Grunwald and Giemsa stains and examining at 100x magnification.

5.4 Chapter summary

C5P is of the pyrolysin family of subtilases, for which no structural information is available. After 3 rounds of screening, crystallization conditions were found for recombinant C5P. Cubic crystals formed in 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1. The crystal form of C5P was p-hexagonal (p=primative) form (3-6x axis of

symmetry). Resolution of crystallography data was beyond the scope of this project. However, if the hypothesis holds true, the structure of C5P should reveal potential interactions of VRs and the A domain with substrate binding sites and/or the substrate.

Activity of purified recombinant C5a peptidase (minus cell wall spanning domains) and the pro+protease domains were assessed in *in vitro* digestion assays. Commercially purchased synthetic human C5a and recombinant His tagged C5a were used as substrate for recombinant C5P and C5Pppd proteins. Both commercially purchased synthetic C5a and His tagged C5a were digested by recombinant C5P. Recombinant pro+protease domains of C5a peptidase demonstrated no activity on commercial synthetic C5a or His-C5a. The presence of the modifier or A-domain is reported to have a stabilizing effect on the protease domain in closely related subtilases, which may account for this demonstrated lack of activity. Both commercial synthetic C5a and His tagged C5a were digested by recombinant C5P. The shift down in electrophoretic mobility of the digested products suggests the release of a small peptide fragment.

Aliquots of the His tagged C5a digest mix were purified by RPHPLC and analyzed by Mass Spectrometry. The expected molecular weight of His tagged C5a if the scisscile bond for C5P is between His67 and Lys 68 is 9375.7. The molecular weight of the released fragment analyzed by Mass Spectrometry was 9391 Da, accurate to 1 Da/10,000 Da. This difference could be accounted for by the oxidation of a methionine residue to the sulphoxide. The discrepancy between molecular weights is therefore within the range of accuracy for Mass Spectrometry, and supports the scissile bond to be between His67 and Lys68.

The recombinant pro+protease domains of C5a peptidase demonstrated no activity on commercial synthetic C5a or His-C5a. The function of the modifier or A-domain is unknown, however it is speculated to be essential for proteolytic activity and to be involved in regulating proteolytic activity and/or specificity of the protease (Seizen *et al*, 1999). The demonstrated lack of activity of C5Pppd supports the speculation that the A-domain is essential for proteolytic activity.

Recombinant C5P also demonstrates proteolysis of denatured BSA. This activity has been reported for native C5P. C5P activity was inhibited by ZnAc and ZnCl, the Zn effect is suggested to be caused by oxidation of sulfhydryl groups on C5a. C5P was also inhibited at pH 5 which is expected as this pH is below the pI of the active site Histidine residue. Inhibition was also demonstrated by pAMPSF which is an irreversible inhibitor of serine proteases with lysine or arginine substrate specificity. PMSF, a serine protease inhibitor, did not demonstrate inhibition of C5P at or 2x above its effective concentration of 100 μg/mL. Group B streptococci (GBS) C5a peptidase has been reported to be sensitive to PMSF (Bohnsack *et al*, 1991). However, GBS C5P is distinct from the Group A streptococci (GAS) (Cleary *et al* 1992). GAS C5a peptidase has been reported to be sensitive to calcium chloride, manganese chloride and zinc acetate, and not sensitive to EDTA (Stafslien and Cleary, 2000). GAS C5P sensitivity to other inhibitors has not been reported.

In order to relate specificity studies of C5a peptidase to biological effect the chemotactic potential of the breakdown products of C5a was investigated. This was assessed by looking at neutrophil migration under agarose. The chemotactic potential of C5a was not affected after addition of C5P inactivated by ZnAc. The chemotactic potential of His-C5a and commercial C5a was inactivated when cleaved by recombinant C5a peptidase. Recombinant C5P therefore demonstated appropriate biological activity.

Chapter 6 Summary and General Discussion

Streptococcal C5a peptidase is a cell surface subtilisin-like serine protease, also known as subtilases, with an aspartic acid, histidine, and serine catalytic triad. The catalytic or protease domain of C5P is the segment with sequence homology to the subtilisins within which, lactic acid bacteria cell envelope proteinases (CEPs) all have a few large inserts outside of a structurally conserved core. These inserts are predicted to be surface loops. Whereas the subtilisins in general show broad substrate specificity profiles, the CEPs of lactic acid bacteria demonstrate remarkable substrate specificity. The greater specificity of the CEPs is held attributable to changes in the variable regions within a structurally conserved core and to the presence of the A-domain. The research hypothesis was that the variable region insertions and deletions, and the A domain influence the specificity of streptococcal C5a peptidase. This was addressed by studying the biochemical and structural properties of C5P.

6.1 Cloning, expression and purification of recombinant C5P and derivatives

Specific fragments of Streptococcal C5a Peptidase (C5P) were cloned, expressed and purified. These specific fragments included the full coding region of C5P with the exception of the cell wall spanning domains, C5Pmod (modifier or A domain), and C5Pppd (pro region plus protease domain).

Sequencing confirmed that the C5P fragments had been successfully cloned. The query sequences do not share 100% identity with the *S. pyogenes* C5P sequences found in a BLAST search, which may reflect variation between strains or errors introduced during cloning. For all fragments only one clone was sequenced. Sequencing of a second clone could confirm if variations were due to PCR errors or variation between strains. The fragments were cloned from *S. pyogenes* strain B220 for which no sequence information was available. The significant alignments found in a BLAST search were with the C5a peptidase gene from *S. pyogenes* strain cs24 and *S. agalactiae* (GBS) strains I30 and I25.

The recombinant C5P fragments were cloned and expressed as glutathione S-transferase (GST) fusion proteins in *E. coli* XL1. The fusion proteins were purified from sonicates of transformed *E.* coli XL-1 by affinity absorption to glutathione resin. The site-specific protease, PreScission^{TM}, was used to release the recombinant C5P proteins from the GST tag, which remained bound to the glutathione resin. However, lower molecular weight contaminants co-eluted off the glutathione resin. The contaminating proteins would most likely be either breakdown or premature termination products of the recombinant C5P proteins. SE FPLC failed to separate the C5P proteins from contaminants. The recombinant proteins were further purified by IE FPLC. When analyzed by SDS-PAGE the recombinant proteins appeared to be purified to homogeneity. However, the purified recombinant proteins continued to degrade upon storage.

The problem with lower molecular weight contaminants had previously been encountered with the purification of the protease domain of C5P (summer project 1999). In an attempt to overcome this problem the protease domain was expressed in the *E.coli* BL21 strain, which is a mutant strain wherein a major serine protease, Lon, has been deleted. However, there was no discernable difference in the amount of degradation when expression of the protease domain was carried out in XL1 or BL21 strains. To determine whether degradation was a consequence of autoproteolysis, a serine^{512} to alanine^{512} active site knockout construct of the protease domain of C5P ligated into a pGEX6P3 vector was introduced by transformation into both XL1 and BL21 cells. The contaminants were still present; which suggests they are not products of *E. coli* Lon protease proteolysis, nor the result of autoproteolysis. The contaminants may be breakdown products produced by the activity of an *E. coli* protease that was co-purified with the recombinant proteins.

