



# LUND UNIVERSITY

## Streptococcal M protein and human C4BP

Persson, Jenny J

2006

[Link to publication](#)

*Citation for published version (APA):*

Persson, J. J. (2006). *Streptococcal M protein and human C4BP*. [Doctoral Thesis (compilation), Department of Laboratory Medicine]. Department of Laboratory Medicine, Lund University.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Streptococcal M protein and human C4BP

Jenny Persson

Division of Medical Microbiology  
Department of Laboratory Medicine  
Lund University, Sweden

LUND  
2006

## **Doktorsexamen i Medicisk Mikrobiologi**

som med vederbörligt tillstånd från Medicinska Fakulteten vid Lunds Universitet för avläggande doktorsexamen i medicinsk vetenskap kommer att offentligen försvaras i Segerfalksalen, Wallenbergs Neurocentrum, Sölvegatan 17, Lund, lördagen den 29 april 2006, kl. 09.30

## **Fakultetsopponent**

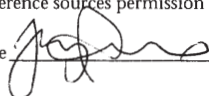
Professor Brian G. Spratt, Imperial College London, London, Storbritannien

Organization LUND UNIVERSITY  Division of Medical Microbiology Department of Laboratory Medicine Sölvegatan 23 223 62 Lund	Document name DOCTORAL DISSERTATION	
	Date of issue April 29, 2006	
	Sponsoring organization	
Author(s) Jenny Persson		
Title and subtitle Streptococcal M protein and human C4BP		
Abstract Antigenic variation of surface proteins allows microorganisms to evade the immune system of the infected host. This phenomenon represents an apparent paradox, because the variable protein must retain an important function, while its antigenic properties vary extensively. The surface associated M protein of <i>Streptococcus pyogenes</i> , a common human pathogen, exhibits antigenic variation due to an N-terminal hypervariable region (HVR). The HVRs of many M proteins bind the human complement regulator C4b-binding protein (C4BP), which down-regulates deposition of complement on the bacterial surface and thereby protects the bacteria against complement-mediated phagocytosis. Different immunological, biochemical and structural aspects of this biologically important interaction is the focus of the four papers included in this thesis. C4BP-binding HVRs exhibit remarkable sequence divergence, yet bind the same ligand. In the first study, we found that such HVRs can be studied in isolated form, as synthetic peptides, thus allowing us to directly characterize the HVRs and their interaction with C4BP. Our data indicate that the peptides bind to the same region in C4BP and assume similar folds, although they are antigenically unrelated. In the second study, we show that such a synthetic peptide can be used to purify human C4BP by a simple one-step affinity-chromatography method. The third study was focused on the sequence divergence among C4BP-binding HVRs. Remarkably, analysis of seven HVRs demonstrated that they completely lack residue identities. However, use of site-specific mutagenesis to substitute relatively conserved residues in the M22 protein indicated that the predicted coiled-coil structure of the HVR is crucial for ability to bind C4BP. Interestingly, change of single residues that do not affect C4BP-binding induced major immunological changes. Together, the data in paper III indicate that HVRs of C4BP-binding M proteins have an extraordinary capacity for sequence change and antigenic variability, while retaining a specific ligand-binding function. In paper IV, we studied the three-dimensional structure of C4BP-binding HVRs, using peptides derived from the M4 and M22 proteins. No structure could be obtained, but the data clearly indicate that the central parts of the HVRs are folded as coiled-coils, both in solution and in complex with C4BP, while the termini are flexible. Remarkably, the peptides derived from M4 and M22 appear to adopt similar structures, in spite of a limited number of residue identities.		
Key words: <i>Streptococcus pyogenes</i> , innate immunity, sequence variability, antigenic variation, M protein, human C4BP		
Classification system and/or index terms (if any):		
Supplementary bibliographical information:		Language English
ISSN and key title: 1652-8220		ISBN 91-85481-78-5
Recipient's notes	Number of pages 135	Price
	Security classification	

Distribution by (name and address)

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date March 21, 2006

# Streptococcal M protein and human C4BP

Jenny Persson

Division of Medical Microbiology  
Department of Laboratory Medicine  
Lund University, Sweden



*Academic dissertation*

LUND  
2006

## CONTENTS

<b>CONTENTS</b>	4
<b>LIST OF PAPERS</b>	6
<b>ABBREVIATIONS</b>	7
<b>SUMMARY</b>	8
<b>INTRODUCTION</b>	9
<b>THE COMPLEMENT SYSTEM</b>	10
Activation of complement	10
<i>Classical pathway</i>	10
<i>Lectin pathway</i>	12
<i>Alternative pathway</i>	13
Functions of complement	14
<i>Opsonization and phagocytosis</i>	14
<i>Membrane attack complex</i>	14
<i>Anaphylatoxins</i>	15
<i>Additional roles for complement</i>	15
Regulation of complement activation	16
<i>Regulators in plasma</i>	16
<i>C4BP</i>	18
<i>Cell-bound regulators</i>	20
 <b>STREPTOCOCCUS PYOGENES</b>	21
Infections and sequelae	22
Classification	24
Genomics	25
Virulence factors	25
The M protein family	27
<i>The Mga regulon and the emm locus</i>	28
<i>Mrp, M and Enn proteins</i>	30
<i>The coiled-coil structure and M protein</i>	31
<i>Binding of human plasma proteins to M proteins</i>	33
<i>Binding of human plasma proteins to Mrp and Enn proteins</i>	36
<i>M protein and human C4BP</i>	36
 <b>VARIABILITY IN MICROBIAL SURFACE PROTEINS</b>	39
Antigenic variation	39
The hypervariable region in streptococcal M proteins	41
 <b>BIOPHYSICAL METHODOLOGY – IN BRIEF</b>	42
Circular Dichroism	42
Nuclear Magnetic Resonance Spectroscopy	44

<b>PRESENT INVESTIGATION</b>	45
<b>PAPER I:</b> Isolated hypervariable regions derived from streptococcal M proteins specifically bind human C4b-binding protein: implications for antigenic variation	45
<b>PAPER II:</b> Single-step purification of human C4b-binding protein (C4BP) by affinity chromatography on a peptide derived from a streptococcal surface protein	46
<b>PAPER III:</b> Extreme sequence divergence but conserved ligand-binding specificity in <i>Streptococcus pyogenes</i> M protein	47
<b>PAPER IV:</b> Streptococcal M protein: structural studies of the hypervariable region, free and bound to human C4BP	50
<b>CONCLUSIONS</b>	53
<b>SAMMANFATTNING PÅ SVENSKA</b>	54
<b>ACKNOWLEDGMENTS</b>	56
<b>REFERENCES</b>	58
<b>APPENDICES: papers I-IV</b>	77

## LIST OF PAPERS

This thesis is based on the following papers, which are referred to in the text as **papers I-IV**.

- I. Morfeldt E., Berggård K., Persson J., Drakenberg T., Johnsson E., Lindahl E., Linse S. and Lindahl G. (2001) Isolated hypervariable regions derived from streptococcal M proteins specifically bind human C4b-binding protein: implications for antigenic variation. *J Immunol* 167: 3870-3877<sup>1</sup>
- II. Persson J. and Lindahl G. (2005) Single-step purification of human C4b-binding protein (C4BP) by affinity chromatography on a peptide derived from a streptococcal surface protein. *J Immunol Methods* 297: 83-95<sup>2</sup>
- III. Persson J., Beall B., Linse S., and Lindahl G. (2006) Extreme sequence divergence but conserved ligand-binding specificity in *Streptococcus pyogenes* M protein. *PLoS Pathogens*, under revision.
- IV. André I., Persson J., Blom A. M., Nilsson H., Drakenberg T., Lindahl G. and Linse S. (2006) Streptococcal M protein: structural studies of the hypervariable region, free and bound to human C4BP. *Biochemistry*. In press.<sup>3</sup>

Printed papers are reproduced with permission from <sup>1</sup>Copyright 2001. The American Association of Immunologists, Inc., <sup>2</sup>Copyright 2005, Elsevier Press B.V. and <sup>3</sup>Biochemistry, in press, unpublished work copyright 2006, American Chemical Society.

## **ABBREVIATIONS**

aa	amino acid
ARF	Acute rheumatic fever
C1-inh	C1-inhibitor
C4BP	C4b-binding protein
CCP	Complement control protein
CD(followed by number)	Cluster of differentiation
CD	Circular dichroism
CR	Complement receptor
DAF	Decay accelerating factor
FH	Factor H
FHL-1	Factor H-like protein 1
fi	Factor I
HVR	Hypervariable region
Ig	Immunoglobulin
kDa	kiloDalton
MAC	Membrane attack complex
MASP	MBL-associated serine protease
MBL	Mannose-binding lectin
MCP	Membrane cofactor protein
Mga	Multigene regulator of group A streptococcus
NMR	Nuclear magnetic resonance
OF	Opacity factor
PSGN	Post-streptococcal glomerulonephritis
RCA	Regulators of complement activation
SOF	Serum opacity factor
SVR	Semivariable region



## SUMMARY

Antigenic variation of surface proteins allows microorganisms to evade the immune system of the infected host. This phenomenon represents an apparent paradox, because the variable protein must retain an important function, while its antigenic properties vary extensively. The surface associated M protein of *Streptococcus pyogenes*, a common human pathogen, exhibits antigenic variation due to an N-terminal hypervariable region (HVR). The HVRs of many M proteins bind the human complement regulator C4b-binding protein (C4BP), which down-regulates deposition of complement on the bacterial surface and thereby protects the bacteria against complement-mediated phagocytosis. Different immunological, biochemical and structural aspects of this biologically important interaction is the focus of the four papers included in this thesis.

C4BP-binding HVRs exhibit remarkable sequence divergence, yet bind the same ligand. In the first study, we found that such HVRs can be studied in isolated form, as synthetic peptides, thus allowing us to directly characterize the HVRs and their interaction with C4BP. Our data indicate that the peptides bind to the same region in C4BP and assume similar folds, although they are antigenically unrelated. In the second study, we show that such a synthetic peptide can be used to purify human C4BP by a simple one-step affinity-chromatography method.

The third study was focused on the sequence divergence among C4BP-binding HVRs. Remarkably, analysis of seven HVRs demonstrated that they completely lack residue identities. However, use of site-specific mutagenesis to substitute relatively conserved residues in the M22 protein indicated that the predicted coiled-coil structure of the HVR is crucial for ability to bind C4BP. Interestingly, change of single residues that do not affect C4BP-binding induced major immunological changes. Together, the data in paper III indicate that HVRs of C4BP-binding M proteins have an extraordinary capacity for sequence change and antigenic variability, while retaining a specific ligand-binding function.

In paper IV, we studied the three-dimensional structure of C4BP-binding HVRs, using peptides derived from the M4 and M22 proteins. No structure could be obtained, but the data clearly indicate that the central parts of the HVRs are folded as coiled-coils, both in solution and in complex with C4BP, while the termini are flexible. Remarkably, the peptides derived from M4 and M22 appear to adopt similar structures, in spite of a limited number of residue identities.

## INTRODUCTION

Throughout history, infectious diseases have caused a number of infamous pandemics, each in its own time killing vast numbers of people. In the 14<sup>th</sup> century, the Black Death killed 30-40% of the European population, and smallpox and measles practically eradicated the Aztec and Inca populations during the 16<sup>th</sup> century. More people were killed by the post-World War I influenza pandemic, the Spanish flu, than by the war itself, and in modern times millions are dying from AIDS each year (Morens *et al.*, 2004; Harrison, 2006). With the awareness that arose, in the second part of the 19<sup>th</sup> century, of the existence of contagious matter (the germ theory) and the introduction and soon widespread use of penicillin in the early 20<sup>th</sup> century, the common expectancy was that the era of infectious diseases would soon come to an end. In the 1960's, some American scientists therefore declared that "the war against the microbes has been won". Today, with newly emerging and re-emerging diseases, and the development of resistance against antibiotics and other substances in many microbes, it is clear that this battle is far from being won. At present, >25% of yearly deaths worldwide can be ascribed to infections, and twice as many people are killed by viruses, parasites and bacteria as by neoplastic diseases (Morens *et al.*, 2004). Therefore, the necessity of increased knowledge about microbes and their interplay with the human host is more evident than ever.

In the combat against invading pathogenic microorganisms, the human being has two defense systems, innate and acquired immunity. The acquired, or adaptive, immune system evolves throughout life, and adapts to every new microbial encounter. It is composed of B cells and T cells and is responsible for production of specific antibody responses and development of immunological memory. The innate immune system is unspecific in its nature and does not adapt as a result of infection. This system consists of physical barriers such as skin and mucosal membranes, phagocytic cells such as neutrophils and macrophages, and a vast number of bactericidal molecules present in blood, mucosa, tears and saliva. An important part of the innate immune system is the complement system, which plays a key role in the defense against microbes.

This thesis is focused on the interaction between the surface exposed M protein of the bacterium *Streptococcus pyogenes*, which causes disease in humans, and the human plasma protein C4b-binding protein (C4BP), which is involved in regulation of the complement system. The binding of C4BP to the surface of *S. pyogenes* protects the bacterium against complement-mediated killing and thereby aids the bacterium in the fight for survival and multiplication in the human host.

## THE COMPLEMENT SYSTEM

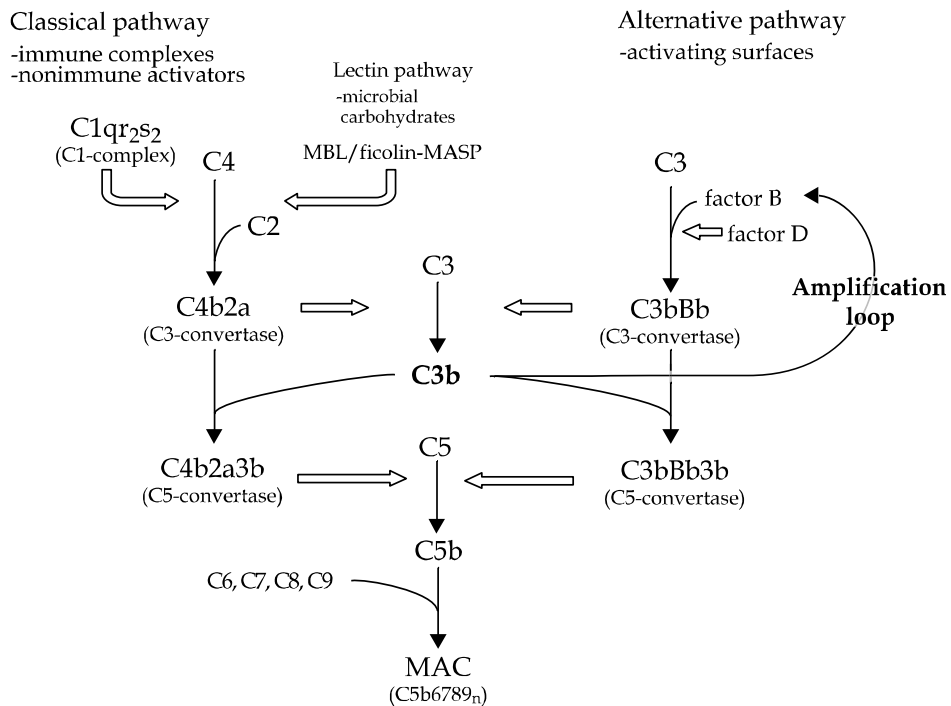
The complement system is a major part of innate immunity and is essential as an early participator in the defense against invading microbes. Moreover, the complement system is implicated in diverse tasks such as clearance of immune-complexes and apoptotic cells, and stimulation of acquired immune responses. The importance of complement in human immunity is demonstrated by individuals with certain complement-deficiencies, who suffer from recurring microbial infections, immune-complex diseases and autoimmune disorders (Walport, 2001a, b; Verschoor and Carroll, 2004).

At the end of the 19<sup>th</sup> century, the human complement system was discovered and described as a “complement” to antibody-mediated killing of microbes and was also found to have the ability to lyse red blood cells of other species (Morgan and Harris, 1999). Many of the individual players in the complement system were isolated and characterized in the 1950’s and 1960’s, and were designated C1-C9, in order of discovery. Notably, the order of action of these proteins is C1-C4-C2-C3-C5-C6-C7-C8-C9. However, complement also comprises several other components and is a cascade system of >35 soluble and surface-bound proteins, where one enzymatically cleaved molecule cleaves and activates the next (Morgan and Harris, 1999; Prodinger *et al.*, 2003). As discussed below, complement can be activated through one of three pathways, the classical, lectin or alternative pathway, which result in formation of so called C3-convertases. As signified by their names, the C3-convertases cleave and activate C3, the key molecule of complement. At the C3 conversion step, the three pathways converge into one (Figure 1).

### **Activation of complement**

#### ***Classical pathway***

The classical activation pathway was the first complement pathway described (Morgan and Harris, 1999). The traditional belief has been that activation of the classical pathway (CP) is completely dependent on specific antibodies, but it is now accepted that a number of non-immune activators exist, such as C reactive protein (CRP) or bacterial lipopolysaccharide (LPS) (Gewurz *et al.*, 1993). Moreover, there is evidence that broadly reacting so called natural IgM antibodies play an important role in triggering of the CP through spontaneous low-level activation (Ochsenbein and Zinkernagel, 2000; Manderson *et al.*, 2001). The first



**Figure 1.** Schematic outline of the human complement system and its activation. Modified from Carlsson (2005) and Sandin (2005).

molecule in the CP is C1, a large hetero-oligomer composed of three types of subunits, C1q, C1r and C1s, in a 1:2:2 complex. C1q has a collagen-like tail and six globular heads, and the elongated C1r<sub>2</sub>:C1s<sub>2</sub> component binds between these parts in C1q. The C1r<sub>2</sub>:C1s<sub>2</sub> component, which is formed only in the presence of Ca<sup>2+</sup>, is responsible for the enzymatic activity of the intact C1-complex. The globular heads of the C1q molecule bind to the Fc-part of IgG and IgM, but importantly, simultaneous engagement of multiple heads must occur to initiate activation. Such multiple-point interaction can only be caused by aggregated Ig, *i.e.* IgG bound to an antigen, preventing constant triggering of C1 by fluid-phase IgG, or by insufficient amounts of surface-bound Ig. Upon multiple head engagement, a conformational change is induced within C1q, inducing the auto-activation of the C1r proenzyme, which in turn cleaves and activates C1s. C1s then exerts its enzymatic activity on the next molecule of the cascade, C4. When C4 is cleaved, a short peptide, C4a, is released and a highly reactive

thioester is exposed in the remaining C4b molecule. The thioester anchors the C4b molecule to amino or hydroxyl groups on the activating surface, or is rapidly hydrolyzed, if no surface is present in the immediate vicinity. In an  $Mg^{2+}$ -dependent manner, plasma C2 may now bind to the membrane-attached C4b and may then be cleaved by the C1-complex. The cleavage of C2 releases the fragment C2b and leaves the larger C2a bound to C4b, causing formation of the CP C3-convertase C4b2a, which can cleave and activate the key molecule of the complement system, C3. In plasma, C3 is the most abundant of the complement proteins (1-2 mg/ml), and it is essential for full function of complement, regardless of activation pathway. Cleavage of C3 by C4b2a causes formation of the anaphylatoxin C3a (see below) and C3b, which contains a highly reactive thioester and may bind to a neighboring surface, like C4b. At this step, the CP converges with the lectin and alternative pathways (Morgan and Harris, 1999; Prodinge *et al.*, 2003; Verschoor and Carroll, 2004).

### ***Lectin pathway***

The most recently discovered, but probably the most evolutionarily ancient, pathway is the lectin pathway (LP) (Ikeda *et al.*, 1987; Fujita, 2002). Activators of the LP are molecules of the innate immune system: mannose-binding lectin (MBL, a so called collectin) and a group of proteins known as ficolins (Gadjeva *et al.*, 2001; Holmskov *et al.*, 2003). These proteins are present in plasma and on mucosal surfaces, and may directly recognize certain bacterial surface carbohydrates, such as mannose and glucose groups. The activators of the LP all contain a collagenous region and a carbohydrate-binding part composed of either a C-type lectin (the collectins) or fibrinogen-like domains (the ficolins). When MBL or a ficolin binds to a bacterium, the LP is activated by the MBL-associated serine proteases (MASPs). The MBL/ficolin-MASP complexes share several similarities with the C1-complex, where MBL and the ficolins structurally resemble C1q, and MASP-2 is a functional equivalent of C1s, cleaving C4 and C2, thus causing the formation of a classical pathway C3-convertase, C4b2a. In addition, a MBL-MASP-1 complex may directly cleave C3, but the physiological relevance of this activity is unclear because of its relative inefficiency (Holmskov *et al.*, 2003). Because activation of the LP results in formation of the CP C3-convertase, the LP may be considered a branch of the CP. Interestingly, MBL has been shown to bind to several important pathogens, among them *S. pyogenes* (Neth *et al.*, 2000), and more recently, one of the ficolins was reported to specifically interact with LTA, a constituent of the Gram-positive cell wall (Lynch *et al.*, 2004), suggesting that the LP may be active in the early defense against invading bacteria.

### *Alternative pathway*

In plasma, spontaneous hydrolysis of the internal thioester in C3 to form C3(H<sub>2</sub>O) occurs continuously (Pangburn *et al.*, 1981). The C3(H<sub>2</sub>O) molecule has many features of C3b, but does not have the ability to attach to adjacent surfaces, thus resides in the fluid phase. Factor B (fB), a protein with structural and functional similarities to the CP component C2, binds C3(H<sub>2</sub>O) in an Mg<sup>2+</sup>-dependent manner and may subsequently become cleaved by factor D, a serine protease present in plasma in minute amounts. The cleavage product Bb remains attached to C3(H<sub>2</sub>O), resulting in formation of C3(H<sub>2</sub>O)Bb, a fluid-phase C3-convertase which probably is continuously formed but rapidly degraded. This soluble C3-convertase may cleave additional C3, in the same manner as the CP convertase, into C3a and C3b. Via the exposed internal thioester, C3b may attach to a surface and form the AP C3-convertase, C3bBb, which then initiates a positive feed-back loop, generating C3b that can deposit onto the same surface. The formation of this amplification loop is a key property of the AP and may be its most important function. Notably, C3b does not discriminate between self and non-self surfaces, but on human cells the amplification loop will normally not arise and cause complement deposition, due to the presence of surface-bound complement regulators. Moreover, surface-exposed sialic acid increases the affinity of the soluble inhibitor factor H (FH, see below) for surface-deposited C3b. Such inert surfaces are referred to as “non-activating”, while *e.g.* bacterial surfaces, not containing sialic acid or complement inhibitors, are referred to as activating surfaces (Morgan and Harris, 1999; Proding *et al.*, 2003; Verschoor and Carroll, 2004).

The constant low-level hydrolysis of C3, or C3 “tick-over”, and subsequent deposition of C3b on a foreign surface is one mechanism by which the AP and its amplification loop may be activated. However, it is likely that a more important manner of activation *in vivo* is through generation of the first C3b via the classical or lectin pathways, followed by accelerated C3b deposition through the AP loop (Morgan and Harris, 1999). In agreement with this proposition, a number of recent studies indicate that initiation of complement deposition on several important human pathogens occur via the CP (Brown *et al.*, 2002; Barnes and Weiss, 2003; Carlsson *et al.*, 2003; Ferguson *et al.*, 2004; Ren *et al.*, 2004; Carlsson *et al.*, 2005).

## **Functions of complement**

### ***Opsonization and phagocytosis***

The most important function of the complement system is probably to cause opsonization, *i.e.* to promote recognition and killing of invading microorganisms by phagocytic cells. Phagocytosis is promoted by an activated complement cascade through the deposition of C3b and iC3b (so called inactivated C3b, see “*Regulators in plasma*”), and to some extent C4b, on the surface of the foreign particle. A microbe opsonized by complement is visible to neutrophils and macrophages that display specific complement receptors (CR). There are a number of receptors specific for different complement components or fragments, where the main receptor involved in phagocytosis is CR3, which binds iC3b. CR4, another receptor for iC3b, and receptors for C3b, C4b, C1q and MBL have also been implicated in phagocytosis, but the function of these receptors in phagocytosis is not as extensively documented as that of CR3. Invading pathogens may also be opsonized by antibodies, primarily of the IgG class, and for optimal phagocytosis to occur, complement and CRs act in concert with antibodies and Fc-receptors (Brown and Gresham, 2003; Prodingen *et al.*, 2003; Verschoor and Carroll, 2004).

### ***Membrane attack complex***

The final product of the complement system is the membrane attack complex (MAC), which is formed through the terminal pathway. The first actors in the terminal pathway are the C5-convertases, which bind to and cleave component C5. These convertases are formed when an additional C3b molecule associates with a preexisting C3-convertase, produced via the classical or the alternative pathway, resulting in formation of C4b2a3b or C3bBb3b, respectively, and shifting the convertase specificity from C3 to C5. The C2a and Bb units of the convertases cleave C5 in the last enzymatic step of the complement cascade, generating the very potent anaphylatoxin C5a and the larger fragment C5b, which remains bound to C3b in the convertase. Complement components C6 and C7 then associate with C5b. The C5b67 trimer may be inserted into adjacent membranes, where it recruits C8 and C9, causing the formation of a pore in the affected membrane and lysis of the cell through osmotic shock. Oligomerization of C9 molecules in the complex widens the pore and is necessary for efficient lysis. Importantly, only some gram-negative bacteria are sensitive to lytic killing by MAC. The thick peptidoglycan cell wall of gram-positive bacteria, such as *S. pyogenes*, prevents insertion of MAC into the membrane. Thus, gram-positive bacteria are resistant to lysis mediated by the terminal pathway of complement. Moreover, many gram-negative

bacteria have surface structures that make them resistant to the MAC. Therefore, the MAC is much less important than the C3b opsonin for the biological function of the complement system.

### ***Anaphylatoxins***

During the progression of the complement cascade, a number of small fragments are released upon cleavage of certain components. The C3a and C5a fragments are potent pro-inflammatory mediators and are known as anaphylatoxins, while no function for C4a has been found in humans. The anaphylatoxins exert important biological functions, such as recruitment and activation of leukocytes, release of inflammatory cytokines, contraction of smooth muscle and upregulation of Fc-receptors on phagocytes (Kumar *et al.*, 2006). Receptors for C3a and C5a are present on a variety of cells, mainly on myeloid lineage cells such as neutrophils and macrophages, but also on epithelial and endothelial cells and possibly also B and T cells (Morgan and Harris, 1999; Prodinge *et al.*, 2003; Verschoor and Carroll, 2004).

### ***Additional roles for complement***

Immune complexes (IC) are antigen-antibody aggregates that may be deposited in capillaries and cause inflammation. ICs with deposited C3- or C4-derived fragments are recognized by complement receptor 1 (CR1) on erythrocytes and are directed to the liver and spleen for removal (Morgan and Harris, 1999; Prodinge *et al.*, 2003; Verschoor and Carroll, 2004). Complement has also been implicated in clearance of cells undergoing apoptosis (programmed cell death). The C1q, MBL and iC3b components have been reported to play a role in the uptake of apoptotic cells by macrophages, and the complement regulator C4b-binding protein, which is studied in this thesis, is recruited to the surface of apoptotic cells, and may inhibit inflammation by down-regulation of complement activation (Webb *et al.*, 2002; Kask *et al.*, 2004).

Complement also provides a link between innate and acquired immunity. The surface-localized protein CR2 (CD21) is a member of the B cell coreceptor complex and may interact with iC3b, C3dg and C3d (see below). Any of these complement products may, when attached to an antigen that is recognized by the B cell receptor, cross-link the B cell receptor with the coreceptor complex, which may lead to a lowered threshold of B cell activation and subsequent antibody production (Carroll, 2004). It has also been suggested that complement plays a role in the T cell response to infection, but the mechanism remains obscure (Carroll,



2004). Interestingly, complement may be involved in the induction of regulatory T cells, because costimulation of the T cell receptor and the cell-bound complement regulator CD46 induces a regulatory phenotype in T cells (Kemper *et al.*, 2003; Price *et al.*, 2005).

Finally, there is evidence that complement has important functions in bone and organ regeneration. Moreover, complement regulators serve to protect the reproductive system and the fetus, and possibly have a direct role in fertilization (Morgan and Harris, 1999; Mastellos and Lambris, 2002).

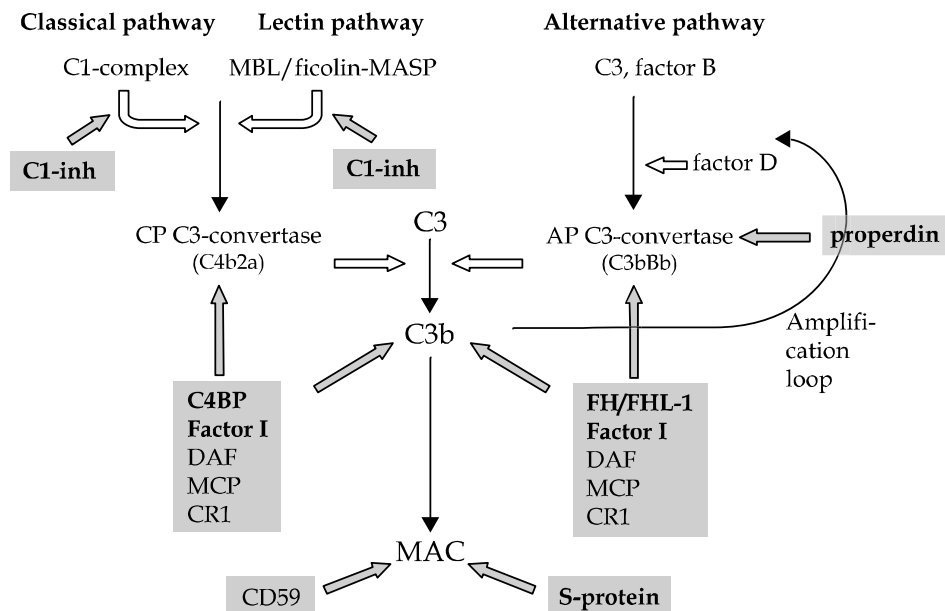
### **Regulation of complement activation**

The complement system is an extremely potent system, capable of causing considerable damage to human tissue. Two crucial constraints prevent unlimited complement activation. First, the short half-life of the activated components C3b and C4b and the instability of the convertases confines complement deposition to the close proximity of the site of activation. Second, the presence of numerous plasma and cell-bound regulators, that are active at different steps in the cascade, effectively controls activation of the system (Figure 2).

#### ***Regulators in plasma***

The C1-inhibitor (C1-inh), a representative of the serine protease inhibitors (serpins), is the only known regulator of the first step of activation via the CP. An assembled C1-complex may undergo spontaneous activation at low levels, in the absence of antibodies. The C1-inh may reversibly bind to an assembled C1-complex and inhibit such autoactivation. Moreover, the C1-inh may irreversibly bind to, and dissociate C1r and C1s from C1q in an activated C1-complex, thereby limiting CP activation. C1-inh may also inhibit the LP by binding to active MASPs in an MBL/ficolin-MASP complex.

At the stage of the C3-convertases, C4b-binding protein (C4BP) and factor H (FH), or the FH splice variant factor H-like protein-1 (FHL-1), may inhibit the formation and accelerate the decay of CP and AP convertases through binding to C4b and C3b, respectively. Moreover, C4BP and FH/FHL-1 act as cofactors to the serine protease factor I (fI) in the degradation of C4b and C3b. Because C4BP is central to the work described in this thesis, this particular regulator will be more thoroughly described in a separate section. When bound to FH, C3b is cleaved by fI at two sites, generating the large, surface-attached iC3b and releasing the smaller C3f. Of note, iC3b can not form a C3-convertase and a fI-cleaved C3b molecule is thus excluded from the AP amplification loop. However, iC3b is an important



**Figure 2.** Schematic representation of the regulation of the complement system. Complement regulators (grey boxes) are either plasma proteins (bold) or cell membrane proteins. All regulators except properdin (which stabilizes the AP C3-convertase) inhibit activation at the step indicated. Modified from Johnsson (1999) and Sandin (2005).

opsonin and interacts with CR3, the main complement receptor active in stimulation of phagocytosis. Thus, iC3b is not “inactive”, as suggested by its (unfortunate) name.

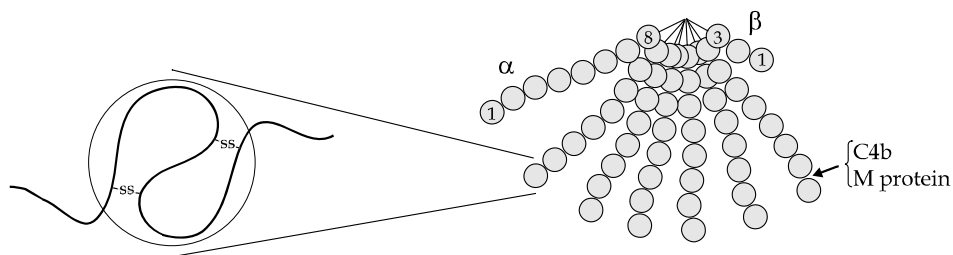
The only positive regulator of complement, properdin, is a stabilizer of the AP convertase C3bBb. It binds to an existing convertase and considerably prolongs its half-life in serum. In addition, properdin may inhibit fI-mediated inactivation of C3b.

S-protein (vitronectin) is one of several molecules with the ability to bind to the soluble precursor of MAC, C5b67, inhibiting its insertion into a cell membrane, thereby limiting lysis mediated by the terminal pathway of complement.

The anaphylatoxins are regulated by carboxypeptidase-N which removes a C-terminal Arg residue from C3a and C5a, converting them into C3a<sub>desArg</sub> and C5a<sub>desArg</sub>. C3a<sub>desArg</sub> is unable to interact with the C3a-receptor, and is considered biologically inactive, while C5a<sub>desArg</sub> still possesses limited proinflammatory capacity (Morgan and Harris, 1999; Proding *et al.*, 2003; Verschoor and Carroll, 2004).

### **C4BP**

C4BP is a large (~570 kDa), heavily glycosylated protein, present in normal human plasma in concentrations of 150-300 mg/l (Scharfstein *et al.*, 1978; Dahlbäck, 1991). However, during the acute phase response the concentration of C4BP may be increased significantly (Barnum and Dahlbäck, 1990). C4BP is a major fluid-phase regulator of the CP of complement. By binding to C4b, C4BP inhibits the formation of the CP C3-convertase and may also accelerate the decay of an existing convertase. Moreover, C4BP acts as a cofactor in the degradation of C4b by fI, which cleaves C4b at two sites, releasing the larger fragment C4c, while the smaller C4d fragment remains attached to the surface via the thioester. C4BP has also been suggested to act as a cofactor in the degradation of C3b, but the physiological relevance of this interaction is uncertain (Blom *et al.*, 2003).



**Figure 3.** Representation of a CCP domain (left) (Modified from Morgan and Harris, 1999) and human C4BP (right). A CCP domain contains ~60 aa held together by two disulphide bridges. The major C4BP isoform in human plasma contains seven  $\alpha$ -chains and one  $\beta$ -chain. These subunits are composed of eight and three CCP domains, respectively. The natural ligand C4b and streptococcal M protein bind at the CCP1-2 interface, as indicated.

C4BP belongs to a family of proteins known as regulators of complement activation (RCA). The members of the RCA-family are proteins composed of domains referred to as short consensus repeats (SCRs) or complement control protein domains (CCPs) (Liszewski *et al.*, 1996; Morgan and Harris, 1999). Such domains are ~60 aa long and are stabilized by two disulphide bridges (Figure 3). The C4BP molecule is normally composed of seven 70 kDa  $\alpha$ -chains and one 45 kDa  $\beta$ -chain, containing eight and three CCPs, respectively. In addition to this  $7\alpha:1\beta$  form, C4BP also exists in the isoforms  $7\alpha:0\beta$  and  $6\alpha:1\beta$  (Hillarp *et al.*, 1989; Criado García *et al.*, 1995). The chains of C4BP are held together by disulphide bonds in a

central core, via their C-terminal  $\alpha$ -helical regions, giving the molecule an octopus-like shape (Dahlbäck *et al.*, 1983; Perkins *et al.*, 1986) (Figure 3). Each  $\alpha$ -chain has a C4b-binding site in CCP1-2 (Dahlbäck *et al.*, 1983; Blom *et al.*, 1999), but under physiological conditions only four C4b molecules can bind to C4BP, apparently due to sterical hindrance (Scharfstein *et al.*, 1978; Ziccardi *et al.*, 1984).

The C4BP  $\beta$ -chain binds the anticoagulant plasma component protein S (PS), and in humans virtually all  $\beta$ -chain-containing C4BP circulates in complex with this protein (Dahlbäck and Stenflo, 1981; Dahlbäck, 1991). PS bound to C4BP cannot exert its inhibitory function in the coagulation cascade, and in agreement with this observation patients that express high levels of C4BP  $\beta$ -chain seem to suffer an increased risk of thrombotic disease (Esparza-Gordillo *et al.*, 2004). Interestingly, in an acute phase response, expression of the non- $\beta$  C4BP isoform is elevated to a greater extent than the  $\beta$ -chain-containing isoforms (García de Frutos *et al.*, 1994; Criado García *et al.*, 1995). Because of the high-affinity association of the  $\beta$ -chain with PS, and the molar excess of PS as compared to C4BP in plasma, the elevated synthesis of the non- $\beta$  forms may reflect a mechanism to maintain appropriate levels of free PS (Griffin *et al.*, 1992). Moreover, C4BP is recruited to apoptotic cells as a result of the interaction between C4BP-bound PS and membrane structures exposed specifically on the apoptotic cell surface (Webb *et al.*, 2002, 2003). Because of the necessity to not induce inflammation upon apoptosis, the engagement of C4BP may be of importance in down-regulating complement at the surface of apoptotic cells.

C4BP also interacts with heparin and serum amyloid P component (SAP) (Hessing *et al.*, 1990; Schwalbe *et al.*, 1990). Neither of these interactions is fully understood, but a possible function of heparin-binding is to localize C4BP to host cell surfaces, thus protecting self tissue from complement deposition and damage. The interaction with SAP is even less understood and may represent an *in vitro* artifact, because SAP was not found to circulate with C4BP under physiological conditions (Sen and Heegaard, 2002). Recently, a novel C4BP ligand was described. Brodeur *et al.* (2003) reported that C4BP binds to CD40 on B cells and may induce activation of these cells independently of the standard activator CD40L. The authors speculate that because C4BP is found associated with B cells in locales with virtually undetectable levels of CD40L, the CD40-C4BP interaction may be a redundancy mechanism playing important roles for B cell survival, proliferation and differentiation when CD40L is not present (Brodeur *et al.*, 2003).