6.2 Purification of native C5P from the cell wall of S. pyogenes

Attempts were made to purify the full length C5P, minus the cell wall spanning domains, from *Streptococcus pyogenes* strain B220 to be used as a positive control in activity assays. A protein of the approximate molecular weight of C5P was harvested from the cell wall of S. *pyogenes* digested with mutanolysin. This protease

demonstrated proteolytic activity on His-C5a. The protein yield was low; only approximately 2 μg C5P per litre of *S. pyogenes* culture was harvested. The protease was purified on a hydroxyapatite column, however, not to homogeneity. There was not considered to be enough protein extracted to continue with the purification process. Approximately 4 μg C5P per litre of *S. pyogenes* culture of C5P was purified to homogeneity from a mutanolysin digest of GBS by ion-exchange chromatography, following elution from a hydroxyapatite column (Bohnsack *et al*, 1991). Parameters of the extraction process may require improvement, however, it is also possible that the *S. pyogenes* strain B220 may express a comparatively lower level of C5P.

6.3 Cloning of a C5P Ser512 to Ala512 active site mutant

A knock-out mutation of the active site serine of the protease was attempted in two ways. The expressed protein was intended for use as a negative control in activity assays, The first method involved the digestion of $C5P^{S512\text{-}A512}$ clone in a pGEX6P3 vector, which had previously been engineered by Dr Jakki Cooney (IMBS), and ligation into a pGEX6P3 vector containing a C5P insert cut with the same restriction enzymes. The restriction enzymes *Bgl*II and *Sac*I were used as they had restriction sites that flanked the site of the desired mutation by several residues. Whilst both restriction sites were cut within the C5P insert, *Sac*I did not cut within the $C5Ppd^{S512\text{-}A512}$ insert. This could have been due to a PCR error within this restriction site, which could be confirmed by DNA sequencing. The mutation in the *SacI* site could have been corrected by PCR. The second method attempted involved introduction of the mutation by PCR using forward and reverse oligonucleotides encoding the serine to alanine mutation over the Ser 512 codon. This method was unsuccessful, probably due to problems within the PCR reactions due to the size of the amplified region (1kb). Other methods of site-directed mutagenesis which could have been used include oligonucleotide-directed mutagenesis using plasmid (double stranded) DNA. The dsDNA fragment cloned into a plasmid vector is denature and mutagenic and selection oligonucleotides are hybridized to the template strand. The mutant strand is synthesized by primer extension using T4 DNA polymerase as it has no 5'-3' exonuclease activity or strand displacement activity. Nicks at the 5' ends of the primers are sealed by T4 DNA ligase and the heteroduplex is propagated and selected

for by transforming a mutation deficient strain of *E. coli.* QuikChange^{R} is a commercial site-directed mutagenesis kit supplied by Stratagene which requires a miniprep of the DNA fragment of interest cloned into a plasmid vector.

6.4 Cloning, expression, and purification of synthetic human C5a

The substrate of Streptococcal C5a Peptidase is human C5a. C5a is a 74 residue fragment of the protein C5. Synthetic human C5a was purchased for use in activity assays. As a more economical source of substrate, synthetic human C5a was also cloned and expressed as a GST fusion protein. Two protein fragments were released off the Glutathione resin which when combined corresponded to approximately the expected molecular weight of the synthetic human C5a. The synthetic gene encodes a protein sequence (AARISL/GPR) that resembles the recognition site of PreScission^{TM} (LEVLPQ/GPL) in that both have a hydrophobic patch followed by Gly in the P1 site, and Pro in the P2 site. The sizes of the synthetic human C5a products suggest PreScission^{TM} may be recognizing this putative site.

To determine if PreScission^{TM} does cleave C5a at this site commercial synthetic human C5a was incubated with PreScission^{TM} under conditions used in the activity assays. No activity was demonstrated. This result was not conclusive however, as both the substrate and the incubation conditions were different. The GST tagged C5a may not have assumed a correct conformation resulting in the accessibility of this site to PreScission^{TM}.

C5a peptidase cleaves a seven residue fragment from the carboxyl terminus of C5a. Release of this fragment from the 35 kDa GST-C5a fusion protein would have been difficult to detect by SDS-PAGE, therefore the GST-C5a fusion protein would not have been suitable for C5P activity assays. Synthetic human C5a was therefore sub-cloned as a His tagged fusion protein. The small size of the His tag (3.4 kDa) would enable the cleavage of His-C5a by C5P to be analyzed by SDS-PAGE. His tagged C5a was purified from inclusion bodies by a process of denaturation with 6 M guanidinium-HCl to solubilise the protein followed by renaturation with 1 mM oxidized and 1 mM reduced glutathione. His-C5a was further purified by Chelating Sepharose Chromatography.

6.5 In vitro activity assays

In order to determine the effect on substrate specificity of the A domain of C5P, purified recombinant full length C5P and C5P without the A domain (C5Pppd) were compared for specificity of activity using *in vitro* digestion assays. Both commercially purchased synthetic C5a and His tagged C5a were digested by recombinant C5P. A shift down in electrophoretic mobility of the digested products suggested the release of a small peptide fragment. The recombinant pro+protease domains of C5a peptidase demonstrated no activity on commercial synthetic C5a or reecombinant His-C5a. The function of the modifier or A-domain is unknown; however it is speculated to be essential for proteolytic activity and to be involved in regulating proteolytic activity and/or specificity of the protease (Seizen *et al*, 1999). The demonstrated lack of activity of C5Pppd supports the speculation that the A-domain is essential for proteolytic activity.

Aliquots of the His tagged C5a digest mix were purified by RPHPLC and analyzed by Mass Spectrometry. The expected molecular weight of His tagged C5a if the scisscile bond for C5P is between His67 and Lys 68 is 9375.7. The molecular weight of the released fragment analyzed by Mass Spectrometry was 9391 Da, accurate to 1 Da/10,000 Da. This difference could be accounted for by the oxidation of a methionine residue to the sulphoxide. The discrepancy between molecular weights is therefore within the range of accuracy for Mass Spectrometry, and supports the scissile bond to be between His67 and Lys68. Oxidation of methionine residues can be prevented by the addition of an excess of methionine to out compete the available O_2. Methionine sulphoxide can be reduced by treatment with a thiol reagent such as dithiothreitol or N-methylmercaptoacetamide. The more polar nature of the sulfoxide will cause a protein that has been oxidized to elute slightly earlier on reverse-phase HPLC than the reduced form of the protein. Analysis by HPLC could be used to determine if the protein had been oxidized prior to MS analysis.