Many pathogenic microbes have developed mechanisms for protection against complement deposition. One such mechanism is the recruitment of soluble complement

regulators which down-regulate complement deposition on the microbial surface (Lindahl *et al.*, 2000). A number of bacterial species have been shown to bind human C4BP, of which the first described and most extensively studied interaction is that with *Streptococcus pyogenes* (Thern *et al.*, 1995; Jenkins *et al.*, 2006). Different aspects of this interaction are the focus of this thesis, and it will be discussed in greater detail below. C4BP has also been shown to bind some other pathogens. All tested strains of the causative agent of whooping cough, *Bordetella pertussis*, bind C4BP, and the binding is significantly reduced in bacterial mutants lacking the virulence factor filamentous hemagglutinin (Berggård *et al.*, 1997). Moreover, all strains of *Neisseria gonorrhoeae* (gonococcus), a Gram negative bacterium that causes the sexually transmitted disease gonorrhea, bind C4BP to the major neisserial surface proteins Por1A and Por1B, and contributes to serum resistance (complement resistance) in these strains (Ram *et al.*, 2001). Interestingly, gonorrheal disease is found only in humans, and a recent report indicates that the species specificity of gonococci may be due to the interaction with human C4BP, but not with C4BP from other species (Ngampasutadol *et al.*, 2005). In control experiments in the same study, the plague bacterium *Yersinia pestis* also bound C4BP. In addition, C4BP has been demonstrated to bind to gonococcal type IV pili (Blom *et al.*, 2001), to the outer membrane protein A (OmpA) of *Escherichia coli* K1 (Prasadarao *et al.*, 2002), to the ubiquitous surface proteins A1 and A2 (UspA1 and UspA2) of the mucosal pathogen *Moraxella catarrhalis* (Nordström *et al.*, 2004) and to the surface of *N. meningitidis*, possibly via the outer membrane protein PorA and/or the sialic acid capsule (Jarva *et al.*, 2005). In all cases analyzed, bacteria-bound C4BP retains its cofactor activity, suggesting that the interaction may be of importance in escaping complement attack, but conclusive evidence for a role in pathogenesis has only been provided for *S. pyogenes* M protein (Carlsson *et al.*, 2003).

### **Cell-bound regulators**

Decay accelerating factor (DAF, CD55) binds to and accelerates the decay of C3- and C5-convertases formed on host cell membranes. CD55 is a GPI-anchored protein that is present on erythrocytes, platelets and all nucleated cells except natural killer (NK) cells and some subsets of T cells.

Membrane cofactor protein (MCP, CD46) is expressed by all nucleated human cells, but not by erythrocytes. In contrast to CD55, CD46 is not active in decay of the convertases, but acts as a cofactor to fl in the degradation of C3b and C4b. Like CD55, CD46 is functional only if the substrate and CD46 are bound to the same cell membrane.

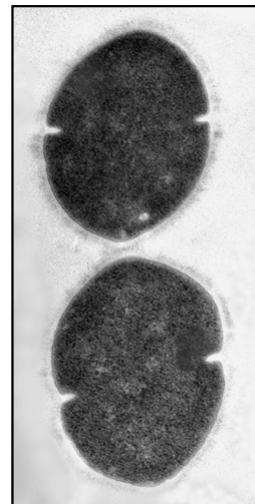
Complement receptor 1 (CR1, CD35) is a transmembrane glycoprotein mainly found on circulating cells. CR1 acts as a cofactor to  $f_1$  in the degradation of C3b and C4b, and accelerates decay of the CP and AP C3-convertases. As previously mentioned, CR1 is abundant on erythrocytes and mediates clearance of ICs, which probably represents the most important physiological function of CR1.

CR1, CD55 and CD46 are all members of the RCA-family of proteins and they regulate both the CP and AP of activation. Interestingly, CD55 and CD46 are exploited as cellular receptors by many pathogens (Lindahl *et al.*, 2000). For the subject described in this thesis, it is of particular interest that streptococcal M proteins have the ability to bind CD46. This interaction does not seem to be directly concerned with deposition or regulation of complement, but may contribute to adhesion of the bacteria to host tissue (Okada *et al.*, 1995; Rezcallah *et al.*, 2005). Moreover, a recent study suggests a new function for the M protein-CD46 interaction, the induction of a regulatory phenotype in naïve CD4+ T cells (Price *et al.*, 2005).

The major regulator of the terminal pathway of complement is the GPI-anchored glycoprotein CD59. This protein is expressed at high levels by virtually all cells in the human body and binds to C8 in the C5b678 complex, thereby restricting C9 aggregation and lytic pore formation.

### *STREPTOCOCCUS PYOGENES*

The Greek words strepto and coccus mean “string of” and “round”, and illustrate the typical appearance of bacteria in the *Streptococcus* genus. This genus comprises Gram-positive bacteria that are spherical in shape, typically growing in pairs or chains. Most species of streptococci are facultative anaerobes, i.e. they are able to grow both in absence and presence of oxygen. The first attempts to classify streptococci were made ~100 years ago and were based on the ability of the bacteria to cause lysis of red blood



**Figure 4.** Electron micrograph of *Streptococcus pyogenes*. The surface exposed M protein, a major virulence factor, is visible as a tuft-like layer surrounding each bacterium (Swanson *et al.*, 1969). (Electron micrograph, courtesy of Eric Carlemalm.)

cells, which is observed on blood-agar plates. Shortly thereafter, these different hemolytic reactions were more extensively described as alpha ( $\alpha$ , incomplete or green hemolysis), beta ( $\beta$ , complete or clear hemolysis) and gamma ( $\gamma$ , no hemolysis) hemolysis (Stevens and Kaplan, 2000). The first serological classifications of streptococci were based on group-specific carbohydrates (groups A - O) present in the cell-wall of different species in this genus (Lancefield, 1933, 1962).

*Streptococcus pyogenes* (Figure 4) is a  $\beta$ -hemolytic, chain-forming coccus and is often referred to as Lancefield group A *Streptococcus* (GAS) because of its cell-wall group A carbohydrate.

### **Infections and sequelae**

According to a recent estimate, >500,000 deaths each year can be attributed to acute *S. pyogenes* infections or consequences thereof, making this bacterium one of the most important human pathogens globally. Moreover, the number of individuals with current infections was estimated to ~120 million, and the number of new cases each year was estimated to ~620 million. Most of these infections are superficial, with low risk of death, but with enormous socio-economical consequences (Carapetis *et al.*, 2005b).

Importantly, *S. pyogenes* is a strict human pathogen, *i.e.* it only causes disease in humans. Although *S. pyogenes* is usually not considered part of the normal flora, asymptomatic carriage of the pathogen is not uncommon, particularly in school children and younger children attending day-care centers (Cunningham, 2000; Courtney *et al.*, 2002; Bisno *et al.*, 2003).

The diseases caused by *S. pyogenes* cover a wide spectrum, from mild infections of the throat and skin to life-threatening deep-tissue infections. The main routes of infection are through the upper respiratory tract or abrasions in the skin (Courtney *et al.*, 2002). The most common *S. pyogenes* disease is acute pharyngitis, or “strep throat”, which mainly affects children (Carapetis *et al.*, 2005b). The main symptoms of pharyngitis are sore throat, pain upon swallowing and fever. Moreover, general swelling of the pharynx and pus-containing, white patches on the tonsils are common visible characteristics of this disease (Bisno, 2001). Before the introduction of antibiotics, scarlet fever was a common complication of pharyngitis and often had a deadly outcome due to septic shock. The sporadic cases of scarlet fever that now are seen in the western world are generally mild and can be described as streptococcal pharyngitis with skin rash (Stevens and Kaplan, 2000).

Impetigo, or pyoderma, is a purulent infection of the skin and is the second most common manifestation of *S. pyogenes* disease. Impetigo is characterized by superficial skin lesions, which develop into rupturing blisters that become covered by a thick crust, generally appearing on the extremities, trunk or face of young children (Bisno and Stevens, 1996; Stevens and Kaplan, 2000). Pharyngitis and impetigo are both mild, generally self-limiting diseases. However, if left untreated, infection may spread to deeper tissues or, from pharyngitis in particular, acute rheumatic fever may develop (see below).

Erysipelas and cellulitis are skin infections that, in contrast to pharyngitis and impetigo, mainly occur in adults and the elderly (Stevens and Kaplan, 2000). Both infections present with local signs of inflammation, fever and at times inflammation of lymphoid tissue. However, while erysipelas develops in the superficial layers of the skin and cutaneous lymphatic vessels, cellulitis involves deeper subcutaneous tissues (Bisno and Stevens, 1996).

In earlier centuries, puerperal sepsis (childbed fever) was a common cause of death among postpartum women. These infections were caused by *S. pyogenes* (Lancefield and Hare, 1935). In the 1840's, when it was recognized by Semmelweis that the majority of these women were contaminated by their physicians, and that the infection could be avoided by careful hand hygiene, the incidence of the disease was much reduced and it is rare today (Semmelweis, 1861; Stevens and Kaplan, 2000). Other invasive and life-threatening syndromes caused by *S. pyogenes* are necrotizing fasciitis and streptococcal toxic shock syndrome, disorders that unfortunately have increased dramatically over the last decades and have awarded *S. pyogenes* massive medial attention as “flesh-eating bacteria” (Bisno and Stevens, 1996; Stevens, 1999).

Two non-suppurative sequelae are associated with *S. pyogenes* infection, acute rheumatic fever (ARF) and post-streptococcal glomerulonephritis (PSGN). ARF is believed to be an autoimmune disease that is caused by cross-reactivity of antibodies directed against streptococcal surface proteins with human tissue. The major manifestations of ARF are inflammation of the heart valves, skin, joints and CNS (Cunningham, 2000; Stevens and Kaplan, 2000). In PSGN, the most common clinical signs are hematuria, back pain and hypertension. The pathogenesis behind PSGN is not fully understood, but theories include deposition of bacterial proteins or immune complexes, or both, in the glomeruli of the kidneys (Stevens and Kaplan, 2000). It is a widely accepted concept that *S. pyogenes* strains can be either rheumatogenic or nephritogenic, and that PSGN may follow either pharyngitis or impetigo, while ARF almost invariably is a consequence of respiratory tract infection (Bisno and Stevens, 1996; Stevens and Kaplan, 2000). However, in populations with high



frequencies of mild *S. pyogenes* infections, the concept that strains cause either ARF or PSGN is distorted (Carapetis *et al.*, 2005a). This distortion is highlighted by the high incidence of ARF in the aboriginal population in Australia, a population where impetigo cases significantly outnumber respiratory tract infections (Bessen *et al.*, 2000; McGregor *et al.*, 2004a; Carapetis *et al.*, 2005a).

*S. pyogenes* infection can most often be effectively treated with antibiotics, and penicillin is usually the drug of choice. Fortunately, all strains are still susceptible to penicillin and similar antibiotics, but strains resistant to other classes of antibiotics, *e.g.* erythromycin, occur with relatively high frequency (Stevens and Kaplan, 2000). In cases of severe necrotic *S. pyogenes* disease, treatment with antibiotics may be insufficient and surgical removal of involved tissue is often required.

### **Classification**

The cell-wall associated M protein is a highly diverse surface protein that allows for subdivision of *S. pyogenes* strains into M types. Importantly, all clinical isolates express an M protein (Kehoe, 1994), which gives rise to protective antibodies specific for the given M type. Initially, M typing was achieved serologically by demonstration of a reaction with type-specific antibodies raised in rabbits (Lancefield, 1962). Today, new types are identified through sequencing of the variable 5' end of the *emm* gene (Beall *et al.*, 1996). This sequence encodes the variable N-terminal region of M protein, which elicits the type-specific antibodies mentioned above. At present, >100 M types have been recognized, where types 1-81 were defined using the original serological method, types 82-93 were first determined by sequence and subsequently by serology (Facklam *et al.*, 1999), while types 94-124 were defined by sequence only (Facklam *et al.*, 2000; Facklam *et al.*, 2002). To separate types that are characterized serologically from those defined solely by sequence, actual serotypes are referred to as M types, while sequence types are denoted *emm* types (Facklam *et al.*, 2002). However, for simplicity, serological M types and sequence determined *emm* types will not be distinguished in this thesis. Thus, all types will be referred to as M types.

Classification of isolates based on the expression of the so called serum opacity factor (OF or SOF) is commonly used and separates strains into two main lineages, OF+ and OF-. For unknown reasons, the OF type of a particular strain is strongly correlated with its M type (Gooder, 1961; Widdowson *et al.*, 1970). In some studies, throat isolates were almost

invariably OF-, while skin strains were both OF+ and OF- (Cunningham, 2000; McGregor *et al.*, 2004b).

Additionally, strains can be categorized with regard to the so called T antigens. These antigens are trypsin resistant surface proteins and ~25 distinct serotypes have been defined (Kehoe, 1994; Johnson *et al.*, 2006). Certain T types are often associated with certain M types, although a single M type can express more than one T antigen and a particular T antigen can be expressed by different M types (Kehoe, 1994; Beall *et al.*, 1997). For many years, the function of the T antigens remained unclear. However, in a recent report light was shed on the role of the T antigens, when the T6 antigen was shown to form pilus-like structures on the bacterial surface. Interestingly, mice immunized with recombinant pilus proteins were protected against infection with *S. pyogenes* (Mora *et al.*, 2005).

### **Genomics**

The genome sequence of an *S. pyogenes* strain was first determined for an M1 strain (Ferretti *et al.*, 2001), and subsequently the genome sequence of a second M1 strain (Sumby *et al.*, 2005), two strains of type M3 (Beres *et al.*, 2002; Nakagawa *et al.*, 2003), and one each for strains of type M6 (Banks *et al.*, 2004), M18 (Smoot *et al.*, 2002) and M28 (Green *et al.*, 2005) have been published. Unpublished sequence data are also available for an M5 strain ([http://www.sanger.ac.uk/Projects/S\\_pyogenes](http://www.sanger.ac.uk/Projects/S_pyogenes)). These genomes contain ~1.9 Mbp, and are predicted to encode 1700-1900 proteins. The genomic variability observed between these strains is mainly due to prophages or other foreign elements, while the core chromosome is very similar in all sequenced genomes (Green *et al.*, 2005). Of note, 7 of the 8 genomes sequenced so far represent OF- strains. The M28 strain is the only OF+ strain sequenced to date, but work is in progress on an OF+ strain of type M49 (<http://www.uni-rostock.de/fakult/medfak/mibi/page/forschi/research.html>).

### **Virulence factors**

The ability of a pathogen to cause disease is dependent on the expression of a number of virulence factors, which may be defined as microbial products essential for initiation and maintenance of an infection but not for growth *in vitro*. As stated above, *S. pyogenes* causes many different diseases, both mild and lethal. The ability to cause such a wide variety of diseases may be due to the production of a large number of virulence factors (Cunningham,

2000; Ferretti *et al.*, 2001). The first genome sequence of an *S. pyogenes* strain revealed >40 putative virulence associated genes (Ferretti *et al.*, 2001). It is important to realize that not all strains express all virulence factors, and that different strains may employ different modes of infecting humans. In addition, local environmental factors at separate sites of infection may influence the virulence factors expressed (Cunningham, 2000; Stevens and Kaplan, 2000; Bisno *et al.*, 2003).

The putative and known virulence factors of *S. pyogenes* may be divided into those associated with the bacterial surface and those secreted into the surrounding medium. The surface localized M protein is commonly regarded as the most important virulence factor of *S. pyogenes*, and its ability to allow the bacterium to resist phagocytosis and multiply in human blood has long been recognized (Lancefield, 1962; Fischetti, 1989). This antiphagocytic property of M protein can at least partly be explained by its ability to interfere with complement deposition on the bacterial surface (Jacks-Weis *et al.*, 1982; Whitnack and Beachey, 1982; Carlsson *et al.*, 2003; Carlsson *et al.*, 2005). Some M proteins may also contribute to virulence by acting as adhesins (Okada *et al.*, 1995; Courtney *et al.*, 2002). The hyaluronic acid capsule of *S. pyogenes* also confers phagocytosis resistance (Wessels *et al.*, 1991) and may contribute to adhesion (Cywes and Wessels, 2001). In a primate model of *S. pyogenes* colonization, the M protein and the hyaluronic acid capsule were both found to be important for throat colonization and persistence, a finding attributed to their ability to confer phagocytosis resistance, adhesion, or both (Ashbaugh *et al.*, 2000).

Several surface components, other than M protein and capsule, have been implicated in adhesion to human cells (Courtney *et al.*, 2002). Among them are SOF (which may be cell-bound or secreted), lipoteichoic acid (LTA) and protein F (Courtney *et al.*, 1992; Hanski and Caparon, 1992; Oehmcke *et al.*, 2004), which all bind fibronectin. Moreover, a collection of surface proteins interacting with *e.g.* fibrinogen, vitronectin and collagen have been implicated in adhesion (Cunningham, 2000; Lukomski *et al.*, 2001; Rasmussen and Björck, 2001; Bisno *et al.*, 2003). As mentioned above, surface-localized pili have recently been identified in *S. pyogenes*, and it seems possible that they play an important role in adhesion (Mora *et al.*, 2005). An extensively studied surface protein is the C5a-peptidase, which is expressed by both *S. pyogenes* and *S. agalactiae* and has proteolytic activity specific for the complement-derived chemoattractant C5a, a potent recruiter and activator of neutrophils. Accordingly, the C5a-peptidase may reduce the number of infiltrating neutrophils and thereby promote the establishment of an infection (Wexler *et al.*, 1985; Lindahl *et al.*, 2000). Another

surface protein, designated GRAB, binds human  $\alpha_2$ -macroglobulin, a property that may contribute to virulence (Rasmussen *et al.*, 1999; Nyberg *et al.*, 2004).

*S. pyogenes* isolates also produce a number of putative virulence factors secreted into the surrounding medium. Two hemolysins are expressed by virtually all strains, the pore-forming streptolysin O (SLO) and the peptide streptolysin S (SLS). Both hemolysins are potent cytolytic agents and are toxic to a number of host cells in addition to erythrocytes (Bisno *et al.*, 2003; Wessels, 2005). Most *S. pyogenes* strains also express streptokinase, a protein that binds to plasminogen and thereby induces plasmin activity, which promotes fibrinolysis and clot dissolution (Bisno *et al.*, 2003). The pyrogenic exotoxins (Spe proteins) are often associated with severe streptococcal disease and belong to a group of proteins referred to as superantigens. These proteins may non-specifically activate certain sets of T cells, leading to production of massive amounts of proinflammatory factors that trigger several signal cascades and the complement, coagulation and fibrinolysis systems, a recognized hallmark of the streptococcal toxic shock syndrome (Bisno *et al.*, 2003). Indeed, there is considerable evidence that Spe-like toxins play a key role in the pathogenesis of the streptococcal toxic shock syndrome (Kotb *et al.*, 2002). A number of Spe:s have been described to date, with SpeA, B and C being the most thoroughly characterized. Other factors with superantigen activity are SSA (streptococcal superantigen), SMEZ (streptococcal mitogenic exotoxin Z) and SMEZ-2 (Bisno *et al.*, 2003). SpeB has been classified among the superantigens but is probably more important in its role as a cysteine proteinase that may cleave various host or bacterial proteins (Collin and Olsen, 2001a). Two other secreted enzymes implicated in virulence are EndoS, which deglycosylates IgG (Collin and Olsen, 2001b) and IdeS (Mac-1), which cleaves IgG in the hinge region (Lei *et al.*, 2001; von Pawel-Rammingen *et al.*, 2002).

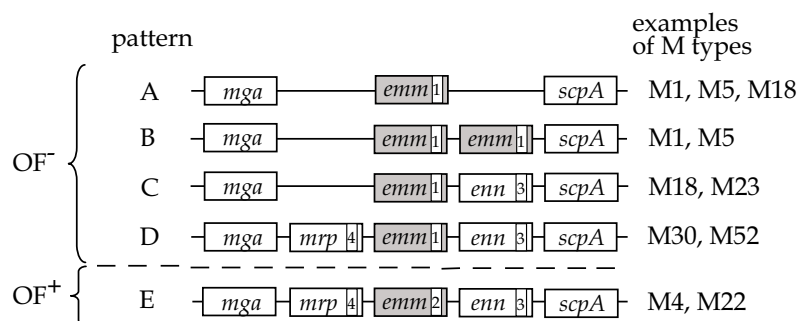
### **The M protein family**

In the 1920's, Rebecca Lancefield first identified the M proteins as serologically distinct surface antigens (Lancefield, 1928). The M protein is a wall-attached coiled-coil protein that gives rise to protective antibodies and has antiphagocytic properties (Lancefield, 1962; Fischetti, 1989). It is a key virulence determinant, and is the most extensively characterized virulence factor of *S. pyogenes*. Initially, *S. pyogenes* strains were thought to express only one M protein, but it is now recognized that a strain may express one, two or three members of the M protein family (Kehoe, 1994), now known as the M (Emm), Mrp and Enn proteins (encoded by genes *emm*, *mrp* and *enn*, respectively). Although all three of these proteins may

have similar function(s) and properties, the designation “M protein” is usually used only for the originally described antiphagocytic M protein, while the Mrp and Enn proteins are designated “M-like proteins”. Of note, the designation Emm has also been used for the M protein expressed by OF+ strains (Thern *et al.*, 1998; Courtney *et al.*, 2006). It should be noted that the classical property of M proteins, the ability to confer resistance to phagocytosis, is a property not only of “true” M proteins but also of Mrp proteins (Podbielski *et al.*, 1996; Thern *et al.*, 1998; Courtney *et al.*, 2006) and possibly of some Enn proteins.

### ***The Mga regulon and the emm locus***

The genes encoding the M, Mrp and Enn proteins are situated in tandem on the bacterial chromosome. These genes comprise the *emm* locus and are located between the *mga* gene, which encodes the positive regulator Mga (Kreikemeyer *et al.*, 2003), and the *scpA* gene, which encodes the previously mentioned C5a-peptidase (Kehoe, 1994) (Figure 5). In addition to the M protein family and the C5a-peptidase, the Mga protein positively controls the expression of several other proteins implicated in virulence, *e.g.* SOF (McLandsborough and Cleary, 1995), and the fibronectin-binding protein Fba (Terao *et al.*, 2001). Together, these Mga-regulated genes comprise the Mga regulon.



**Figure 5.** Organization of M protein family genes (the *emm* locus) in OF- and OF+ strains. The genes *mga* and *scpA*, encoding the positive regulator Mga and the C5a-peptidase, respectively, surround the M protein family genes. The pattern (A-E) is determined by the presence of different subfamily genes (1-4), as indicated (Hollingshead *et al.*, 1993; Bessen and Hollingshead, 1994). Note that all patterns include an *emm* gene (represented in grey), which is expressed in all strains and comes in two variants (subfamily 1 and subfamily 2). Modified from Carlsson (2005) and Sandin (2005).

All strains of *S. pyogenes* have an Mga regulon, but the genes constituting the regulon differ between strains. The Mga protein was initially believed to be the response regulator of a two-component system. However, because a sensory protein has not been found, and phosphorylation of Mga has not been demonstrated, this is still an open question. It is known that Mga responds to elevated CO<sub>2</sub>, iron limitation and increased temperature, and one possibility is that Mga is directly controlled by these environmental factors (Kreikemeyer *et al.*, 2003).

As mentioned above, the *emm* locus of an *S. pyogenes* strain may include one, two or three genes encoding members of the M protein family. Differences in the 3' region separate these genes into four distinct subfamilies (SF1-4) (Hollingshead *et al.*, 1993), and based on the combination of SF genes present, the *emm* locus of a strain may come in one of five possible patterns (A-E) (Bessen and Hollingshead, 1994) (Figure 5). Of note, the *emm* locus of OF+ strains invariably contains three genes, while the *emm* locus of OF- strains are more diverse and may contain 1-3 genes. With few exceptions, strains of patterns A-C are associated with throat infections and pattern D strains with skin infections, while pattern E strains are found at either infection site (Bessen *et al.*, 1996; McGregor *et al.*, 2004b). However, as stated above, this association is not clear-cut in areas with very high infection rates.

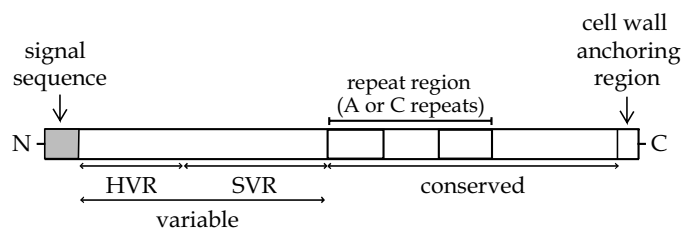
Among clinical isolates of *S. pyogenes*, pattern A and pattern E strains are probably the most important, reflected in the fact that 7 of the 8 genome sequences were determined for OF- pattern A strains, while one was an OF+ pattern E strain. The combined data of Li *et al.* (2003) and McGregor *et al.* (2004) indicate that the vast majority of isolates from invasive disease are pattern A-C or pattern E strains and the isolates were almost evenly distributed between these two groups of strains. Notably, among the pattern A-C strains, strains of pattern A were most prevalent.

Strains of different patterns are believed to have arisen, during evolution, from an ancestral strain of pattern A, *i.e.* the different genes in the *emm* locus may have evolved from an SF1 *emm* gene (Figure 5) (Bessen and Hollingshead, 1994). Recombination events such as gene duplication and subsequent divergence are proposed to explain the generation of the other types of genes in this locus, and strains of patterns D and E are considered the most recently emerged (Bessen and Hollingshead, 1994; Kehoe *et al.*, 1996). Interestingly, SF1 and SF2 genes, encoding two distinct types of M proteins, are mutually exclusive and define two separate lineages of strains (Hollingshead *et al.*, 1993).

There are several lines of evidence for horizontal gene transfer between *S. pyogenes* strains (Kehoe *et al.*, 1996). Examples of M proteins with a mosaic structure that may have arisen through horizontal gene transfer and recombination are found *e.g.* in a pattern B strain of serotype M5 (Whatmore and Kehoe, 1994), and in pattern E strains of serotypes M4 and M60 (Hedén and Lindahl, 1993).

### ***Mrp, M and Enn proteins***

All proteins in the M protein family have similar overall organization and are anchored to the cell wall peptidoglycan via an LPXTG-motif in the C-terminal region (Navarre and Schneewind, 1999; Barnett and Scott, 2002), with the N-terminal protruding into the environment (Figure 6). The C-terminal region is relatively conserved between M proteins, whereas the N-terminal part is more variable. The C-terminal half of an M protein includes a repeat region, which may contain repeats of type A or type C (Hollingshead *et al.*, 1986; Heath and Cleary, 1989; O'Toole *et al.*, 1992). The variability increases towards the N-terminus and the variable half of the protein can be divided into a semivariable region (SVR) and a hypervariable region (HVR). A signal sequence, which has similar sequence in all M proteins, is cleaved off upon secretion. Thus, in a surface exposed M protein, the HVR is the protein region most distal from the bacterial surface (Kehoe, 1994).



**Figure 6.** Schematic representation of an M protein. The C-terminal half of the M protein is relatively conserved while the N-terminal half is more variable and can be separated in a hypervariable region (HVR) and a semivariable region (SVR). The conserved region contains repeats of type A or C. Note that in the surface-expressed protein, the signal sequence (grey) has been cleaved off.

**Mrp** proteins, encoded by *mrp* genes, have conserved repeat region with three so called A-repeats. The variability in the N-terminal part is more limited than among M proteins (encoded by *emm* genes) (Bessen and Hollingshead, 1994). Mrp proteins have been shown to

confer phagocytosis resistance (Podbielski *et al.*, 1996; Thern *et al.*, 1998), and a recent study showed that antibodies against Mrp4 were opsonizing for phagocytosis (Courtney *et al.*, 2006).

**M** proteins are antiphagocytic proteins that give rise to protective immunity and are found in all *S. pyogenes* strains (Fischetti, 1989). The ability of M proteins to confer resistance to phagocytosis is usually analyzed in the so called Lancefield test, which employs fresh human blood and a small bacterial inoculum. After incubation (37°C, usually for 3 h) bacterial growth is determined. Presence of M protein inhibits recognition by neutrophils (PMNs) and allows rapid bacterial growth (Schnitzler *et al.*, 1995; Schnitzler *et al.*, 1999), while strains lacking M protein are phagocytosed. Because of their general importance and because they are the subject of this thesis work, M proteins are described in more detail below.

The conserved repeat region in M proteins has so called C-repeats. Based on antigenic differences within the C-repeat region, M proteins may be divided into two subclasses (I and II), where class I proteins are generally expressed by OF- strains and class II proteins by OF+ strains (Bessen *et al.*, 1989). The M proteins contain a ~50-100 residue N-terminal hypervariable region (HVR). As mentioned above, the HVR is stable within a strain, allowing the identification of >100 different M types (Facklam *et al.*, 2002). Thus, the number of known M types is small, compared to the large number of possible sequence variants, suggesting that these M types have been selected because of their superior fitness.

**Enn** proteins, like M proteins, have a C-repeat region in the conserved half of the protein. The *enn* gene is present in many strains, but is often silent, or expressed at very low levels during *in vitro* growth (Haanes and Cleary, 1989; Jeppson *et al.*, 1992). The function of Enn proteins is not clear. They are not known to elicit protective antibodies and exhibit limited sequence variation in the N-terminal region. Even so, the variability of Enn proteins appears to be intermediate between that of M and Mrp proteins (M > Enn > Mrp) (Bessen and Hollingshead, 1994).

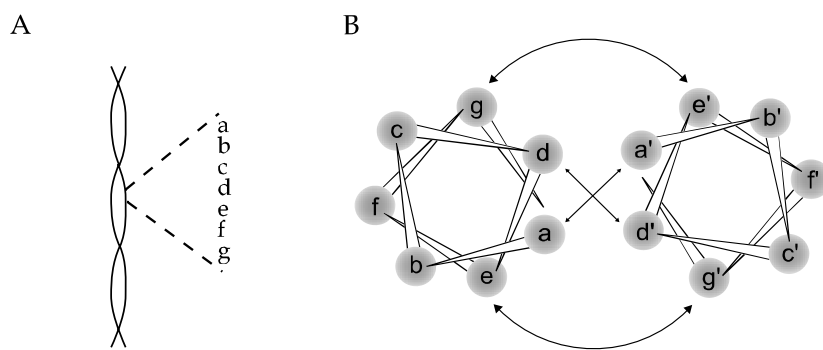
### ***The coiled-coil structure and M protein***

Streptococcal M proteins are homodimeric, 50-60 nm long proteins that largely adopt a so called coiled-coil structure, one of the most commonly found tertiary structures in proteins (Figure 7A) (Lupas and Gruber, 2005). In a coiled-coil, two or more  $\alpha$ -helices twist around each other to form a rope-like fibrillar structure. A coiled-coil dimer, a very common type of coiled-coil, may be composed of two identical or two different helices, thus forming a homo-



or a heterodimer. Moreover, the helices in a coiled-coil can be arranged in a parallel or an anti-parallel fashion, *i.e.* the N- and C-termini of the two helices can be oriented in the same or opposite directions.

The amino acid sequence of coiled-coil proteins exhibits a characteristic seven-residue repeat pattern, with the residue positions designated **a-g** (Figure 7) (Crick, 1953). In Figure



**Figure 7.** Schematic representations of a dimeric coiled-coil protein. **A.** Representation of the rope-like structure of the coiled-coil. The protein is characterized by a heptadic sequence pattern where the residues are designated **a-g**, as indicated. **B.** Helical wheel diagram of a cross-section of the dimeric coiled-coil. The arrows represent the main interactions between residues in the two chains: hydrophobic interactions involving **a** and **d** residues, and electrostatic interactions between **e** and **g** residues

7B, a cross-section of a dimeric coiled-coil protein is depicted in a so called helical wheel diagram (Mason and Arndt, 2004). This representation shows a single heptad repeat of each helix and the arrangement of the **a-g** residues in the coiled-coil. The main inter-helix interactions are also illustrated. These residues are often denoted **a-g** in one helix of the coil and **a'-g'** in the other. Positions **a** and **d** (or **a'** and **d'**) are found on the same face of a helix, and compose the inner core between the two helices of the coiled-coil. The side chain of residue **a** mainly contacts the side chain of an **a'** residue in the other helix, and the same is true for **d** and **d'** residues. The **a/a'** and **d/d'** positions are usually occupied by hydrophobic residues and the stability of the coiled-coil is largely determined by interactions between these hydrophobic residues in the structural core, where improved stability follows from increased hydrophobicity (Hu *et al.*, 1990; Mason and Arndt, 2004). Residues in positions **e** and **g** are

solvent exposed, and they are often polar or charged, accounting for electrostatic interactions between and/or within the helices. The **e** and **g** residues are positioned so that **e** in one helix interacts with **g'** in the other. Residues in positions **e** and **g** are less important for structural stability than those in positions **a** and **d**, but may be important for the local conformation and parallel/anti-parallel helix orientation of the coiled-coil (Hu *et al.*, 1993; Zeng *et al.*, 1997; Mason and Arndt, 2004). The **b**, **c** and **f** residues are often hydrophilic residues, located opposite to the **a** and **d** residues in the helix, and on the outer face of the coiled-coil (Figure 7B). Accordingly, these residues are exposed to the surroundings and may be important for inter-protein interactions rather than for the formation or stability of the coiled-coil structure as such. In an undistorted  $\alpha$ -helix, 3.6 residues make up each turn. However, because of a slight twist to the helix in a coiled-coil, 3.5 residues compose a full turn and consequently one heptad repeat completes two turns (Lupas and Gruber, 2005). Many coiled-coils are dimeric proteins, but different oligomerization states are common and coiled-coils containing up to five helices exist (Mason and Arndt, 2004).

The distribution of residues in the heptadic pattern of M proteins is often non-optimal, with Asn residues frequently occupying position **a** (Manjula and Fischetti, 1980; Nilson *et al.*, 1995; Cedervall *et al.*, 1997). Such deviation from the most favorable pattern often reflects the individual conformation of the coiled-coil, at the expense of structural stability (Mason and Arndt, 2004). For example, the Mrp4 protein deviates less from the optimal heptadic pattern than does the M4 protein expressed by the same strain, and consequently has higher stability (Cedervall *et al.*, 1997).

### ***Binding of human plasma proteins to M proteins***

M proteins interact with a range of host plasma proteins, but the plasma proteins bound may vary between different M proteins. These interactions may confer different properties, such as phagocytosis resistance, ability to adhere or ability to invade tissues.

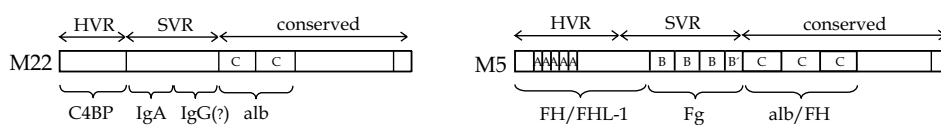
The available evidence suggests that M proteins can be divided into two major groups, depending on their ability to bind either fibrinogen or C4BP (Carlsson *et al.*, 2005). These two groups apparently correspond to class I and class II M proteins, which show slight variations in the sequence of the C-repeats (Bessen *et al.*, 1989). Because fibrinogen and C4BP play important roles for the function of M proteins, the interaction with these plasma proteins will be considered first in this section. The interaction with C4BP will also be described in a separate section, due to its importance for this thesis work. Several other

interactions with human plasma proteins will be described more briefly. Binding of human plasma proteins to two representative M proteins, M5 and M22, is summarized in Figure 8.

*C4BP.* Many *S. pyogenes* strains express an M protein that binds the classical pathway regulator C4BP (Thern *et al.*, 1995; Johnsson *et al.*, 1996; Pérez-Caballero *et al.*, 2000). The C4BP-M protein interaction is central to the work presented in this thesis and will be discussed in greater detail in a separate section below.

*Fibrinogen.* All *S. pyogenes* strains bind human Fg (Kronvall *et al.*, 1979a). Because early work showed that some M proteins bind Fg (Kantor, 1965), it has often been assumed that all M proteins bind Fg, but this is not correct. While many strains express a single M protein that binds Fg (pattern A in Figure 5), the Fg-binding property of OF+ strains is due to the Mrp protein, not the M protein (O'Toole *et al.*, 1992; Stenberg *et al.*, 1992; Podbielski *et al.*, 1996; Thern *et al.*, 1998). It is not yet known whether certain strains (*e.g.* those of pattern D) may express two Fg-binding proteins, one Mrp protein and one M protein.

Binding of Fg contributes to phagocytosis resistance through inhibition of complement deposition on the bacterial surface (Whitnack and Beachey, 1982; Carlsson *et al.*, 2005; Courtney *et al.*, 2006). Interestingly, analysis of the M5 system demonstrated that complement deposition does not occur via the alternative pathway, as has often been assumed, but via the classical pathway, and M5-bound Fg was shown to interfere with this deposition via the classical pathway (Carlsson *et al.*, 2005). Complexes between M protein and Fg have also been suggested to play a role in increased vascular leakage of plasma proteins in patients with necrotizing disease (Herwald *et al.*, 2004). Moreover, a recent study indicates that Fg plays an important role in blocking binding of specific antibodies directed against M protein (Sandin *et al.*, 2006).



**Figure 8.** Binding of human plasma proteins to two representative M proteins, M22 and M5. HVR, hypervariable region; SVR, semivariable region.

*Albumin.* Many *S. pyogenes* strains bind human serum albumin (Kronvall *et al.*, 1979b). Molecular analysis demonstrated that the binding is due to the C-repeat region of M proteins

(Åkesson *et al.*, 1994; Retnoningrum and Cleary, 1994). The function of this interaction remains unclear. However, studies by Sandin *et al.* (2006) demonstrated that albumin bound to the C-repeat region of an M protein may play an important role in evasion of specific antibodies directed against that region.