Specificity of C5P is still controversial and has been studied only by one group, and they have presented two different answers to the question of where the scissile bond

lies. The scissile bond for C5P is reported to be either between His67 and Lys 68 or Lys68 and Asp69 of C5a. Sequencing of the 74 residue C5a after incubation with C5P, extracted by a limited tryptic digest of C5P from *S. pyogenes* strain M49T14 (135kDa), suggested a six-residue peptide was lost from the carboxyl terminal (Wexler et al., 1985). These authors later amended this when they found the scissile bond to be between His67 and Lys68, rather than between Lys68 and Asp69, which would give loss of a seven-residue peptide (Cleary *et al.*, 1992). In this experiment extraction of the C5P was performed with muralytic enzyme phage C lysin, and a synthetic 20-residue peptide of the C-terminal of C5a was used as the substrate. The location of the cleavage site is controversial, as both the method of C5P extraction, and the C5a substrate were different in these two experiments. C5P derived from the tryptic digest ranged in molecular weight from 103,000 to 114,000 *Mr*, whereas native C5P extracted with detergent from the cell wall is 135,000 to 137,000 *Mr* (Wexler and Cleary, 1985). The cell wall spanning and anchoring domains of C5P have 94 and 41 residues, respectively (Siezen, 1999). This suggests trypsin cleaves within the A domain, which is a putative modifier of specificity of C5P. It has been proposed that C5P is a highly specific endopeptidase, and that its function is to eliminate the chemotactic and macrophage-activating anaphylatoxin C5a from the foci of infection (Cleary *et al.*, 1992), although evidence in support of specificity is not very substantial.

Recombinant C5P also demonstrated proteolysis of denatured BSA, which had been reported for native C5P (Cleary et al., 1992). C5P activity was inhibited by ZnAc and ZnCl, the Zn effect is suggested to be caused by oxidation of sulfhydryl groups on C5a (Stafslien and Cleary, 2000). C5P was also inhibited at pH 5 which is expected as this pH is below the pI of the active site Histidine residue. Inhibition was also demonstrated by pAMPSF which is an irreversible inhibitor of serine proteases with lysine or arginine substrate specificity. PMSF, a serine protease inhibitor, did not demonstrate inhibition of C5P at or 2x above its effective concentration of 100 µg/mL. Group B streptococci (GBS) C5a peptidase has been reported to be sensitive to PMSF (Bohnsack *et al*, 1991). Group A streptococci (GAS) C5P sensitivity to PMSF has not been reported. GBS C5P is biochemically distinct from the GAS (Cleary *et al* 1992). GAS C5a peptidase has been reported to be sensitive to calcium

chloride, manganese chloride and zinc acetate, and not sensitive to EDTA (Stafslien and Cleary, 2000). GAS C5P sensitivity to other inhibitors has not been reported.

6.6 Under agaraose polymorphonuclear neutrophil migration

In order to relate specificity studies of C5a peptidase to biological effect the chemotactic potential of the breakdown products of C5a was investigated. This was assessed by looking at neutrophil migration under agarose. Under agarose migration assays involves the migration of polymorphonuclear neutrophils (PMNs) through agarose towards the chemotactic factor C5a. Samples from C5P digest mixtures, undigested recombinant His tagged C5a, and commercially purchased synthetic human C5a were analyzed for chemotactic potential in under-agarose neutrophil migration assays. PMNs were purified from fresh human blood collected from a healthy volunteer. Migration was assessed by staining with May-Grunwald and Giemsa stains and examination by light microscopy at 100 x magnification. The chemotactic potential of His-C5a and commercial C5a was inactivated when cleaved by recombinant C5a peptidase. The chemotactic potential of C5a was not affected after addition of C5P inactivated by ZnAc. The chemotactic potential of C5a was inactivated when cleaved by C5a peptidase, therefore, recombinant C5P demonstrated appropriate biological activity.

6.7 Screening for crystallization conditions of recombinant C5P

C5P is of the pyrolysin family of subtilases (subtilisin like serine proteases), for which no structural information is available. Comparative sequence analysis demonstrated C5P shares sequence homology with the subtilase *Lactococcus lactis* PrtP, (Siezen, 1999; Siezen et al., 1993; Siezen et al., 1991; Siezen and Leunissen, 1997). In contrast to subtilisins, C5P and PrtP are both highly specific proteases. With respect to the conserved core structure of subtilisins and subtilases, both PrtP and C5P have large insertions in variable regions (VRs), particularly the VR13. Siezen *et al* postulated that variations within the VRs modulate substrate specificity of subtilases (Siezen, 1999). Furthermore, the characteristic cleavage pattern of casein by PrtP appears to be modulated by the A domain. C5P is a highly specific endopeptidase,

which cleaves the chemotactic and macrophage-activating anaphylatoxin C5a on the carboxyl side of His67 (Cleary et al., 1992). Homology to PrtP suggests the specificity of C5P may also be modified by the A domain and the VRs (Siezen, 1999). Structural studies commenced with screening for crystallization conditions of the purified recombinant C5P fragments. Resolution of crystollographic data was beyond the scope of this project. However, if the hypothesis holds true, the structure of C5P should reveal potential interactions of VRs and the A domain with substrate binding sites and\or the substrate.

Purified recombinant C5Pmod concentrated to 0.87 mg/mL was screened for crystallization conditions with the Hampton Crystal Screens one and two employing the hanging drop vapour diffusion method following the method described for recombinant C5P. No crystals were formed in these screens, possibly because the protein was not concentrated enough. Although the manufacturer of the Hampton Crystal Screens recommend that the sample concentration is between 5 to 25 mg/mL, it is suggested that the protein be concentrated to 10 mg/mL for screening (Drenth, 1994). Recombinant C5Pppd was not screened for crystallization conditions as the protein was not stable after purifying to homogeneity. Cloning the protease domain with a fragment of the modifier domain may confer some stability on the expressed recombinant protein.

After 3 rounds of screening, crystallization conditions were found for recombinant C5P. Cubic crystals formed in 2.7 M $(NH_4)_2SO_4$ in 0.2 M Hepes pH 7.1. The crystal form of C5P was p-hexagonal (p=primative) form (3-6x axis of symmetry). It was possible that crystallization was occurring with a breakdown product of the recombinant C5P protein as the crystals took several weeks to months to form. Considerable breakdown of the unscreened C5P protein occured over this period (demonstrated by analysis of a sample analyzed by SDS-PAGE. To determine if it was a breakdown product of C5P that was crystallizing, a crystal was dissolved in crystallization buffer and analyzed by SDS-PAGE. The size of the principal constituent/s were compared to a sample of the unscreened C5P protein.

The amino-terminal of the unscreened C5P protein constituent that was of the same molecular weight as the crystallized protein was sequenced. The N-terminal sequence

was TIRDLN. This sequence corresponds to residues 101 to 106 of C5P. Comparison sequence analysis predicted residues 32-101 of C5P to constitute a putative pro-sequence Siezen *et al* 1999.

X-ray diffraction of the C5P crystal was performed by Dr Jakki Cooney and Dr Todd Kagawa using a Rigaku R-AxisIIC detector with a RU-200B rotating anode X-ray source at the institute of molecular biosciences, Massey University. Data were processed using HKL and CCCP4 data processing packages. The data was not of sufficient resolution to be used towards providing a solution of the structure. Although C5P shares homology with the subtilisins, for which structures are available. The conserved core of the subtilisins is interrupted with insertions and deletions in variable regions, therefore Molecular Replacement (MR) may not provide a solution of the structure. Dr Jakki Cooney and Dr Todd Kagawa expressed recombinant C5P in a strain of *E. coli* auxotrophic for methionine and grown on minimal media supplemented with amino acids using selenomethionine in place of methionine. Selenium is an anomalous scatterer and can give phase information that can produce high quality electron density maps by Multiwavelength Anomalous Dispersion (MAD). This approach was beyond the scope of this thesis and may be pursued by others at a later date.

6.8 Pro-region processing

Most subtilases are synthesized as pre-pro-enzymes. The 'pre' refers to an amino-terminal signal peptide, which is cleaved during translocation through the cell membrane. Cleavage of the pro-segment is generally required for activation of the subtilase, however this has not been established for C5P (Siezen, 1999). Removal of the predicted propeptide of C5P has been reported to be unnecessary for activation of the protease (Chen and Cleary, 1990).

The pro-sequence of the closely related subtilases PrtP and PrtB are removed by autocatalytic processing, (Siezen *et al* 1997), which suggests that the pro region of the protease was being processed by one residue, towards the amino terminal, from the predicted processing site. Both the recombinant C5P and C5Pppd breakdown to a

product that may correspond to the loss of the seventy residue pro region (7.7kDa). Cleary *et al* (1992) state that there is no indication that C5P is activated by further processing subsequent to removal of the signal peptide. However, the amino terminal sequence of C5P released from the cell surface by Streptococcal Cysteine Proteinase (SCP) starts at residue 90, (Berge and Bjorck 1995). C5a-mediated granulocyte migration was inhibited by the SCP released C5P, suggesting that amino acids 32-89 are not required for proteolytic activity. This supports the possibility of pro region processing.

The protein sequences from six residues either side of the C5a scissile bond (P6'-P6) and six residues either side of the C5P pro-region processing site suggested by N-terminal sequencing were compared for similarity using the SIM alignment tool for protein available at the ExPASy web site, http://us.expasy.org/. The significance of protein sequence similarity was evaluated to have an optimal similarity score of 0 using the PRSS tool. The isoelectric point, aliphatic, and GRAVY indices, of these sequences were also significantly different. (The PRSS tool and GRAVY indices are also available at the ExPASy web site). These sequences do not resemble each other; therefore, if the pro-region is removed by autoproteolysis, C5P may have a more relaxed substrate recognition than has been alluded to previously.

To determine if processing is necessary for activation of the protease, a mutation could be introduced within the predicted pro-region processing site of C5P in order to inhibit processing. The activity of the expressed pro-region mutant protein could then be compared with the activity of expressed C5P cloned without the pro-region. The pro-region of a protein can function as an intramolecular chaperone necessary for correct folding as it decreases the activation energy in the folding reaction. Without the pro-region a protein may be suspended in the molten globule stage of protein folding and be broken down before assuming its correct conformation (Seizen et al., 1995). It may therefore be necessary to clone C5P with the pro-region and use the protein as a positive control once the pro-region has been processed.

6.9 Immunological studies

It has been suggested that antibody response to C5P may account for the partial non-M type-specific immunity that is observed in adults (O'Connor and Darip, 1991).The C5P of GBS has been suggested as a candidate for a GBS polysaccharide-protein conjugate vaccine (Cheng *et al*, 2001) as it is a highly conserved surface protein. GBS C5P was found to be opsonic and to induce macrophage killing of GBS (Cheng et al 2002). The C5P of GAS is also a potential candidate for a GAS polysaccharide-protein conjugate vaccine. Antiserum raised against GAS C5P extracted from strain CS101 was capable of neutralizing cell-bound C5P activity in six different strains of *S. pyogenes,* suggesting that GAS C5P may have only one antigenic type (O'Connor and Cleary, 1986). Intranasal immunization with a recombinant protein fragment of C5P from M49 streptococci stimulated significant levels of IgA and Ig G antibody production, which were shown to inhibit nasapharyngeal colonization of mice by strains representing the major subdivisions of *S. pyogenes* (Ji et al., 1997). Antibodies raised by this M49 C5P fragment were able to neutralize the C5P activity of M1, M6, M11, as well as M49 streptococci, which supports the suggestion that anti-C5P antibodies lack serotype specificity. Examination by enzyme linked immunosorbent assay (ELISA) of human sera and saliva found measurable concentrations of anti-C5P IgG and IgA, in samples from healthy adults and convalescent patients of Group A streptococcal pharyngitis, whilst young children that were less likely to have been exposed to *S. pyogenes* lacked anti-C5P immunoglobulins. Full-length recombinant C5P could be used to screen convalescent serum for anti-C5P antibodies using the enzyme-linked immunosorbent assay in order to determine C5P immunogenicity in NZ.

6.10 Investigation of C5P polymorphisms

All strains of GBS express C5P, however, some stains exhibit little or no C5a-ase activity, which has been attributed to a genetic polymorphism (Bohnsack et al., 2000). The ubiquitous expression of C5P, irrespective of C5a-ase activity, led to speculation that C5P may have a second important function. Antibody against C5P was found to be opsonic (Cheng et al., 2001) which suggested that the cell surface protease may also function as an adhesin or invasin (Cheng et al., 2002). C5P was identified from a phagemid clone, selected on immobilized fibronectin, from a shotgun phage display

library constructed from GBS chromosomal DNA (Beckmann et al., 2002). GBS C5P was demonstrated to bind Hep2 and A549 human epithelial cells, and both GAS and GBS C5P were demonstrated to bind fibronectin, in enzyme-linked immunosorbent assays. Rabbit anti-C5P blocked invasion of A549 cells by GBS, but did not inhibit adherence, indicating that C5P can function as an invasin, independent of adherence (Cheng et al., 2002). C5P could be extracted from NZ clinical isolates of *S. pyogenes* and assayed for activity. Based on observations of subtilases closely related to C5P, any variation or relaxation found in substrate specificity is expected to correlate with differences in sequence within either VRs or the A domain.