*Factor H/FHL-1.* Many M proteins interact with different regulators of the human complement system (other than C4BP) (Horstmann *et al.*, 1988; Johnsson *et al.*, 1998; Lindahl *et al.*, 2000). Two regulators of the alternative pathway of complement, factor H (FH) and its splice variant factor H-like protein 1 (FHL-1) have been suggested to contribute to the phagocytosis resistance of *S. pyogenes* by protecting the bacteria from alternative pathway complement deposition (Horstmann *et al.*, 1988; Johnsson *et al.*, 1998). However, this hypothesis has not been confirmed; on the contrary, emerging data suggest that neither of these proteins (FH and FHL-1) is implied in M protein-mediated phagocytosis resistance (Perez-Casal *et al.*, 1995; Kotarsky *et al.*, 2001; Sandin *et al.*, 2006). Thus, the role of these two complement regulators in *S. pyogenes* infections remains unclear.

*Fibronectin.* Some M proteins bind human fibronectin (Frick *et al.*, 1995; Cue *et al.*, 2001), an interaction that has been implicated in invasion of epithelial cells by *S. pyogenes* (Cue *et al.*, 2001).

*IgA and IgG.* It was reported in 1973 that many *S. pyogenes* strains bind the Fc-part of human IgG (Kronvall, 1973), and shortly thereafter this was found to be the case also for IgA-Fc (Christensen and Oxelius, 1975). In 1989, two separate studies found these interactions to be M protein dependent (Frithz *et al.*, 1989; Heath and Cleary, 1989).

The binding of IgA contributes to phagocytosis resistance (Carlsson *et al.*, 2003), possibly by interfering with the binding of IgA-Fc to its cellular receptor CD89 (Pleass *et al.*, 2001). The IgA-binding region is a well-defined protein domain that can be studied in isolated form, as a synthetic peptide (Johnsson *et al.*, 1999; Sandin *et al.*, 2002).

Many M and M-like proteins are also known to bind IgG-Fc (Heath and Cleary, 1989; Lindahl and Åkerström, 1989; Gomi *et al.*, 1990; Stenberg *et al.*, 1992; Stenberg *et al.*, 1994; Åkesson *et al.*, 1994), but the function of this interaction is not clear.

*Kininogen.* Some M proteins were reported to bind human kininogen (Ben Nasr *et al.*, 1995), but the binding site has not been identified. Analysis of this interaction suggested that streptococcal cysteine protease may cleave the M protein-bound kininogen, thereby causing release of the pro-inflammatory peptide kinin, possibly inducing increased vascular permeability (Ben Nasr *et al.*, 1995; Herwald *et al.*, 1996).

*Plasminogen.* Certain strains of the OF+ lineage have been shown to express M proteins that bind human plasminogen (Berge and Sjöbring, 1993; Wistedt *et al.*, 1995). Plasminogen bound to the streptococcal surface may be activated by host activators or streptokinase to generate plasmin activity, causing degradation of extracellular matrix proteins and facilitating tissue invasion (Boyle and Lottenberg, 1997; Ringdahl *et al.*, 1998). This interaction causes an important increase in bacterial virulence (Khil *et al.*, 2003; Sun *et al.*, 2004).

#### ***Binding of human plasma proteins to Mrp and Enn proteins***

Description of the Mrp and Enn proteins is beyond the scope of this thesis. Moreover, much less is known about these proteins than about the extensively studied M proteins. However, the ability of the Mrp and Enn proteins to bind plasma proteins will be briefly summarized here, since this information underlines the similarity between these proteins and the classical M proteins and also underlines the remarkable finding that all members of the M protein family bind at least one human plasma protein.

**Mrp.** Mrp proteins have been shown to bind IgG-Fc and Fg (Heath *et al.*, 1990; Stenberg *et al.*, 1992), and to confer resistance to phagocytosis (Podbielski *et al.*, 1996; Thern *et al.*, 1998). Recently a mechanism by which Mrp proteins may confer phagocytosis-resistance was suggested. As in the case of Fg bound to M proteins (discussed above), Fg bound to the Mrp protein of a type 4 strain was shown to protect the bacterium against complement-deposition via the classical pathway (Courtney *et al.*, 2006).

**Enn.** When expressed in *E. coli*, most or all Enn proteins bind IgA-Fc (Bessen, 1994; Thern *et al.*, 1998). However, because of the low, or even absent, expression of many Enn proteins in streptococci, it is unclear whether this interaction plays a role *in vivo*. Recently, the Enn protein in a serotype 18 strain was shown to bind C4BP (Pérez-Caballero *et al.*, 2004), suggesting a possible role for that Enn protein in phagocytosis resistance. However, such a role remains to be demonstrated. Moreover, protein H, a mosaic protein expressed by some M1 strains (Åkesson *et al.*, 1994), has the same genetic location as the Enn proteins, and also binds C4BP (Johnsson *et al.*, 1996).

#### ***M protein and human C4BP***

The binding of human C4BP to *S. pyogenes* was first described by Thern *et al.* (1995), who tested a large number of strains, and found that most of them (~83%) could bind C4BP. For two strains of different M types, M4 and M22, this property was analyzed at the molecular level and found to be due to the M protein. In addition, C4BP bound to the bacteria was

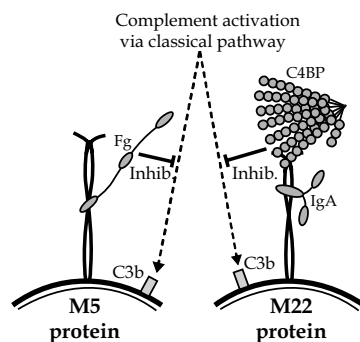
shown to retain its function as a cofactor to factor I (fI) in the degradation of C4b. Subsequent studies of the interaction between C4BP and M protein demonstrated that C4BP binds to the most N-terminal ~50 aa HVR (Johnsson *et al.*, 1996). Moreover, the binding site in C4BP was found to overlap with the binding site for C4b, the natural ligand of C4BP (Figure 2) (Accardo *et al.*, 1996; Blom *et al.*, 2000), but because of the polymeric structure of C4BP, different  $\alpha$ -chains may either bind C4b or the bacterium, probably explaining why bacteria-bound C4BP retains its cofactor function.

The HVRs of C4BP-binding M proteins exhibit extensive sequence and antigenic variability, yet molecular studies indicate that they bind to the same site in C4BP (Johnsson *et al.*, 1996; Blom *et al.*, 2000; **paper I**; **paper III**). A recent paper analyzed the structural basis for this interaction between M protein and human C4BP (Jenkins *et al.*, 2006). Previous studies had shown that CCP1-2 of the C4BP  $\alpha$ -chain were important for binding of M protein (Accardo *et al.*, 1996; Blom *et al.*, 2000) and the study by Jenkins *et al.* (2006) shows that these two CCPs are sufficient for binding. Employing heteronuclear NMR spectroscopy, the structure of CCP1-2 was solved, and the interaction with the M4 protein was studied using a peptide derived from the C4BP-binding HVR of this M protein. Interestingly, a four-residue linker connecting the CCP1-2 domains also plays an important role in binding, because the insertion of two alanine residues into this linker completely abrogated the interaction with M protein. The spectroscopic data suggested that residues in both CCPs are affected upon binding of the M protein-derived peptide. However, because these affected residues are not located on the same face of the  $\alpha$ -chain, it seemed unlikely that they were all directly involved in the binding. Rather, the data were interpreted to reflect a conformational change in C4BP, induced upon interaction with the peptide (Jenkins *et al.*, 2006). Previously, a cluster of positive residues on the C4BP CCP1-2 interface had been suggested to be important for the C4BP-M protein binding (Blom *et al.*, 2000). However, due to the salt-resistant nature of the interaction, hydrophobic interactions and hydrogen bonds are likely to contribute (Blom *et al.*, 2000; Jenkins *et al.*, 2006). Possibly, the positive cluster in C4BP participates in an important “first-step” electrostatic interaction, guiding the M protein to the right location on C4BP, while this cluster may not be involved in the binding *per se* (Jenkins *et al.*, 2006).

The fact that bacteria-bound C4BP is still functional as a cofactor in C4b degradation, suggested that C4BP-binding might play an important role in protection of the bacteria against phagocytosis, by interfering with complement deposition on the bacterial surface (Thern *et al.*, 1995). To analyze this hypothesis, Berggård *et al.* (2001) constructed a strain expressing an M22 protein with a seven residue deletion in the HVR. This strain completely lacked

C4BP-binding, and was partially deficient in its ability to grow in human blood. This result suggested that the C4BP-binding indeed contributed to M protein-mediated resistance to phagocytosis, but binding of C4BP could not alone account for the anti-phagocytic property of the M22 protein (Berggård *et al.*, 2001). However, Carlsson *et al.* (2003) demonstrated that the anti-phagocytic property of the M22 protein can be fully explained as a cooperation between two distinct ligand-binding regions in this M protein, the C4BP-binding region and the IgA-binding region. While the mechanism by which bound IgA protects against phagocytosis remains unclear, this study confirmed that bound C4BP plays a key role in protecting *S. pyogenes* against phagocytosis.

The finding that *S. pyogenes* recruits C4BP and that this interaction is important for the ability of the bacterium to resist phagocytosis was puzzling because C4BP is a regulator of the classical pathway of complement, whereas the general assumption has been that bacteria mainly activate the alternative pathway. However, analysis of complement deposition on an M-negative *S. pyogenes* strain demonstrated that the complement system was activated via the classical pathway, explaining the value of recruiting a classical pathway regulator (Carlsson *et al.* 2003). Thus, C4BP bound to *S. pyogenes* M protein protects the bacteria against attack via the classical pathway (Figure 9, right part). Even if this biological function of M protein-bound C4BP has been demonstrated only in the M22 system, it is reasonable to assume that the role of surface-bound C4BP is similar in other C4BP-binding *S. pyogenes* strains.



**Figure 9.** Unifying model for the inhibition of complement deposition by M proteins. Complement is activated via the classical pathway by *S. pyogenes*, resulting in surface deposition of C3b. However, M protein inhibits the C3b deposition by recruiting a human plasma protein. Some M proteins, such as M5, bind Fg, while other M proteins, such as M22, bind C4BP. As indicated, the M22 protein also binds human IgA, which contributes to phagocytosis resistance by an unknown mechanism (Carlsson *et al.*, 2003). Modified from Carlsson *et al.* (2005).

As mentioned above, human Fg bound to *S. pyogenes* M protein also inhibits complement deposition via the classical pathway. Because all M proteins appear to bind either C4BP or Fg, these findings have allowed the construction of a unifying model for the protection of *S. pyogenes* against complement deposition and subsequent phagocytosis (Carlsson *et al.*, 2005) (Figure 9).

Because M proteins bind multiple ligands and cross-react immunologically due to the more conserved C-terminal part, the binding properties or immunological properties of a ligand-binding domain are not easily analyzed using whole M proteins. Characterization of the interaction between C4BP and M proteins was much facilitated by the finding that the C4BP-binding HVRs of M proteins can be studied in isolated form as synthetic peptides (**paper I**). An important property of these peptides is the presence of a C-terminal cysteine residue, not present in the intact M proteins, which allowed for dimerization via a disulphide bridge. This cysteine residue was included because intact M proteins are dimeric coiled-coils and because dimerization is known to be critical for the ability of M proteins to bind ligands (Åkerström *et al.*, 1992). A peptide not containing a C-terminal cysteine showed very limited binding to C4BP, as compared to the binding of the dimerized peptides, indicating that the dimeric nature of the M protein is indeed essential for ligand binding (**paper I** and **paper II**).

Remarkably, out of all proteins in human serum, C4BP is the only one that binds to HVR-derived peptides, demonstrating that the HVRs are well-defined ligand-binding protein domains with highly specific ligand-binding properties (**paper I**). Because of this binding specificity, the HVR of C4BP-binding M proteins can be used to purify C4BP by a simple single-step procedure (**paper II**). The finding that HVRs of C4BP-binding M proteins are well-defined ligand-binding domains that can be studied in isolated form, as synthetic peptides, was an important finding that enabled much of the work presented in this thesis.

## VARIABILITY IN MICROBIAL SURFACE PROTEINS

Microorganisms constantly evolve to adapt to a changing environment. For pathogens, sequence variability in surface proteins is a common mechanism by which they adjust to survive better in their hosts. The effects of sequence variability may be diverse. A change in receptor specificity may provide the microbe with new tools for infection of a specific cell type (Baranowski *et al.*, 2001; Mayer *et al.*, 2002), or may allow the microbe to change tissue tropism (Weissman *et al.*, 2003). In some pathogens, sequence variability may even permit



the organism to switch between different species (Strömberg *et al.*, 1990; van Luenen *et al.*, 2005). However, most cases of variability probably represent so called antigenic variation, an important phenomenon that will be discussed in this section.

### **Antigenic variation**

In pathogenic microorganisms, variability in a surface structure targeted by protective immunity is commonly referred to as antigenic variation. Importantly, the term antigenic variation implies that the variability confers a selective advantage and promotes virulence. The variable surface structure is most often a protein, but in the literature the term antigenic variation is sometimes also used for changes in *e.g.* carbohydrates (Deitsch *et al.*, 1997). Here, I will only consider antigenic variation in proteins. An important aspect of antigenic variation is that the variable protein must retain an important function, while varying to avoid host immunity, otherwise this structure would probably be lost by mutation.

The significance of antigenic variation can hardly be overestimated, because it is the major obstacle in the development of new vaccines. Because of antigenic variation, efficient vaccines are not available against pathogens such as HIV, *Plasmodium falciparum* (the causative agent of malaria) or influenza A virus. Importantly, antigenic variation may arise through one of two general mechanisms, antigenic shift or antigenic drift (Borst, 1991).

**Antigenic shift** represents a drastic change in protein expression. For example, the expression of the protein may be switched on or off (phase variation), or one sequence may be replaced with another (gene conversion). In gene conversion mechanisms a gene from a silent part of the genome is copied into an expression site (Borst, 1991; Deitsch *et al.*, 1997). A classical example of a pathogen exhibiting antigenic shift is influenza A virus, in which the genes, including those encoding the major surface proteins hemagglutinin (HA) and neuraminidase (NA), may be reassembled to create a virus with new immunologic properties, which may cause a rapidly spreading influenza epidemic (Steinhauer and Skehel, 2002; Stevens *et al.*, 2004).

**Antigenic drift** represents the accumulation of point mutations or small insertions/deletions that gradually change the antigenic properties of a protein. In bacteria, antigenic variation through antigenic drift is most likely a slower process than antigenic shift, because a single small genetic change will not cause the appearance of an immunologically completely different variant. However, due to the lack of proof-reading systems in RNA viruses, and their rapid replication, viruses such as HIV may mutate very rapidly (Wilson and

Cox, 1990; Malim and Emerman, 2001; Steinhauer and Skehel, 2002). In the M protein of *S. pyogenes*, the N-terminal hypervariable region (HVR) exhibits antigenic variation that most likely has arisen through antigenic drift.

### **The hypervariable region in streptococcal M protein**

The ~50-100 aa HVR of M proteins varies extensively in sequence. The HVR is stable within a strain of *S. pyogenes*, allowing the identification of >100 different M types (Facklam *et al.*, 2002). However, the number of known M types is small, compared to the large number of possible sequence variants, suggesting that these M types have been selected because of their superior fitness. The M protein HVRs are targets for protective antibodies and exhibit antigenic variation (Lancefield, 1962; Ashbaugh *et al.*, 2000). The variability in the HVRs arises through antigenic drift, *i.e.* point mutations and small insertions or deletions, and in certain M proteins it may also arise through variation in the size of a repeat region (Fischetti, 1989). Small changes in the gene segment encoding the HVR will give rise to a new allele (subtype) of an existing M type. Definition of a completely new type requires that the first 90 bases encoding mature protein share less than 92% similarity to any previously determined M/*emm* type (<http://www.cdc.gov/ncidod/biotech/strep/doc.htm>).

Because of the extensive sequence variability it was initially suggested that the HVRs do not have a specific ligand-binding function, but inhibit phagocytosis by having a negative charge, causing repulsion of the phagocyte cell surface or IgG antibodies (Fischetti, 1983; Fischetti, 1989). However, it is now known that many HVRs do have a specific ligand-binding function, in spite of the extensive sequence variability. In this regard, M proteins may be divided into two major groups. In one group of M proteins, the HVR binds the complement regulator C4BP (Thern *et al.*, 1995; Accardo *et al.*, 1996; Johnsson *et al.*, 1996), an interaction that is important for phagocytosis resistance because M protein-bound C4BP inhibits surface-deposition of complement via the classical pathway (Berggård *et al.*, 2001; Carlsson *et al.*, 2003). The second group of M proteins do not bind C4BP. In some of these non-C4BP-binding proteins, the HVR has been found to bind the complement regulator FHL-1 (Johnsson *et al.*, 1998), but the physiological relevance of this interaction is unknown (Kotarsky *et al.*, 2001; Sandin *et al.*, 2006). It seems possible that these HVRs also interact with other, unknown ligands, explaining the predicted structural constraints on these HVRs (Li *et al.*, 2003).

Interestingly, it has been observed that mutations that occur in the gene segments encoding HVRs are almost exclusively non-synonymous, which is indicative of a strong immune pressure (Kehoe *et al.*, 1996; Li *et al.*, 2003). The selective pressure to retain a particular structure and function in the M protein HVR, which changes to avoid recognition by specific antibodies, therefore gives rise to a Darwinian evolution in which only the most fit M protein variants have survived during evolution (Phillips, 2002). However, new M types may slowly appear through a similar process.

In addition to their medical importance, M proteins are attractive model systems for studies of antigenic variation. In contrast to *e.g.* gp120 of HIV, which varies also under experimental conditions, the M protein HVR is stable within a strain and is therefore a system where the variability can be studied under more controlled conditions. Moreover, in spite of the extreme sequence variability, many HVRs bind a common ligand, C4BP. These HVRs are well-defined ligand-binding domains that can be studied in isolated form, as synthetic peptides, allowing detailed biochemical and immunological characterization.

## BIOPHYSICAL METHODS - IN BRIEF

In the last paper included in this thesis, an attempt to solve the three-dimensional structure of C4BP-binding HVRs is reported (**paper IV**). This work, performed in collaboration with biophysical chemists, gave me the opportunity to learn the basic principles of two methods that are commonly used in studies of protein structure, circular dichroism (CD) and nuclear magnetic resonance spectroscopy (NMR). In this section, brief and very simplified descriptions of these methods will be given.

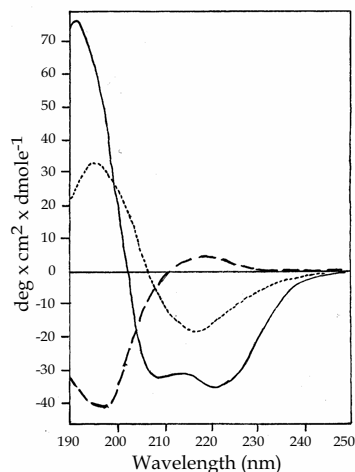
### **Circular Dichroism**

Objects that are not identical to their mirror image counterpart are referred to as being *chiral*. For example, left and right hand gloves are mirror images that differ from each other, *i.e.* the left hand glove will only fit your left hand. Such objects, as the two gloves in a pair, are referred to as *enantiomers*.