6.11 C5P as a target for chemotherapeutics

Mouse infection models have demonstrated C5P is a virulence factor of *S. pyogenes* (Ji et al., 1997; Ji et al., 1996; O'Connor and Cleary, 1987). C5P is therefore a potential target for chemotherapeutic treatment of streptococcal infections. This study was designed to elucidate the structural and biochemical basis of the substrate specificity of C5P, which will assist the design of potent inhibitors of this powerful virulence factor. The recent resurgence of rheumatic fever and increase in the incidence of invasive *S. pyogenes* infections (Carapetis et al., 1995; Stevens, 1999) considered with penicillin failure in treatment of streptococcal infections (Gillespie, 1998; Stevens, 1999), present a strong argument for the need of new chemotherapeutic agents. An inhibitor of C5P could be included in a cocktail of inhibitors designed against individual virulence factors. Such an approach to chemotherapy does not put the pathogen under selective pressure for resistance. Variants of C5a peptides that bind to C5P but are not cleaved could be used as leads for the design of C5P inhibitors. Such variants could be detected by phage display. Initial approaches to developing HIV-1 protease inhibitors were based on characterizing substrate-protease interactions. Inhibitors consisted of small polypeptides that mimicked normal substrate but with the replacement of a nonhydrolyzable isotere at the cleavage site (West and Fairlie, 1995). Because of limitations of peptidomimetic inhibitors, such as poor absorption and oral bioavailability, and high susceptibility to hydrolysis, inhibitors that were structural mimics of peptides with little peptidic character were designed (Reich *et al*, 1995; Ringhofer *et al*, 1999). This was possible because of the cocrystal structure of HIV-1 protease with a peptidomimetic inhibitor. With the solution of the structure of C5P the

design of an inhibitor could follow the approach of the design of the HIV-1 protease inhibitor.

C5P has been shown to be immunogenic and antibodies raised against C5P neutralize both C5P C5a-ase activity and its function as an invasion, and were found to be opsonic (Wexler *et al*, 1985; Ji *et al*, 1997; Cheng et al, 2001; Cheng et al, 2002). Furthermore, C5P may have only one antigenic type (O'Connor and Cleary, 1986; O'Connor and Darip, 1991). These observations have led to C5P being proposed as a candidate for a vaccine. Antibodies raised against C5P could also be used for chemotherapy.

Bibliography

Akesson, P., Sjoholm, A. G., and Bjorck, L. (1996). Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. Journal of Biological Chemistry *271*, 1081-1088.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database searchprograms, Nucleic Acids Res. *25*, 3389-3402.

Beckmann, C., Waggoner, J. D., Harris, T. O., Tamura, G. S., and Rubins, C. E. (2002). Identification of Novel Adhesins from Group B Streptococci by Use of Phage Display Reveals that C5a Peptidase Mediates Fibronectin Binding. Infection and Immunity *70*, 2869-2876.

Berge, A. (1997) Molecular analysis of *Streptococcus pyogenes* and its interactions with the human host, Lund University, Sweeden.

Berge, A., and Bjork, L., (1995). Streptococcal Cysteine Proteinase releases biologically active fragments of streptococcal surface proteins. The Journal of Biological Chemistry *270*, 9862-9867.

Bergeron, F., Leduc, R., and Day, R. (2000). Subtilase-like pro-protein convertases: from molecular specificity to theraputic applications. Journal of Molecular Endocrinology *24*, 1-22.

Bohnsack, J. F., Takahashi, S., Hammitt, L., Miller, D. V., Aly, A. A., and Adderson, E. E. (2000). Genetic Polymorphisms of Group B Streptococcus *scpB* Alter Functional Activity of a Cell-Associated Peptidase that Inactivates C5a. Infection and Immunity *68*, 5018-5025.

Bohnsack, J. F., Zhou, X., Gustin, J. N., Rubens, C. E., Parker, C. J., and Hill, H. R. (1992). Bacterial evasion of the antibody response: human IgG antibodies neutralize soluble but not bacteria-associated Group B streptococcal C5a-ase. The Journal of Infectious Diseases *165*, 315-321.

Bruinenberg, P. G., Doesburg, P., Alting, A. C., Exterkate, F. A., de Vos, W. M., and Siezen, R. J. (1994). Evidence for a large dispensable segment in the subtilisn-like catalytic domain of the *Lactococcus lactis* cell-envelope proteinase. Protein Engineering *7*, 991-996.

Carapetis, J., Robins-Browne, R., Martin, D., Shelby-James, T., and Hogg, G. (1995). Increasing severity of invasive group A streptococcal disease in Australia: clinical and molecular epidemiologicl features and identification of a new virulent M-nontypeable clone. Clinical Infectious Diseases *21*, 1220-1227.

Chen, C. C., and Cleary, P. P. (1989). Cloning and expression of the streptococcal C5a peptidase gene in *Escherichia coli*: linkage to the type 12 M protein gene. Infection and Immunity *57*, 1740-1745.

Chen, C. C., Cleary, P. P., (1990) Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus pyogenes. Journal of Biological Chemistry, *265(6)*, 3161-7.

Cheng, Q., Carlson, B., Pillai, S., Eby, R., Edwards, L., Olmsted, S. B., and Cleary, P. P. (2001). Antibody Against Surface-Bound C5a Peptidase is Opsonic and Initiates Macrophage Killing of Group B Streptococci. Infection and Immunity *69*, 2302-2308.

Cheng, Q., Stafslien, D., Purushothaman, S. S., and Cleary, P. P. (2002). The Group B Streptococcal C5a Peptidase Is Both a Specific Protease and an Invasin. Infection and Immunity *70*, 2408-2413.

Cleary, P. P., Peterson, J., Chen, C., and Nelson, C. (1991). Virulent human strains of Group G streptococci express a C5a peptidase enzyme similar to that produced by Group A streptococci. Infection and Immunity *59*, 2305-2310.

Cleary, P. P., Prahbu, U., Dale, J. B., Wexler, D. E., and Handley, J. (1992). Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and Immunity *1992*, 5219-5223.

Dodson, G., and Wlodawer, A. (1998). Catalytic triads and their relatives. Trends in Biochemical Science *23*, 347-352.

Drenth, J. (1994) Principles of protein X-ray crystallography. New York: Springer-Verlag. ISBN 0-387-94091-X.

Fernandez, H. N., and Hugli, T. E. (1978). Primary structural analysis of the polypeptide portion of human C5a anaphlatoxin. Journal of Biological Chemistry *253*, 6955-6964.

Fernandez-Espla, M. D., Garault, P., Monnet, V., and Rul, F. (2000). *Streptococcus thermophilus* Cell Wall-Anchored Proteinase: Release, Purification, and Biochemical and Genetic Characterization. Applied and Environmental Microbiology *66*, 4774-4778.

Gillespie, S. H. (1998). Failure of penicillin in *Streptococcus pyogenes* pharyngeal infection [comment]. Lancet *352*, 1954-1956.

Granlund, M., Oberg, L., and Sellin, M. (1997). Identification of a novel insertion element, IS*1548*, in Group B streptococci, predominantly in strains causing endocarditis. The Journal of Infectious Diseases *177*, 967-976.

Hill, H. R., Bohnsack, J. F., Morris, E. Z., Augustine, N. H., Parker, C. J., Cleary, P. P., and Wu, J. T. (1988). Group B streptococci inhibit the chemotactic activity of the fifth component of complement. Journal of immunology *141*, 3551-3556.

Hollingshead, S. K., Arnold, J., Readdy, T. L., Yung, D. L., and Bessen, D. E. (1993). Structural heterogeneity of the *emm* gene cluster in group A streptococci. Molecular microbiology *8*, 707-717.