Plane polarized (or linearly polarized) light may be explained as the sum of two circularly polarized light waves, one to the left and one to the right. If plane polarized light is passed through a medium, and the effect by the medium on the left and right circular light

waves is different, the medium is considered to be *optically active*. Chiral materials are optically active. Two features may selectively affect left- and right- polarized light in an optically active medium, the first is the speed of the light wave through the medium, and the second is the extent to which the light wave is absorbed by the medium. The difference in absorbance can be measured if an optically active medium is pulsed alternately with left and right circularly polarized light, and is referred to as *circular dichroism* (CD). Importantly, the CD signal varies extensively with the wave length of the applied light, which has practical implications as discussed below.

Amino acids are chiral, and thus optically active, and the enantiomers are denoted L- or D-form amino acids. Naturally occurring amino acids are synthesized mainly as one enantiomer, the L-form. The peptide bonds in a protein absorb light very effectively in the far-UV region (wave lengths of 170-250 nm). In this UV region, the CD signal provides information about the secondary structure of the protein, but does not assign it to any given sequence. Specific structures, such as *e.g.*  $\alpha$ -helix and anti-parallel  $\beta$ -sheet, give rise to specific spectra (Figure 10), which are known from studies with model peptides and native proteins with known structure. Moreover, the far-UV region is very sensitive to loss in secondary structure and is often used in studies of protein stability.



**Figure 10.** Typical CD spectra of  $\alpha$ -helix (solid line), anti-parallel  $\beta$ -sheet (dotted line) and random coil (*i.e.* lack of secondary structure, dashed line). Adopted from Creighton (1993).

In a typical spectrum of an  $\alpha$ -helix, two distinct minima are visible, one around 208 nm and one around 222 nm (Figure 10). In an  $\alpha$ -helical coiled-coil protein, the values of the CD signals at these minima should relate to each other so that  $CD_{222}/CD_{208} > 1$  (Lau *et al.*, 1984;

Engel *et al.*, 1991). Thus, CD analysis may be used to analyze the secondary structure of a protein, and may also give some insights about tertiary structure, such as coiled-coil formation (Cantor and Schimmel, 1980; Creighton, 1993).

### **Nuclear Magnetic Resonance Spectroscopy**

The NMR technique has been awarded several Nobel prizes, one in physics (1952), two in chemistry (1991 and 2002), and one in medicine (2003). NMR is a powerful and theoretically very complex method, which is not readily explained without involving mathematical models and quantum mechanics. However, it is fully possible to appreciate some general principles of NMR without a deeper understanding of the underlying physics of the actual NMR phenomenon. NMR is a method to study nuclei of atoms in a molecule, and is the only technique available to date that can resolve the three-dimensional structure of a molecule in solution. The nuclei most commonly studied are  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ , where the easiest nucleus to use is the proton ( $^1\text{H}$ ), because of its abundance in nature.

For simplicity, the nucleus of an atom may be envisaged as a small magnet, with a north and a south pole. In an NMR experiment, these polarized nuclei are placed under the influence of a strong magnet, where they align with the applied magnetic field, and are subsequently pulsed with radio waves. The nuclei absorb energy and together change their orientation within the magnetic field, thereby creating a radiofrequency that may be measured. To assure that measurements performed on different NMR spectrometers provide the same information, the emitted frequency of a sample is compared to a reference substance. This measured difference from the reference substance is referred to as the *chemical shift* of a sample nucleus. The chemical shift is specific for a nucleus in a given environment, such that *e.g.* a proton in a methylene group ( $-\text{CH}_2$ ) next to a hydroxyl group ( $-\text{OH}$ ) has a different chemical shift than a proton in a methyl group ( $-\text{CH}_3$ ) next to a hydroxyl group, thereby providing information about the positions of the nuclei in relation to each other, which allows for the determination of a three-dimensional structure. In protein NMR, knowledge about the primary structure of the protein is necessary for an accurate model of the three-dimensional structure. The study of a single type of nucleus is known as one-dimensional (1D) NMR. However, several types of nuclei can be studied simultaneously, in 2D or multidimensional NMR (Van Holde *et al.*, 1988; Rattle, 1995; <http://www.wikipedia.org>; S. Linse and I. André, personal communication).

## PRESENT INVESTIGATION

### **PAPER I: Isolated hypervariable regions derived from streptococcal M proteins specifically bind human C4b-binding protein: implications for antigenic variation**

Pathogenic microorganisms have evolved a number of mechanisms to escape the immune system of the host. One such mechanism is antigenic variation, *i.e.* alterations in a surface structure to avoid immune recognition. The ~50-100 aa HVR of M proteins exhibit antigenic variation and has long been recognized to be targeted by specific antibodies (Fischetti, 1989). However, the function of this region in pathogenesis remained unclear. In 1995, Thern *et al.* found that many M proteins bind human C4BP, and Johnsson *et al.* (1996) attributed the interaction to the HVR in these M proteins. Subsequent studies of the biological function of this binding demonstrated that C4BP plays an important role in phagocytosis resistance (Berggård *et al.*, 2001; Carlsson *et al.*, 2003), a situation that made it of interest to further characterize this interaction. However, because streptococcal M proteins interact with multiple human plasma proteins, the nature of the M protein-C4BP interaction is not easily studied with intact M proteins. To circumvent this problem, peptides corresponding to the HVRs of three C4BP-binding proteins (M2, M4 and M22) were synthesized, in the hope that it would be possible to study the C4BP-binding HVRs in isolated form.

It had been demonstrated previously that the ability of M proteins to bind ligands requires that the M proteins have dimeric structure (Åkerström *et al.*, 1992). Therefore, the synthetic peptides contained a C-terminal cysteine not present in the native protein, to allow dimerization via a disulphide bridge. Interestingly, the dimerized peptides were able to bind C4BP, while a peptide without the C-terminal cysteine residue showed no or little binding to C4BP. Affinity chromatography experiments, in which whole human serum was applied to dimerized peptide immobilized in a column, showed that the M protein-C4BP interaction is highly specific, because out of all proteins present in serum, only C4BP was retained.

Analysis of the sequence of different C4BP-binding HVRs identified only very limited residue identity. Indeed, only three conserved residues could be identified, possibly representing a C4BP-binding motif. However, this putative motif was also found in non-C4BP-binding M proteins, indicating that the three residues were not sufficient to confer ability to bind C4BP. Further analysis of the C4BP-binding peptides suggested that, in spite of the remarkable sequence variability, these peptides bound to the same region in C4BP, at the CCP1-2 interface. Moreover, secondary structure analysis using CD, as well as computational analysis, suggested that the peptides assume similar folds.

Although the C4BP-binding peptides/HVRs display extensive sequence variability, the apparent similarities in folding and C4BP-binding properties suggested that they might possess at least some immunological cross-reactivity. However, analysis exploiting rabbit antiserum against each of the peptides demonstrated that each antiserum recognized only the corresponding peptide, indicating that the HVRs are immunologically unrelated.

In summary, C4BP-binding HVRs represent distinct ligand-binding domains that can be studied as synthetic peptides. Such peptides exhibit extensive sequence variability but have similar ligand-binding properties, suggesting that they have similar three-dimensional structure. Nevertheless several C4BP-binding peptides did not cross-react immunologically.

## **PAPER II: Single-step purification of human C4b-binding protein (C4BP) by affinity chromatography on a peptide derived from a streptococcal surface protein**

In recent years, the interest in C4BP has increased, emphasizing the value of a simple method to purify this protein (Lindahl *et al.*, 2000 ; Nonaka *et al.*, 2001; Webb *et al.*, 2002; Brodeur *et al.*, 2003; Carlsson *et al.*, 2003; Blom *et al.*, 2004). Because previously described techniques to purify C4BP are based on precipitation of human plasma and one or several subsequent chromatographic steps (Villiers *et al.*, 1981; Dahlbäck, 1983), we tried to develop a simple purification technique based on the specific interaction between C4BP and peptides derived from streptococcal M proteins.

A dimerized peptide (M22-N) corresponding to the HVR of the M22 protein was previously shown to bind human C4BP with high specificity (**paper I**), indicating that this peptide could be used to develop a simple affinity chromatography method to purify C4BP. Dimerized M22-N was immobilized in a column, and whole human serum was applied. After extensive washing, bound proteins were eluted and subjected to dialysis against a neutral buffer. The column, containing 5 mg immobilized streptococcal peptide, had a binding capacity of ~3.5 mg C4BP, and could be used repeatedly during a 1-year period with no apparent loss of function.

The purified C4BP was shown to have high molecular weight, indicating that it had the multimeric structure typical of C4BP, and was demonstrated to be more pure than a commercially available preparation of C4BP. Functional analysis demonstrated that the purified C4BP retained its activity as a cofactor to factor I in C4b-degradation. Moreover, binding assays using C4b, the natural ligand for C4BP, and streptococcal M protein, pure or expressed on the surface of whole streptococci, indicated that the binding properties of the

C4BP purified by the single-step method exceeded those of commercially available C4BP. In addition, this affinity chromatography method could be used to deplete serum of C4BP. Indeed, >99.5% of the C4BP present in a serum sample was retained on a column.

In conclusion, a synthetic peptide corresponding to the C4BP-binding region of the M22 protein can be used to purify human C4BP, and to deplete serum of C4BP, by a single-step affinity chromatography method. In addition to its interest for work on C4BP, this result emphasizes the potential of bacterial proteins/peptides for the purification of human proteins.

### **PAPER III: Extreme sequence divergence but conserved ligand-binding specificity in *Streptococcus pyogenes* M protein**

Most cases of sequence variability in surface proteins of pathogenic microorganisms, probably represent antigenic variation. An important aspect of antigenic variation is the requirement on the variable protein to retain a specific function, such as ligand-binding, while varying in sequence to escape immune recognition. It is generally assumed that ligand-binding is dependent on a specific conserved sequence motif (Shakhnovich *et al.*, 1996). Even the extremely variable surface proteins gp120 of HIV and hemagglutinin of influenza virus contain ligand-binding sites including a few highly conserved amino acid residues (Wilson and Cox, 1990; Kwong *et al.*, 1998; Skehel and Wiley, 2000).

We studied the limits of variability in C4BP-binding HVRs of M proteins. Previous analysis of five C4BP-binding HVRs revealed only three conserved residues (**paper I**), corresponding to residues L28, E31 and D40 in the extensively studied M22 protein (Figure 1B in **paper III**). Because of the limited residue identity between these five HVRs, we hypothesized that not even the three conserved residues would be essential for ability to bind C4BP.

Initial analysis was focused on the M22 protein. Studies using fusion proteins showed that residues 1-39 of M22 are sufficient for C4BP-binding. Consequently, residue D40 of M22 is not necessary for C4BP-binding, implying that only two of the three residues mentioned above might be required for C4BP-binding. To analyze whether these two residues are conserved in C4BP-binding HVRs, a large number of *S. pyogenes* strains were screened for ability to bind C4BP. Based on sequence analysis of the M protein HVRs of C4BP-binding strains, two strains, expressing the M4.1 and M114 proteins, were chosen for further studies. The HVRs of these proteins were demonstrated to represent distinct C4BP-binding

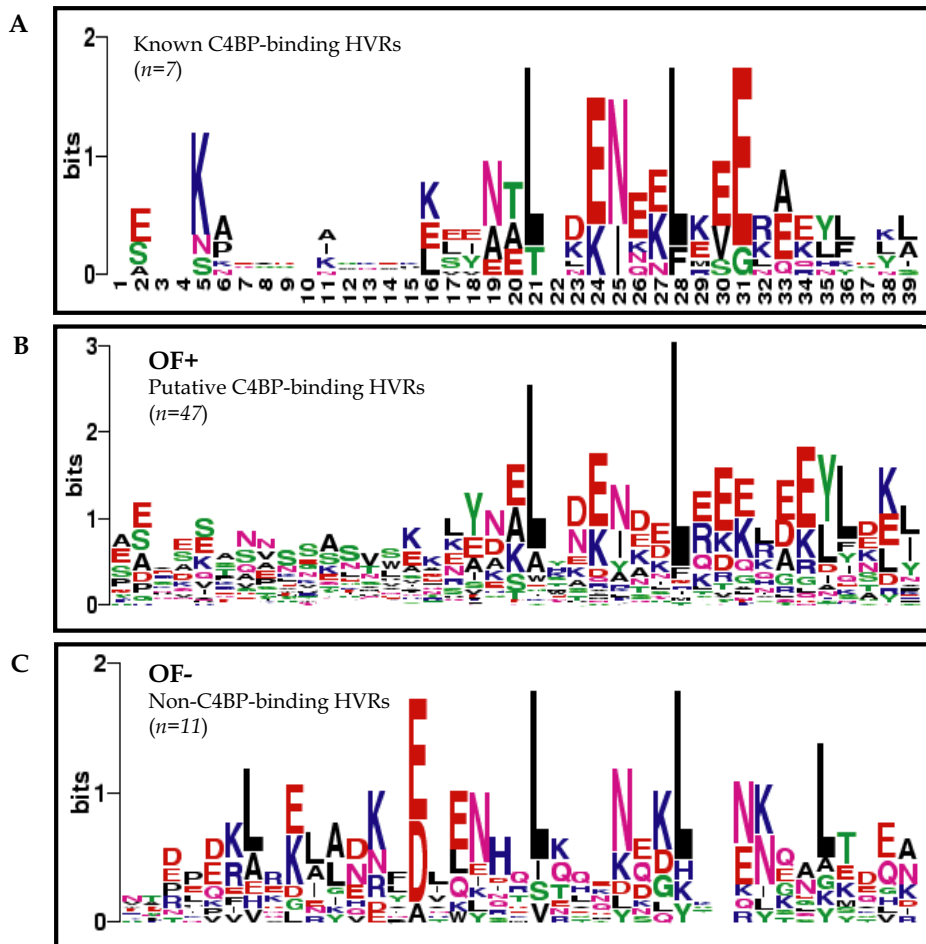


domains, and the addition of the M4.1 and M114 HVR sequences to the alignment performed in **paper I** showed that C4BP-binding HVRs completely lack residue identities.

Although different C4BP-binding HVRs lack residue identities, a sequence logo of the seven known C4BP-binding HVRs indicated that the C-terminal half is less variable (Figure 11A). This finding made it of interest to study the possible role in binding of relatively conserved residues. Four of these residues, L21, E24, L28 and E31, were substituted for alanines and the mutant proteins, containing single aa substitutions, were expressed on the surface of streptococci. Interestingly, substitution of either L21 or L28 completely abolished the ability of the M22 protein to bind C4BP. A computational prediction, and the distribution of residues in the sequence logo, suggested that the central part of the M22 HVR forms a coiled-coil and the L21 and L28 residues were predicted to represent core residues essential for the formation of this coiled-coil. Although it cannot be ruled out that these two Leu residues are directly involved in the binding of C4BP, our mutagenesis data support the conclusion that the centre of the HVR is folded as a coiled-coil (see **paper IV**). In contrast, substitution of residues E24 and E31, which are predicted to be solvent exposed in a coiled-coil, had no effect on C4BP-binding. Thus, it is not clear why these residues are relatively conserved among the C4BP-binding HVRs studied.

It was of interest to know whether the sequence logo of C4BP-binding HVRs would have similar appearance if many more HVRs were included, and whether it is different from a logo derived from non-C4BP-binding HVRs. To analyze additional C4BP-binding HVRs, we compared M protein HVRs in OF+ strains of different M types. It seemed likely that these HVRs promote binding of C4BP, because all OF+ strains bind C4BP, and because binding of C4BP has been attributed to the M protein in all OF+ strains studied at the molecular level (Thern *et al.*, 1995; Johnsson *et al.*, 1996; **paper I**; **paper III**). A logo of the M protein HVRs of 47 OF+ strains tested in **paper III** was therefore created (Figure 11B). This logo is very similar to that shown in Figure 11A. As in Figure 11A, the logo in Figure 11B is dominated by two Leu residues (corresponding to L21 and L28 in M22). In addition, Glu residues corresponding to E24 and E31 in M22 are prominent, although the conservation of these residues is less prominent than in the logo in Figure 11A. Thus, the logos in Figures 11A and 11B have similar overall appearance, with two prominent Leu residues and an abundance of Glu residues, suggesting that these features are important for C4BP-binding.

A logo derived from the HVRs of M proteins in 11 non-C4BP-binding strains (all OF-) is shown in Figure 11C. The M protein-expressing strains used to construct this logo were analyzed in **paper III**. The lack of C4BP-binding in these strains was not due to lack of M



**Figure 11.** Sequence logos of M protein HVRs. **A.** Seven known C4BP-binding HVRs (**paper III**). The shortest known C4BP-binding region in the M proteins were aligned using ClustalW, and the part of the alignment corresponding to the shortest known C4BP-binding region in M22 (residues 1-39) was used to create the logo. **B.** HVRs of M proteins in 47 OF+ strains that bind C4BP (tested in **paper III**). **C.** HVRs of M proteins in 11 non-C4BP-binding OF- strains (tested in **paper III**).

To produce the logos presented in B and C, residues 1-50 of the M proteins were aligned using ClustalW. These alignments were then manually aligned to the highly conserved Leu residues in the logo in panel A, and 39-residue regions corresponding to that in panel A were chosen to create the logos. The height of each letter is proportional to the observed frequency of the corresponding residue at that position, while the overall height of each stack is proportional to the sequence conservation at that position (Crooks *et al.*, 2004).

protein expression, because all strains bound human Fg, a feature known to be due to members of the M protein family. Therefore, in contrast to the hypothetical logo in panel B, the logo of non-C4BP-binding HVRs is based on reliable experimental data. Interestingly, this logo also displays two dominating Leu residues, strengthening the hypothesis that these residues are not involved in C4BP-binding but may be important for formation of a coiled-coil. However, the relatively conserved Glu residues in the C4BP-binding HVRs were not prominent in the non-C4BP-binding HVRs. This analysis suggest that alignments of C4BP-binding and non-C4BP-binding HVRs indeed gives rise to different logos. In particular, comparison of the different logos indicates that the C-terminal half of a C4BP-binding HVR is more negatively charged than the corresponding region in a non-binding HVR.

As mentioned above, substitution of L21 or L28 completely abolished the C4BP-binding ability of the M22 protein, while substitution of E24 or E31 did not affect C4BP-binding. This finding made it of interest to analyze whether the single residue changes, not affecting C4BP-binding, induced antigenic alterations. Remarkably, although substitution of E24 or E31 with Ala did not affect the ability of M22 to bind C4BP, these changes strongly decreased the ability of a polyclonal antiserum, directed against the HVR, to recognize the mutant proteins. This finding implies that single amino acid changes in an HVR may give rise to antigenic variation and that accumulation of such changes may eventually give rise to a new M type.

To conclude, C4BP-binding HVRs represent a unique example of sequence divergence, because they completely lack residue identities but bind the same ligand. Moreover, single residue changes in an HVR, not affecting the ability to bind C4BP, may considerably change the immunological properties of the HVR, supporting classical models for the appearance of antigenic drift through immunological selection.

#### **PAPER IV: Streptococcal M protein: structural studies of the hypervariable region, free and bound to human C4BP**

Streptococcal M proteins are known to largely assume a coiled-coil structure (Fischetti, 1989; Kehoe, 1994; Navarre and Schneewind, 1999). However, it is not known whether the coiled-coil extends through the N-terminal HVR, which has been depicted as non-structured in the literature. Because computational data and studies of secondary structure had indicated that C4BP-binding HVRs adopt similar folds, although their sequences are highly divergent (**paper I**), it was of interest to more accurately study the three-dimensional structure of these

M protein HVRs. The aim of this study was to identify the three-dimensional structure for C4BP-binding HVRs, using heteronuclear NMR spectroscopy.