Hortsmann, R. D., Sievertsen, H. J., Leippe, M., and Fischetti, V. A. (1992). Role of fibrinogen in complement inhibition by streptococcal M protein. Infection and Immunity *60*, 5036-5041.

Ji, Y., Carlson, B., Kondagunta, A., and Cleary, P. P. (1997). Intranasal immunization with C5a peptidase prevents nasopharyngeal colonization of mice by the Group A streptococcus. Infection and Immunity *65*, 2080-2087.

Ji, Y., McLandsborough, L., Kondugunta, A., and Cleary, P. P. (1996). C5a peptidase alters clearance and trafficking of Group A streptococci by infected mice. Infection and Immunity *64*, 503-510.

Ji, Y., Schnitzler, N., DeMaster, E., and Cleary, P. (1998). Impact of M49, Mrp, Enn, and C5a Peptidase proteins on colonization of the mouse oral mucosa by *Streptococcus pyogenes*. Infection and immunity *66*, 5399-5405.

Kossiakoff, A.A., Spencer, S. A. (1981). Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: Neutron structure of trypsin. Biochemistry *20*, 6462-6474.

Madigan, M. T., Martinko, J. M., and Parker, J. (1997). Brock Biology of.Microorganisms, Eighth edn, Prentice Hall International).

Mandecki, W., Mollison, K. W., Bolling, T. J., Powell, B. S., Carter, G. W., and Fox, J. L. (1985). Chemical synthesis of a gene encoding the human complement fragment C5a and its expression in *Escherichia coli*. Proceedings of the National Academy of Science in the USA *82*, 3543-3547

McIver, K., Heath, A. S., Green, B. D., and Scott, J. R. (1995). Specific binding of the activator Mga to promoter sequences of the *emm* and *scpA* genes in the Group A streptococcus. Journal of Bacteriology *177*, 6619-6624.

McIver, K., and Scott, J. R. (1997). Role of *mga* in growth phase regulation of virulence genes of the Group A Streptococcus. Journal of Bacteriology *179*, 5178-5187.

Mims, C., Wakelin, D., Playfair, J., Williams, R., and Roitt, I. (1998). Medical Microbiology, Second edn, Mosby International Limited).

O'Connor, A. P., and Cleary, P. P. (1986a). Localization of the streptococcal C5a peptidase to the surface of Group A streptococci. Infection and Immunity *53*, 432-434.

O'Connor, S. P., and Cleary, P. (1986b). Localization of the streptococcal C5a peptidase to the surface of group A streptococci. Infection and Immunity *53*, 432-443.

O'Connor, S. P., and Cleary, P. P. (1987). In vivo Streptococcus pyogenes C5a peptidase activity: analysis using transposon- and nitrosguanidine-induced mutants. The Journal of Infectious Diseases *156*, 495-504.

O'Connor, S. P., and Darip, D. (1991). The human antibody response to streptococcal C5a peptidase. The Journal of Infectious Diseases *163*, 109-116.

Perona, J. J., and Craik, C. S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Science *4*, 337-360.

Philips, G. N., Flicker, P. F., Cohen, C., Manjula, B. N., and Fischetti, V. A. (1981). Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. Proceedings of the National Academy of Science in the USA *78*, 4689-4693.

Roitt, I. (1997). Roitt's Essential Immunology, Ninth Edition edn, Blackwell Science).

Ross, M., Peters, L., Double staining to increase the sensitivity of protein detection in acrylamide gels. Biotechniques 9 *5*, 532-533.

Reich, S.H., Melnick, M., Davies, J.F., Appelt, K., Lewis, K.K., Fuhry, M.A., Pino, M.P., Trippe, A.J., Nguyen, D., Dawson, H., Wu, B., Musick, L., Kosa, M., Dahil, D., Webber, S., Gehlhaar, D.K., Andrada, D., Shetty, B. (1995) Protein structure-based design of potent orally bioavailable, nonpeptide inhibitors of human immunodeficiency virus protease. Proceedings of the National Academy of Science in the USA, *8*, 3298-3302.

Ringhofer, S., Kallen, J., Dutzler, R., Billich, A., Visser, A. J.W.G., Scholz, D., Steinhauser, O., Schreiber, H., Auer, M., Kungl, A.J. (1999) X-ray structure and conformational dynamics of the HIV-1 protease in complex with the inhibitor SDZ283-910. Agreement of time-resolved spectroscopy and molecular dynamics simulations, 1147-1159.

Siezen, R. J. (1999). Multi-domain, cell-envelope proteinases of lactic acid bacteria. Antonie von Leeuwenhoek *76*, 139-155.

Siezen, R. J., Bruinenberg, P. G., Vos, P., van Alen-Boerrigter, I. J., Nijhuis, M., Alting, A. C., Exterkate, F. A., and de Vos, W. M. (1993). Engineering of the subtilase binding region of the subtilisin-like cell-envelope proteinase of *Lactococcus lactis*. Protein Engineering *6*, 927-937.

Siezen, R. J., de Vos, W. M., Leunissen, J. A. M., and Dijkstra, B. W. (1991). Homology modelling and protein engineering strategy of subtilases, the family ot subtilisin-like serine proteases. Protein Engineering *4*, 719-737.

Siezen, R. J., and Leunissen, J. A. M. (1997). Subtilases: The superfamily of subtilisin-like serine proteases. Protein Science *6*, 501-523. Seizen, R. J., Leunissen, J. A. M., and Schinde, U., eds. (1995). Homology analysis of the propeptides of subtilisin-like serine proteases (subtilases. (R. G. Landes Company).

Simpson, W. J., La Penta, D., Chen, C., and Cleary, P. P. (1990). Coregulation of type 12 M protein and streptococcal C5a peptidase genes in group A streptococci: evidence for a virulence regulon controlled by the *virR* locus. Journal of Bacteriology *172*, 696-700.

Stafslien, D., and Cleary, P. P. (2000). Characterization of the Streptococcal C5a Peptidase Using a C5a-Green Fluorescent Protein Fusion Protein Substrate. Journal of Bacteriology *182*, 3254-3258.

Steitz, T. A., and Shulman, R. G. (1982) Crystallographic and NMR studies of the serine proteases. Annual Review of Biochemistry and Biophysiology. *11*, 419-444

Stevens, D. L. (1999). The flesh-eating bacterium: what's next? Journal of Infectious Diseases *179 Suppl 2*, S366-374.

Stollerman, G. H. (1998a). The changing face of rheumatic fever in the 20th century. Journal of Medical Microbiology *47*, 655-657.

Stollerman, G. H. (1998b). The changing face of rheumatic fever in the 20th century [editiorial]. Journal of Medical Microbiology *47*, 655-657.

Stryer, L. (1995). Biochemistry, Fourth edn (New York, W. H. Freeman and Company).

Vos, P., Boerrigter, I. J., Buist, G., Haandrikman, A. J., Niijuis, M., de Reuver, M. B., Siezen, R. J., Venema, G., de Vos, W. M., and Kok, J. (1991). Engineering of the *Lactococcus lactis*serine proteinase by construction of hybrid enzymes. Protein Engineering *4*, 479-484.

Toth, M. J., Huwyler, L., Boyar, W. C., Braunwalder, A. F., Yarwood, D., Hadala, J., Haston, W. O., Sills, M. A., Seligmann, B., and Galakatos, N. (1994). The pharmacophore of the human C5a anaphylatoxin. Protein Science *3*, 1159-1168.

Wessels, M. R., Moses, A. E., Goldberg, J. B., and DiCesare, T. J. (1991). Hyaluronic acid capsule is a virulence factor for mucoid group A streptococci. Proceedings of the National Academy of Sciences in the USA *88*, 8317-8321.

West, M.L., Fairlie, D.P. (1995) Targeting HIV-1 protease: a test of drug-design methodologies. Trends Pharmocological Science, *16(2)*, 67-75.

Wexler, D. E., Chenoweth, D. E., and Cleary, P. P. (1985). Mechanism of action of the Group A streptococcal C5a inactivator. Proceedings of the National Academy of Science in the USA *82*, 8144-8148.

Wexler, D. E., and Cleary, P. P. (1985). Purification and characteristics of the streptococcal chemotactic factor inactivator. Infection and Immunity *50*, 757-764.

Wexler, D. E., Nelson, R. D., and Cleary, P. P. (1983). Human neutrophil chemotactic response to Group A streptococci: bacteria-mediated interference with complement-derived chemotactic factors. Infection and Immunity *39*, 239-246.

Whitnack, E., and Beachey, E. H. (1995). Biochemical and biological properties of the binding of human fibrinogen to M protein in group A streptococci. Journal of bacteriology *164*, 350-358.

Zhang, X., Boyer, W., Toth, M. J., Wennogle, L., and Gonnella, C. (1997) Structural definition of the C5a C terminus by two-dimensional nuclear magnetic resonance spectroscopy. Proteins: structure, function, and genetics *28*, 261-267.

Appendix A Sequences of oligonucleotides used in C5P cloning

Oligo name	5' –Sequence-3'	Purpose
C5P 421F	GCGAATTGTCTGTTTTGGTTTCGT	sequencing of C5P
C5P 1020R	AGTTTCAGTGAGCTGTTTATCTGG	sequencing of C5P
C5P 1324R	CTGCCAGAGAATAATTTTCTGTTAGG	sequencing of C5P
C5P 1800R	GTTCAGGTGAACCTTGCTTGAGGT	sequencing of C5P
C5P 2224F	CTGGTTAATCTACCACTACCTAAT	sequencing of C5P
C5P 2640R	TGAACCAAGTGTGCTTGCCAAGTC	sequencing of C5P
C5P For	CAGCAGGGATCCAATACTGTGACAGAAGACACTCCT	cloning of C5P
C5P Rev	CAGCAGGAATTCTTAAGAGTGGCCCTCCAATAGCTT	cloning of C5P
C5Pppd For	CAGCAGGGATCCAATACTGTGACAGAAGACACTCCT	cloning of C5Pppd
C5Pppd Rev	CAGCAGGAATTCTTATGCTGCTGAAGCTTTTTTAGC	cloning of C5Pppd
C5Pmod For	CAGCAGGGATCCACGATGTATGTGACAGATAAGGAT	cloning of C5Pmod
C5Pmod Rev	CAGCAGGAATTCTTAAGGTGACAGAAGACACTCCT	cloning of C5Pmod

Appendix B Sequencing

Cloned synthetic human C5a cDNA sequence

```
>gi|208517|gb|K03392.1|SYNHUMC5A  Synthetic human complement fragment C5a
gene, complete cds
          Length = 253

 Score =  452 bits (228), Expect = e-124
 Identities = 228/228 (100%)
 Strand = Plus / Plus


Cloned      1 atgaccctgcagaagaaaatcgaagaaatcgctgctaagtacaaacactctgttgttaaa 60
C5a           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Synthetic  21 atgaccctgcagaagaaaatcgaagaaatcgctgctaagtacaaacactctgttgttaaa 80
Human C5a

Cloned     61 aaatgctgctacgacggtgcttgcgttaacaacgacgaaacttgcgaacagcgtgctgct 120
C5a           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Synthetic  81 aaatgctgctacgacggtgcttgcgttaacaacgacgaaacttgcgaacagcgtgctgct 140
Human C5a


Cloned    121 cgtatctctctgggcccgcgttgcatcaaagcattcactgaatgctgcgttgttgcttct 180
C5a           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Synthetic 141 cgtatctctctgggcccgcgttgcatcaaagcattcactgaatgctgcgttgttgcttct 200
Human C5a


Cloned    181 cagctgcgtgctaacatctctcacaaagacatgcaactgggtcgttaa 226
C5a           ||||||||||||||||||||||||||||||||||||||||||||||||
Synthetic 201 cagctgcgtgctaacatctctcacaaagacatgcaactgggtcgttaa 248
Human C5a
```

A BLASTN search (Altschul, 1997) using default parameters (BLOSOM 60, PAM 30) revealed cloned C5a cDNA shares 100% identity with the complete coding region for synthetic human C5a.