Three streptococcal peptides expressed in *E. coli*. Two peptides, M4-N and M22-N, derived from the HVRs of the M4 and M22 proteins, respectively, correspond to the C4BP-binding regions in these proteins. One peptide, M22-NL, contained the HVR of M22 but also included an adjacent semivariable region that binds IgA-Fc (Stenberg *et al.*, 1994; Sandin *et al.*, 2002). Because of certain properties of the peptides, such as a tendency to aggregate at high concentrations, a three-dimensional structure could not be determined. However, using a range of techniques, we could obtain clear indications concerning the structural properties of the peptides.

Circular dichroism (CD) spectra indicated that the peptides assume a helical structure. Typically, a coiled-coil structure produces a CD spectrum with minima at 208 and 221 nm, where the ratio between the signals at these wave lengths should relate to each other as:  $208/221 > 1$  (Lau *et al.*, 1984; Engel *et al.*, 1991). This result was indeed obtained for the M22-NL peptide. However, this ratio is probably not as useful when a significant part of the molecule is in an unstructured state, as appeared to be the case for the M4-N and M22-N peptides.

For NMR analysis, the peptides were labelled with  $^{15}\text{N}$ , or  $^{13}\text{C}$  and  $^{15}\text{N}$ . These studies were focused mainly on the M4-N peptide, because preliminary NMR analysis of the peptides indicated that this peptide gave rise to superior spectral data. Results obtained from analysis of the M22-N peptide were less clear than for M4-N, but pointed in the same direction. Through the use of different NMR techniques, permitting analysis of nuclei in either the peptide backbone or the amino acid side chains, it was possible to conclude that the M4-N peptide has a structured central part of ~4 heptads, with flexible termini. The peptides were also studied in complex with CCP1-2 of the C4BP  $\alpha$ -chain. Because the signals arising from the peptide backbone are lost for interacting residues, it was possible to map the C4BP-binding region in the streptococcal peptide. The data indicated that the C4BP-binding region, possibly as many as 28 residues, is contained within the structured centre of the peptide. It should be noted that if the peptide binds as a coiled-coil, only about half of the ~28 residues will be on the binding interface.

A recent NMR study of CCP1-2 of the C4BP  $\alpha$ -chain, in complex with the unlabelled M4-N peptide indicates that the peptide mainly binds to CCP2, but that the four-residue linker connecting CCP1-2 is required for binding, and that residues in CCP1 are affected by binding of the peptide (Jenkins *et al.*, 2006). Together with the NMR analysis of M4-N reported in

**paper IV**, these data suggest that the elongated M4-N peptide may interact with both CCP1 and CCP2 of the C4BP  $\alpha$ -chain.

In conclusion, the data in **paper IV** indicate that M4-N has a folded central region of ~4 heptads with flexible termini and that the C4BP-binding surface is contained within these four heptads. Moreover, in spite of the very limited identity between M4-N and M22-N, these peptides probably adopt similar folds, both in solution and in complex with C4BP. Intriguingly, these C4BP-binding HVRs completely lack immunological cross-reactivity, although they specifically bind the same ligand and have similar structure. This apparent contradiction may be explained if C4BP mainly interacts with the backbone of the HVRs, while antibodies recognize the amino acid side chains, which vary extensively between HVRs.

## CONCLUSIONS

- The ~50 residue hypervariable regions (HVRs) of C4BP-binding streptococcal M proteins can be studied as synthetic peptides that retain their C4BP-binding ability. Different C4BP-binding HVRs show limited sequence identity, yet bind with high specificity to the same region in C4BP and seem to assume similar folds. Comparison of three such HVRs showed that they completely lacked immunological cross-reactivity, in spite of their similar properties.
- A synthetic peptide corresponding to the HVR of the M22 protein can be used to isolate C4BP from whole human serum, in a simple one-step affinity chromatography purification method. This highly purified C4BP retains its normal biological functions, including cofactor activity and ligand-binding. Immobilized peptide can also be used to efficiently deplete human serum of C4BP.
- Extended studies of C4BP-binding HVRs showed that they are extremely variable and lack any residue identities. Thus, in contrast to common assumptions that variable proteins must have a limited number of conserved residues to retain a specific function, *Streptococcus pyogenes* M protein exhibits extreme sequence divergence but retains a highly specific ligand-binding property.

The introduction of single amino acid substitutions in the HVR of the M22 protein showed that mutation of relatively conserved residues, likely to be important for the coiled-coil structure of the region, completely abolish the ability to bind C4BP. In contrast, mutation of either of two other relatively conserved residues, which probably are not implicated in structural stability, showed no effect on C4BP-binding. Interestingly, proteins containing these single aa changes were not as effectively recognized as the wild type protein by a polyclonal antiserum. Together, these findings show that HVRs of M proteins have an extraordinary capacity for sequence change and antigenic variability, while retaining a specific ligand-binding function.

- Peptides corresponding to the HVRs of the M4 and M22 proteins are folded as coiled-coils both in solution and when bound to human C4BP. The C4BP-binding surface is elongated and may cover as much as four coiled-coil heptads in the centre of the peptides. The analysis indicated that the two peptides have similar folds, in spite of their very limited sequence identity.

## SAMMANFATTNING PÅ SVENSKA

Mikroorganismer som orsakar sjukdom har ofta väl utvecklade mekanismer för att undvika att upptäckas och avdödas av människans immunförsvar. En av dessa mekanismer kallas *antigen variation* och är ett vanligt förekommande fenomen i den mikrobiella världen. Vid en infektion bildar människans immunsystem antikroppar som känner igen strukturer, t.ex. ett protein, på mikroorganismens yta. Antigen variation innebär att ett sådant protein varierar i struktur så att det inte längre känns igen av de bildade antikropparna. En viktig aspekt gällande antigen variation är dock att det varierande proteinet samtidigt måste bibehålla en viss funktion, d.v.s. variationen begränsas av proteinets nödvändiga egenskaper. Inom den medicinska mikrobiologin är antigen variation ett problematiskt fenomen som medför att det saknas effektiva vaccin mot globalt mycket viktiga sjukdomar som t. ex. AIDS, malaria och influensa.

*Streptococcus pyogenes*, även kallade grupp A streptokocker, är en vanligt förekommande sjukdomsalstrande bakterie. *S. pyogenes* ger vanligen upphov till lindriga sjukdomar såsom halsfluss och svinkoppor, men kan också verka som "mördarbakterier" och ge upphov till livshotande infektioner och allvarliga följsjukdomar. I en nyligen utkommen studie uppskattas att ca 500.000 människor om året avlider i, eller i sviterna av, infektioner med *S. pyogenes*, vilket gör att denna bakterie kan räknas till en av de tio mikroorganismer som årligen orsakar flest dödsfall i världen.

*S. pyogenes* har på sin yta det s.k. M-proteinet, vilket är viktigt för bakteriens förmåga att ge upphov till sjukdom. M-proteiner finns i mer än 100 olika varianter (M-typer; M1, M2, M3 o.s.v.), vilka definieras av en specifik del av M-proteinet, den s.k. *hypervariabla regionen* (HVR). Vid en infektion med *S. pyogenes* bildar människans immunförsvar antikroppar riktade mot M-proteinets HVR. Dessa antikroppar skyddar individen mot återkommande infektioner med streptokocker som bär denna typ av M-protein, men inte mot infektion med en streptokock som bär en annan typ av M-protein. Sålunda uppvisar HVRerna hos M-proteiner antigen variation. Trots denna variation binder många M-proteiners HVRer det humana proteinet C4BP, ett protein i människans blod vars funktion är att hämma det s.k. komplementsystemet. Komplementsystemet utgör en viktig del av vårt immunförsvar och har förmågan att leda till avdödning av invaderande mikroorganismer. C4BP bundet till bakterieytan via M-proteinet skyddar *S. pyogenes* mot angrepp från komplementsystemet. Bindningen av C4BP är därför viktig för bakteriens förmåga att överleva och orsaka sjukdom

i människa. I de fyra delarbeten som ingår i denna avhandling studeras denna för bakterien viktiga interaktion ur biokemiska, immunkemiska och strukturella perspektiv.

I två delarbeten studeras C4BP-bindande HVRers egenskaper, bl.a. med avseende på variabilitet och bindning till C4BP (delarbete I och III). I delarbete I visar vi att C4BP-bindande HVRer kan studeras i form av syntetiska peptider (d.v.s. konstgjorda kortare proteiner) som binder C4BP, vilket möjliggör specifika studier av denna interaktion. Studier av tre sådana peptider visade att dessa binder till samma yta i C4BP och att de troligen veckas likartat. Trots detta uppvisar HVRerna/peptiderna en betydande skillnad i aminosyrasekvens (aminosyror är byggstenar med vilka proteiner är sammansatta) och korsreagerar inte immunologiskt (d.v.s. antikroppar som skapats mot en HVR känner inte igen de övriga).

Många mikroorganismer uttrycker proteiner som uppvisar mycket stor variabilitet men bibehåller en specifik bindningsförmåga. Trots denna variabilitet återfinns oftast ett mönster av ett litet antal aminosyror som finns kvar i alla varianter av det varierande proteinet. Dessa konserverade aminosyror utgör sålunda ett ”motiv”, något som ofta antas vara nödvändigt för bibehållen bindning. I delarbete III studerade vi ytterligare C4BP-bindande HVRer och visade att dessa fullständigt saknar sådana konserverade aminosyror. Detta fynd innebär att *S. pyogenes* M proteiner utgör ett unikt exempel med avseende på variabilitet. Vidare visar vi att förändring av enstaka aminosyror, vilka sannolikt påverkar proteinets tredimensionella utseende, helt eliminerar bindningen till C4BP. Förändring av vissa andra aminosyror påverkade däremot inte C4BP-bindningen men hade ändå stor effekt på proteinets immunologiska egenskaper. Således kan en mikroorganism, genom en mycket begränsad förändring av en ytstruktur, undkomma de specifika antikroppar människans immunförsvar skapar.

Försök som utfördes i delarbete I antydde att en peptid som motsvarar en C4BP-bindande HVR kan användas för att enkelt rena C4BP från serum. Tidigare använda metoder för rening av C4BP är tidskrävande, och en mer lättillgänglig reningsmetod skulle vara värdefull. En sådan metod beskrivs i delarbete II. Med hjälp av en peptid motsvarande en C4BP-bindande HVR kunde funktionellt C4BP renas ur serum i ett enda steg.

I delarbete IV gjordes ett försök att bestämma konformationen hos två peptider, motsvarande HVRerna i de C4BP-bindande M proteinerna M4 och M22. En exakt konformation kunde inte fastställas, men resultaten indikerade tydligt att dessa två, sekvensmässigt skilda HVRer, antar samma tredimensionella utseende, både i lösning och bundna till C4BP. Detta fynd stöder den ovan nämnda observationen att C4BP-bindande HVRer kan veckas likartat, trots betydande skillnad i aminosyrasekvens.



## ACKNOWLEDGMENTS

I could not have done any of this without the help of a large number of people. There is no way I can mention all of you here, but I hope you all know who you are and that I am very grateful. Nevertheless, I would particularly like to thank:

*Gunnar Lindahl*, my supervisor. For excellent scientific guidance and never-ending support and encouragement. Thanks to you I feel confident and comfortable, knowing that I will always be well off bringing your ways of science with me. I would also like to thank you for giving me the time and space I much needed during the roughest of times.

The Lindahl group: *Margaretha Stålhammar-Carlemalm*, if it wasn't for you we would all spend most of our time trying to find stuff (thank you also for nice days in Vrigstad!), *Lotta Sandin*, my personal loan-shark (without you I would starve) and *Johan Waldemarsson*, the Master of Tråkiga Ordvitsar. You all make the GL lab a great place to work at.

The former members of the GL group, *Eskil Johnsson*, *Eva Morfeldt*, *Jessica Schaumburg*, *Karin Berggård*, *Thomas Areschoug* and *Ulla Regnér*. Especially, I would like to thank Eva for supervising me during my first period in the lab, and Ulla for the outrageous number of binding-tests you helped me run. Thank you!

My co-authors and co-workers *Sara Linse* and *Ingemar André*, who taught me just about everything I know about protein chemistry and biophysics. Thank you for hours of explanations (I'm quite sure I still don't really get all of it...) and for being so very patient with me (that, of course, does not include you, Ingemar ☺).

*Susanna Cardell*, my mentor. I don't know if we ever figured what the mentor-system was (or is), but at least we tried. Thank you for lots of good advice and many laughs.

*Martin Stenström*, for being there when I needed it most. I am deeply in debt to you and I can only hope that you know how much you mean to me.

My partners in crime...The PK! *Anna Tranberg, Ellen Hägerdal, Emma Johansson-Frigyesi, Karin Gustavsson* and *Malin Späth*. For being extraordinarily good friends with which I am privileged to have shared years of tea-drinking (among many other things), and for plain and simple being there through everything. I love you! And thank you *Attila, Robbie, Seb, Stu* and *Felix* for being such great guys and for making my friends happy.

*Anja Morell, Anna Kristoffersson* and *Ulrika Emanuelsson* for all the hard work you do (or did...) to keeping the choir running, and for being great friends! Anja, thanks to you, choir music and singing is an important part of my life, how can I ever repay you!?!

My old stable-pals, *Catti Ogén* and *Jenny Mattisson*. We have shared many, many bizarre moments, on and off the horseback. Even though we don't keep in such close touch as before, you are always on my mind and you mean a whole lot to me!

The people closest to me, my darling nutcase of a father *Bosse* (the best dad in the world!) and my "step-mom" ☺ – the world's sweetest *Cilla*, my beloved siblings *Johan, Stefi* and *Joachim* (my biggest idols) and the rest of the Östlund families, *Bengt, Malin, Axel, Cecilia, Marika, Mikaela* och *Tove*. Your unconditional love and support carries me. I am proud to call you all family!

*Lilian Larsson*, my mother and the bravest woman I ever knew. I wish you could have been here to share this with me.

*Fredric*. Baby. You have my heart.

## REFERENCES

- Accardo, P., Sánchez-Corral, P., Criado, O., García, E., and Rodríguez de Córdoba, S. (1996) Binding of human complement component C4b-binding protein (C4BP) to *Streptococcus pyogenes* involves the C4b-binding site. *J Immunol* **157**: 4935-4939.
- Åkerström, B., Lindahl, G., Björck, L., and Lindqvist, A. (1992) Protein Arp and protein H from group A streptococci. Ig binding and dimerization are regulated by temperature. *J Immunol* **148**: 3238-3243.
- Åkesson, P., Schmidt, K.H., Cooney, J., and Björck, L. (1994) M1 protein and protein H: IgGFc- and albumin-binding streptococcal surface proteins encoded by adjacent genes. *Biochem J* **300**: 877-886.
- Ashbaugh, C.D., Moser, T.J., Shearer, M.H., White, G.L., Kennedy, R.C., and Wessels, M.R. (2000) Bacterial determinants of persistent throat colonization and the associated immune response in a primate model of human group A streptococcal pharyngeal infection. *Cell Microbiol* **2**: 283-292.
- Banks, D.J., Porcella, S.F., Barbian, K.D., Beres, S.B., Philips, L.E., Voyich, J.M., DeLeo, F.R., Martin, J.M., Somerville, G.A., and Musser, J.M. (2004) Progress toward characterization of the group A *Streptococcus* metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. *J Infect Dis* **190**: 727-738.
- Baranowski, E., Ruiz-Jarabo, C.M., and Domingo, E. (2001) Evolution of cell recognition by viruses. *Science* **292**: 1102-1105.
- Barnes, M.G., and Weiss, A.A. (2003) Activation of the complement cascade by *Bordetella pertussis*. *FEMS Microbiol Lett* **220**: 271-275.
- Barnett, T.C., and Scott, J.R. (2002) Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J Bacteriol* **184**: 2181-2191.
- Barnum, S.R., and Dahlbäck, B. (1990) C4b-binding protein, a regulatory component of the classical pathway of complement, is an acute-phase protein and is elevated in systemic lupus erythematosus. *Complement Inflamm* **7**: 71-77.
- Beall, B., Facklam, R., and Thompson, T. (1996) Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. *J Clin Microbiol* **34**: 953-958.
- Beall, B., Facklam, R., Hoenes, T., and Schwartz, B. (1997) Survey of emm gene sequences and T-antigen types from systemic *Streptococcus pyogenes* infection isolates collected in San Francisco, California; Atlanta, Georgia; and Connecticut in 1994 and 1995. *J Clin Microbiol* **35**: 1231-1235.

- Ben Nasr, A., Herwald, H., Müller-Esterl, W., and Björck, L. (1995) Human kininogens interact with M protein, a bacterial surface protein and virulence determinant. *Biochem J* **305**: 173-180.
- Beres, S.B., Sylva, G.L., Barbian, K.D., Lei, B., Hoff, J.S., Mammarella, N.D., Liu, M.-Y., Smoot, J.C., Porcella, S.F., Parkins, L.D., Campbell, D.S., Smith, T.M., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M., and Musser, J.M. (2002) Genome sequence of a serotype M3 strain of group A Streptococcus: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. *PNAS* **99**: 10078-10083.
- Berge, A., and Sjöbring, U. (1993) PAM, a novel plasminogen-binding protein from *Streptococcus pyogenes*. *J Biol Chem* **268**: 25417-25424.
- Berggård, K., Johnsson, E., Mooi, F.R., and Lindahl, G. (1997) *Bordetella pertussis* binds the human complement regulator C4BP: role of filamentous hemagglutinin. *Infect Immun* **65**: 3638-3643.
- Berggård, K., Johnsson, E., Morfeldt, E., Persson, J., Stålhammar-Carlemalm, M., and Lindahl, G. (2001) Binding of human C4BP to the hypervariable region of M protein: a molecular mechanism of phagocytosis resistance in *Streptococcus pyogenes*. *Mol Microbiol* **42**: 539-551.
- Bessen, D., Jones, K.F., and Fischetti, V.A. (1989) Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. *J Exp Med* **169**: 269-283.
- Bessen, D.E. (1994) Localization of immunoglobulin A-binding sites within M or M-like proteins of group A streptococci. *Infect Immun* **62**: 1968-1974.
- Bessen, D.E., and Hollingshead, S.K. (1994) Allelic polymorphism of *emm* loci provides evidence for horizontal gene spread in group A streptococci. *Proc Natl Acad Sci USA* **91**: 3280-3284.
- Bessen, D.E., Sotir, C.M., Readdy, T.L., and Hollingshead, S.K. (1996) Genetic correlates of throat and skin isolates of group A streptococci. *J Infect Dis* **173**: 896-900.
- Bessen, D.E., Carapetis, J.R., Beall, B., Katz, R., Hibble, M., Currie, B.J., Collingridge, T., Izzo, M.W., Scaramuzzino, D.A., and Sriprakash, K.S. (2000) Contrasting molecular epidemiology of group A streptococci causing tropical and nontropical infections of the skin and throat. *J Infect Dis* **182**: 1109-1116.
- Bisno, A.L., and Stevens, D.L. (1996) Streptococcal infections of skin and soft tissues. *N Engl J Med* **334**: 240-245.
- Bisno, A.L. (2001) Acute pharyngitis. *N Engl J Med* **344**: 205-211.