C5Pppd Sequence Alignment

```
 Score =  959 bits (2479), Expect = 0.0
 Identities = 525/540 (97%), Positives = 528/540 (97%), Gaps = 1/540 (0%)

Query: 1   NTVTEDTPATEQAVETPQPTAVSEEAPSSSKETKIPQTPGDAEETVADDANDLAPQAPAK 60
           NTVTEDTP TEQAVETPQPTAVSEE P SSKETK PQTP DAEET+ADDANDLAPQAPAK
Sbjct: 32  NTVTEDTPVTEQAVETPQPTAVSEEVP-SSKETKTPQTPDDAEETIADDANDLAPQAPAK 90

Query: 61  TADTPATSKATIXDLNDPSQVKTLQEKASKGAGTVVAVIDAGFDKNHEAWRLTDKTKARY 120
           TADTPATSKATI DLNDPSQVKTLQEKA KGAGTVVAVIDAGFDKNHEAWRLTDKTKARY
Sbjct: 91  TADTPATSKATIRDLNDPSQVKTLQEKAGKGAGTVVAVIDAGFDKNHEAWRLTDKTKARY 150

Query: 121 QSKEDLEKAKKEHGITYGEWVNDKVAYYHDYSKDGKTAVDQEHGTHVSGILSGNAPSETK 180
           QSKEDLEKAKKEHGITYGEWVNDKVAYYHDYSKDGKTAVDQEHGTHVSGILSGNAPSETK
Sbjct: 151 QSKEDLEKAKKEHGITYGEWVNDKVAYYHDYSKDGKTAVDQEHGTHVSGILSGNAPSETK 210

Query: 181 EPYRLEGAMPEAQLLLMRVEIVNGLADYARNYAQAIRDAVNLGAKVINMSFGNAALAYAN 240
           EPYRLEGAMPEAQLLLMRVEIVNGLADYARNYAQAIRDAVNLGAKVINMSFGNAALAYAN
Sbjct: 211 EPYRLEGAMPEAQLLLMRVEIVNGLADYARNYAQAIRDAVNLGAKVINMSFGNAALAYAN 270

Query: 241 LPDETKKAFDYAKSKGVSIVTSAGNDSSFGGKTRLPLADHPDYGVVGTPAAADSTLTVAS 300
           LPDETKKAFDYAKSKGVSIVTSAGNDSSFGGKTRLPLADHPDYGVVGTPAAADSTLTVAS
Sbjct: 271 LPDETKKAFDYAKSKGVSIVTSAGNDSSFGGKTRLPLADHPDYGVVGTPAAADSTLTVAS 330

Query: 301 YSPDKQLTETAMVKTDDHQAKEMPVLSTNRFEPTKAYDYAYANRGMKEDDFKDVKGKIAL 360
           YSPDKQLTETAMVKTDD Q KEMPVLSTNRFEP KAYDYAYANRGMKEDDFKDVKGKIAL
Sbjct: 331 YSPDKQLTETAMVKTDDQQDKEMPVLSTNRFEPNKAYDYAYANRGMKEDDFKDVKGKIAL 390

Query: 361 IERGDIDFKDKIANAKKAGAVGVLIYDNQDKGFPIELPNVDQMPAAFISRKDGLLLKDNS 420
           IERGDIDFKDK+ANAKKAGAVGVLIYDNQDKGFPIELPNVDQMPAAFISRKDGLLLKDN
Sbjct: 391 IERGDIDFKDKVANAKKAGAVGVLIYDNQDKGFPIELPNVDQMPAAFISRKDGLLLKDNP 450

Query: 421 KKTITFNATPKVLPTASDTKLSRFSSWGLTADGNIKPDIAAPGQDILSSVANNKYAKLSG 480
           +KTITFNATPKVLPTAS TKLSRFSSWGLTADGNIKPDIAAPGQDILSSVANNKYAKLSG
Sbjct: 451 QKTITFNATPKVLPTASGTKLSRFSSWGLTADGNIKPDIAAPGQDILSSVANNKYAKLSG 510

Query: 481 TSMSAPLVAGIMGLLQKQYETQYPDMTPSERLDLAKKVLMSSATALYDEDEKAYFSPRQQ 540
           TSMSAPLVAGIMGLLQKQYETQYPDMTPSERLDLAKKVLMSSATALYDEDEKAYFSPRQQ
Sbjct: 511 TSMSAPLVAGIMGLLQKQYETQYPDMTPSERLDLAKKVLMSSATALYDEDEKAYFSPRQQ 570

Query: 541 GAGAVDAKKASAA  554
           GAGAVDAKKASAA
Sbjct: 571 GAGAVDAKKASAA  584
```

Results of a BLAST blastp translated cloned C5Pppd sequence alignment using default parameters (BLOSOM 60, PAM 30). The alignment shown is cloned C5Pppd (query) with a *S. pyogenes* C5Psequence accession number P15926 (Chen. and Cleary, 1990)

Alignment of cloned C5P and *S. pyogenes* J05229.1 C5P

```
Alignment Parameters

blosum62 matrix

        Gap Weight:     12      Average Match:  2.912
     Length Weight:      4   Average Mismatch: -2.003

           Quality:   4846             Length:   1002
             Ratio:  4.945              Gaps:        2
Percent Similarity: 98.059   Percent Identity: 97.753

       Match display thresholds for the alignment(s):
                   | = IDENTITY
                   : =    2 (conservative substitution)
                   . =    1 (partially conservative substitution)


                .         .         .         .         .
 32 ntvtedtpvteqavetpqptavseevp.ssketktpqtpddaeetiadda 80
    |||||||| |||||||||||||||| | |||||| |||| |||||:||||
  1 NTVTEDTPATEQAVETPQPTAVSEEAPSSSKETKIPQTPGDAEETVADDA 50
          .         .         .         .         .
 81 ndlapqapaktadtpatskatirdlndpsqvktlqekagkgagtvvavid 130
    |||||||||||||||||||| ||||||||||||||| ||||||||||
 51 NDLAPQAPAKTADTPATSKATIXDLNDPSQVKTLQEKASKGAGTVVAVID 100
          .         .         .         .         .
131 agfdknheawrltdktkaryqskedlekakkehgityqewvndkvayyhd 180
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 AGFDKNHEAWRLTDKTKARYQSKEDLEKAKKEHGITYGEWVNDKVAYYHD 150
          .         .         .         .         .
181 yskdgktavdqehgthvsgilsgnapsetkepyrlegampeaqlllmrve 230
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 YSKDGKTAVDQEHGTHVSGILSGNAPSETKEPYRLEGAMPEAQLLLMRVE 200
          .         .         .         .         .
231 ivngladyarnyaqairdavnlgakvinmsfgnaalayanlpdetkkafd 280
    ||||||||||||||||||||||||||||||||||||||||||||||||||
201 IVNGLADYARNYAQAIRDAVNLGAKVINMSFGNAALAYANLPDETKKAFD 250
          .         .         .         .         .
281 yakskgvsivtsagndssfggktrlpladhpdygvvgtpaaadstltvas 330
    |||||||||||||||||||||||||||||||||||||||
251 YAKSKGVSIVTSAGNDSSFGGKTRLPLADHPDYGVVGTPAAADS...... 294
          .         .         .         .         .
331 yspdkqltetamvktddqqdkempvlstnrfepnkaydyayanrgmkedd 380
                  | | |||||||||||| ||||||||||||||||
295 ..............DHQAKEMPVLSTNRFEPTKAYDYAYANRGMKEDD 328
          .         .         .         .         .
381 fkdvkgkialiergdidfkdkvanakkagavgvliydnqdkgfpielpnv 430
    ||||||||||||||||||:||||||||||||||||||||||||||||||
329 FKDVKGKIALIERGDIDFKDKIANAKKAGAVGVLIYDNQDKGFPIELPNV 378
          .         .         .         .         .



431 dqmpaafisrkdglllkdnpqktitfnatpkvlptasgtklsrfsswglt 480
    |||||||||||||||||| .|||||||||||||||| ||||||||||||
379 DQMPAAFISRKDGLLLKDNSKKTITFNATPKVLPTASDTKLSRFSSWGLT 428
          .         .         .         .         .
481 adgnikpdiaapgqdilssvannkyaklsgtsmsaplvagimgllqkqye 530
    ||||||||||||||||||||||||||||||||||||||||||||||||||
429 ADGNIKPDIAAPGQDILSSVANNKYAKLSGTSMSAPLVAGIMGLLQKQYE 478
```

Alignment was produced using the unix-1 alignment tool bestfit. The *S. pyogenes* J05229.1 C5P sequence (Chen and Cleary (1990) was retrieved from the GenBank sequence database. There was a gap in the sequence between residues 324-346