- Bisno, A.L., Brito, M.O., and Collins, C.M. (2003) Molecular basis of group A streptococcal virulence. *Lancet Infect Dis* **3**: 191-200.
- Blom, A.M., Webb, J., Villoutreix, B.O., and Dahlbäck, B. (1999) A cluster of positively charged amino acids in the C4BP  $\alpha$ -chain is crucial for C4b binding and factor I cofactor function. *J Biol Chem* **274**: 19237-19245.
- Blom, A.M., Berggård, K., Webb, J.H., Lindahl, G., Villoutreix, B.O., and Dahlbäck, B. (2000) Human C4b-binding protein has overlapping, but not identical, binding sites for C4b and streptococcal M proteins. *J Immunol* **164**: 5328-5336.
- Blom, A.M., Rytönen, A., Vasquez, P., Lindahl, G., Dahlbäck, B., and Jonsson, A.-B. (2001) A novel interaction between type IV pili of *Neisseria gonorrhoeae* and the human complement regulator C4b-binding protein. *J Immunol* **166**: 6764-6770.
- Blom, A.M., Kask, L., and Dahlback, B. (2003) CCP1-4 of the C4b-binding protein alpha-chain are required for factor I mediated cleavage of complement factor C3b. *Mol Immunol* **39**: 547-556.
- Blom, A.M., Villoutreix, B.O., and Dahlbäck, B. (2004) Complement inhibitor C4b-binding protein - friend or foe in the innate immune system? *Mol Immunol* **40**: 1333-1346.
- Borst, P. (1991) Molecular genetics of antigenic variation. *Immunol Today* **12**: A29-A33.
- Boyle, M.D., and Lottenberg, R. (1997) Plasminogen activation by invasive human pathogens. *Thromb Haemost* **77**: 1-10.
- Brodeur, S.R., Angelini, F., Bacharier, L.B., Blom, A.M., Mizoguchi, E., Fujiwara, H., Plebani, A., Notarangelo, L.D., Dahlbäck, B., Tsitsikov, E., and Geha, R.S. (2003) C4b-binding protein (C4BP) activates B cells through the CD40 receptor. *Immunity* **18**: 837-848.
- Brown, E.J., and Gresham, H.D. (2003) Phagocytosis. In *Fundamental immunology*. Paul, W.E. (ed). New York: Lippincott Williams and Wilkins, pp. 1105-1150.
- Brown, J.S., Hussell, T., Gilliland, S.M., Holden, D.W., Paton, J.C., Ehrenstein, M.R., Walport, M.J., and Botto, M. (2002) The classical pathway is the dominant complement pathway required for innate immunity to *Streptococcus pneumoniae* infection in mice. *Proc Natl Acad Sci U S A* **99**: 16969-16974.
- Cantor, C.R., and Schimmel, P.R. (1980) *Biophysical Chemistry; Part II*. San Francisco: Freeman & Co.
- Carapetis, J.R., McDonald, M., and Wilson, N.J. (2005a) Acute rheumatic fever. *Lancet* **366**: 155-168.

- Carapetis, J.R., Steer, A.C., Mulholland, E.K., and Weber, M. (2005b) The global burden of group A streptococcal diseases. *Lancet Infect Dis* **5**: 685-694.
- Carlsson, F., Berggård, K., Stålhammar-Carlemalm, M., and Lindahl, G. (2003) Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. *J Exp Med* **198**: 1057-1068.
- Carlsson, F. (2005) Resistance to phagocytosis in *Streptococcus pyogenes*. Role of surface M protein. *Doctoral thesis*. Lund University, pp. 48.
- Carlsson, F., Sandin, C., and Lindahl, G. (2005) Human fibrinogen bound to *Streptococcus pyogenes* M protein inhibits complement deposition via the classical pathway. *Mol Microbiol* **56**: 28-39.
- Carroll, M.C. (2004) The complement system in regulation of adaptive immunity. *Nat Immunol* **5**: 981-986.
- Cedervall, T., Johansson, M.U., and Åkerström, B. (1997) Coiled-coil structure of group A streptococcal M proteins. Different temperature stability of class A and C proteins by hydrophobic- nonhydrophobic amino acid substitutions at heptad positions a and d. *Biochemistry* **36**: 4987-4994.
- Christensen, P., and Oxelius, V.-A. (1975) A reaction between some streptococci and IgA myeloma proteins. *APMIS* **83**: 184-188.
- Collin, M., and Olsen, A. (2001a) Effect of SpeB and EndoS from *Streptococcus pyogenes* on human immunoglobulins. *Infect Immun* **69**: 7187-7189.
- Collin, M., and Olsen, A. (2001b) EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J* **20**: 3046-3055.
- Courtney, H.S., von Hunolstein, C., Dale, J.B., Bronze, M.S., Beachey, E.H., and Hasty, D.L. (1992) Lipoteichoic acid and M protein: dual adhesins of group A streptococci. *Microb Pathog* **12**: 199-208.
- Courtney, H.S., Hasty, D.L., and Dale, J.B. (2002) Molecular mechanisms of adhesion, colonization, and invasion of group A streptococci. *Ann Med* **34**: 77-87.
- Courtney, H.S., Hasty, D.L., and Dale, J.B. (2006) Anti-phagocytic mechanisms of *Streptococcus pyogenes*: binding of fibrinogen to M-related protein. *Mol Microbiol* **59**: 936-947.
- Creighton, T.E. (1993) *Proteins: structures and molecular properties*. New York: W. H. Freeman and Company.

- Criado García, O., Sánchez-Corral, P., and Rodríguez de Córdoba, S. (1995) Isoforms of human C4b-binding protein. II. Differential modulation of the *C4BPA* and *C4BPB* genes by acute phase cytokines. *J Immunol* **155**: 4037-4043.
- Crick, F.H.C. (1953) The packing of  $\alpha$ -helices: Simple coiled-coils. *Acta Crystallogr* **6**: 689-697.
- Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.
- Cue, D., Lam, H., and Cleary, P.P. (2001) Genetic dissection of the *Streptococcus pyogenes* M1 protein: regions involved in fibronectin binding and intracellular invasion. *Microb Pathog* **31**: 231-242.
- Cunningham, M.W. (2000) Pathogenesis of group A streptococcal infections. *Clin Microbiol Rev* **13**: 470-511.
- Cywes, C., and Wessels, M.R. (2001) Group A Streptococcus tissue invasion by CD44-mediated cell signalling. *Nature* **414**: 648-652.
- Dahlbäck, B., and Stenflo, J. (1981) High molecular weight complex in human plasma between vitamin K-dependent protein S and complement component C4b-binding protein. *Proc Natl Acad Sci USA* **78**: 2512-2516.
- Dahlbäck, B. (1983) Purification of human C4b-binding protein and formation of its complex with vitamin K-dependent protein S. *Biochem J* **209**: 847-856.
- Dahlbäck, B., Smith, C.A., and Müller-Eberhard, H.J. (1983) Visualization of human C4b-binding protein and its complexes with vitamin K-dependent protein S and complement protein C4b. *Proc Natl Acad Sci USA* **80**: 3461-3465.
- Dahlbäck, B. (1991) Protein S and C4b-binding protein: components involved in the regulation of the protein C anticoagulant system. *Thromb Haemost* **66**: 49-61.
- Deitsch, K.W., Moxon, E.R., and Wellems, T.E. (1997) Shared themes of antigenic variation and virulence in bacterial, protozoal, and fungal infections. *Microbiol Mol Biol Rev* **61**: 281-293.
- Engel, M., Williams, R.W., and Erickson, B.W. (1991) Designed coiled-coil proteins: synthesis and spectroscopy of two 78-residue alpha-helical dimers. *Biochemistry* **30**: 3161-3169.
- Esparza-Gordillo, J., Soria, J.M., Buil, A., Souto, J.C., Almasy, L., Blangero, J., de Córdoba, S.R., and Fontcuberta, J. (2004) Genetic correlation between plasma levels of C4BP isoforms containing  $\beta$  chains and susceptibility to thrombosis. *J Med Genet* **41**: e5.

- Facklam, R., Beall, B., Efstratiou, A., Fischetti, V., Johnson, D., Kaplan, E., Kriz, P., Lovgren, M., Martin, D., Schwartz, B., Totolian, A., Bessen, D., Hollingshead, S., Rubin, F., Scott, J., and Tyrrell, G. (1999) *emm* typing and validation of provisional M types for group A streptococci. *Emerg Infect Dis* **5**: 247-253.
- Facklam, R., Beall, B., Efstratiou, A., Kriz, P., Tyrrell, G., Lovgren, M., Martin, D., and Kaplan, E. (2000) Nomenclature of new *emm* types of group A streptococci. In *Streptococci and streptococcal diseases*. Martin, D.R. and Tagg, J.R. (eds). Securacopy, Porirua, New Zealand.: Proc. XIV Lancefield Intl. Symp., pp. 797-803.
- Facklam, R.F., Martin, D.R., Lovgren, M., Johnson, D.R., Efstratiou, A., Thompson, T.A., Gowan, S., Kriz, P., Tyrrell, G.J., Kaplan, E., and Beall, B. (2002) Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: *emm*103 to *emm*124. *Clin Infect Dis* **34**: 28-38.
- Ferguson, J.S., Weis, J.J., Martin, J.L., and Schlesinger, L.S. (2004) Complement protein C3 binding to *Mycobacterium tuberculosis* is initiated by the classical pathway in human bronchoalveolar lavage fluid. *Infect Immun* **72**: 2564-2573.
- Ferretti, J.J., McShan, W.M., Ajdic, D., Savic, D.J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A.N., Kenton, S., Lai, H.S., Lin, S.P., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A., and McLaughlin, R. (2001) Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci U S A* **98**: 4658-4663.
- Fischetti, V. (1983) Requirements for the opsonic activity of human IgG directed to type 6 group A streptococci: net basic charge and intact Fc region. *J Immunol* **130**: 896-902.
- Fischetti, V.A. (1989) Streptococcal M protein: molecular design and biological behavior. *Clin Microbiol Rev* **2**: 285-314.
- Frick, I.M., Crossin, K.L., Edelman, G.M., and Björck, L. (1995) Protein H - a bacterial surface protein with affinity for both immunoglobulin and fibronectin type III domains. *EMBO J* **14**: 1674-1679.
- Frithz, E., Hedén, L.-O., and Lindahl, G. (1989) Extensive sequence homology between IgA receptor and M proteins in *Streptococcus pyogenes*. *Mol Microbiol* **3**: 1111-1119.
- Fujita, T. (2002) Evolution of the lectin-complement pathway and its role in innate immunity. *Nat Rev Immunol* **2**: 346-353.
- Gadjeva, M., Thiel, S., and Jensenius, J.C. (2001) The mannan-binding-lectin pathway of the innate immune response. *Curr Opin Immunol* **13**: 74-78.



- Garcia de Frutos, P., Alim, R., Härdig, Y., Zoller, B., and Dahlbäck, B. (1994) Differential regulation of alpha and beta chains of C4b-binding protein during acute-phase response resulting in stable plasma levels of free anticoagulant protein S. *Blood* **84**: 815-822.
- Gewurz, H., Ying, S.-C., Jiang, H., and Lint, T.F. (1993) Nonimmune activation of the classical complement pathway. *Behring Inst Mitt* **93**: 138-147.
- Gomi, H., Hozumi, T., Hattori, S., Tagawa, C., Kishimoto, F., and Björck, L. (1990) The gene sequence and some properties of protein H. A novel IgG-binding protein. *J Immunol* **144**: 4046-4052.
- Goeder, H. (1961) Association of a serum opacity reaction with serological type in *Streptococcus pyogenes*. *J Gen Microbiol* **25**: 347-352.
- Green, N.M., Zhang, S., Porcella, S.F., Nagiec, M.J., Barbian, K.D., Beres, S.B., LeFebvre, R.B., and Musser, J.M. (2005) Genome sequence of a serotype M28 strain of group A streptococcus: potential new insights into puerperal sepsis and bacterial disease specificity. *J Infect Dis* **192**: 760-770.
- Griffin, J.H., Gruber, A., and Fernandez, J.A. (1992) Reevaluation of total, free, and bound protein S and C4b-binding protein levels in plasma anticoagulated with citrate or hirudin. *Blood* **79**: 3203-3211.
- Haanes, E.J., and Cleary, P.P. (1989) Identification of a divergent M protein gene and an M protein-related gene family in *Streptococcus pyogenes* serotype 49. *J Bacteriol* **171**: 6397-6408.
- Hanski, E., and Caparon, M. (1992) Protein F, a fibronectin-binding protein, is an adhesin of the group A streptococcus *Streptococcus pyogenes*. *Proc Natl Acad Sci USA* **89**: 6172-6176.
- Harrison, D. (2006) Smittorna som dödat miljoner. *Populär Historia* **1**: 25-28.
- Heath, D.G., and Cleary, P.P. (1989) Fc-receptor and M-protein genes of group A streptococci are products of gene duplication. *Proc Natl Acad Sci USA* **86**: 4741-4745.
- Heath, D.G., Boyle, M.D.P., and Cleary, P.P. (1990) Isolated DNA repeat region from *fcrA76*, the Fc-binding protein gene from an M-type 76 strain of group A streptococci, encodes a protein with Fc-binding activity. *Mol Microbiol* **4**: 2071-2079.
- Hedén, L.-O., and Lindahl, G. (1993) Conserved and variable regions in protein Arp, the IgA receptor of *Streptococcus pyogenes*. *J Gen Microbiol* **139**: 2067-2074.
- Herwald, H., Collin, M., Müller-Esterl, W., and Björck, L. (1996) Streptococcal cysteine proteinase releases kinins: a novel virulence mechanism. *J Exp Med* **184**: 665-673.

- Herwald, H., Cramer, H., Mörgelin, M., Russell, W., Sollenberg, U., Norrby-Teglund, A., Flodgaard, H., Lindbom, L., and Björck, L. (2004) M protein, a classical bacterial virulence determinant, forms complexes with fibrinogen that induce vascular leakage. *Cell* **116**: 367-379.
- Hessing, M., Vlooswijk, R.A.A., Hackeng, T.M., Kanters, D., and Bouma, B.N. (1990) The localization of heparin-binding fragments on human C4b-binding protein. *J Immunol* **144**: 204-208.
- Hillarp, A., Hessing, M., and Dahlbäck, B. (1989) Protein S binding in relation to the subunit composition of human C4b-binding protein. *FEBS Lett* **259**: 53-56.
- Hollingshead, S.K., Fischetti, V.A., and Scott, J.R. (1986) Complete nucleotide sequence of type 6 M protein of the group A streptococcus. Repetitive structure and membrane anchor. *J Biol Chem* **261**: 1677-1686.
- Hollingshead, S.K., Readdy, T.L., Yung, D.L., and Bessen, D.E. (1993) Structural heterogeneity of the *emm* gene cluster in group A streptococci. *Mol Microbiol* **8**: 707-717.
- Holmskov, U., Thiel, S., and Jensenius, J.C. (2003) Collections and ficolins: humoral lectins of the innate immune defense. *Annu Rev Immunol* **21**: 547-578.
- Horstmann, R.D., Sievertsen, H.J., Knobloch, J., and Fischetti, V.A. (1988) Antiphagocytic activity of streptococcal M protein: selective binding of complement control protein factor H. *Proc Natl Acad Sci USA* **85**: 1657-1661.
- Hu, J.C., O'Shea, E.K., Kim, P.S., and Sauer, R.T. (1990) Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* **250**: 1400-1403.
- Hu, J.C., Newell, N.E., Tidor, B., and Sauer, R.T. (1993) Probing the roles of residues at the e and g positions of the GCN4 leucine zipper by combinatorial mutagenesis. *Protein Sci* **2**: 1072-1084.
- Ikeda, K., Sannoh, T., Kawasaki, N., Kawasaki, T., and Yamashina, I. (1987) Serum lectin with known structure activates complement through the classical pathway. *J Biol Chem* **262**: 7451-7454.
- Jacks-Weis, J., Kim, Y., and Cleary, P.P. (1982) Restricted deposition of C3 on M<sup>+</sup> group A streptococci: correlation with resistance to phagocytosis. *J Immunol* **128**: 1897-1902.
- Jarva, H., Ram, S., Vogel, U., Blom, A.M., and Meri, S. (2005) Binding of the complement inhibitor C4bp to serogroup B *Neisseria meningitidis*. *J Immunol* **174**: 6299-6307.

- Jenkins, H.T., Mark, L., Ball, G., Persson, J., Lindahl, G., Uhrin, D., Blom, A.M., and Barlow, P.N. (2006) Human C4b-binding protein, structural basis for interaction with streptococcal M protein, a major bacterial virulence factor. *J Biol Chem* **281**: 3690-3697.
- Jeppson, H., Frithz, E., and Hedén, L.-O. (1992) Duplication of a DNA sequence homologous to genes for immunoglobulin receptors and M proteins in *Streptococcus pyogenes*. *FEMS Microbiol Lett* **71**: 139-145.
- Johnson, D.R., Kaplan, E.L., VanGheem, A., Facklam, R.R., and Beall, B. (2006) Characterization of group A streptococci (*Streptococcus pyogenes*): correlation of M-protein and *emm*-gene type with T-protein agglutination pattern and serum opacity factor. *J Med Microbiol* **55**: 157-164.
- Johnsson, E., Thern, A., Dahlbäck, B., Hedén, L.O., Wikström, M., and Lindahl, G. (1996) A highly variable region in members of the streptococcal M protein family binds the human complement regulator C4BP. *J Immunol* **157**: 3021-3029.
- Johnsson, E., Berggård, K., Kotarsky, H., Hellwage, J., Zipfel, P.F., Sjöbring, U., and Lindahl, G. (1998) Role of the hypervariable region in streptococcal M proteins: binding of a human complement inhibitor. *J Immunol* **161**: 4894-4901.
- Johnsson, E. (1999) Studies on streptococcal M proteins. Interactions with IgA and human complement regulators. *Doctoral thesis*. Lund University, pp. 128.
- Johnsson, E., Areschoug, T., Mestecky, J., and Lindahl, G. (1999) An IgA-binding peptide derived from a streptococcal surface protein. *J Biol Chem* **274**: 14521-14524.
- Kantor, F.S. (1965) Fibrinogen precipitation by streptococcal M protein. I. Identity of the reactants, and stoichiometry of the reaction. *J Exp Med* **121**: 849-859.
- Kask, L., Trouw, L.A., Dahlbäck, B., and Blom, A.M. (2004) The C4b-binding protein-protein S complex inhibits the phagocytosis of apoptotic cells. *J Biol Chem* **279**: 23869-23873.
- Kehoe, M.A. (1994) Cell-wall-associated proteins in Gram-positive bacteria. In *Bacterial cell wall*. *New Comp Biochem*. Vol. 27. Ghuysen, J.-M. and Hakenbeck, R. (eds). Amsterdam: Elsevier Science B.V., pp. 217.
- Kehoe, M.A., Kapur, V., Whatmore, A.M., and Musser, J.M. (1996) Horizontal gene transfer among group A streptococci: implications for pathogenesis and epidemiology. *Trends Microbiol* **4**: 436-443.

- Kemper, C., Chan, A.C., Green, J.M., Brett, K.A., Murphy, K.M., and Atkinson, J.P. (2003) Activation of human CD4+ cells with CD3 and CD46 induces a T-regulatory cell 1 phenotype. *Nature* **421**: 388-392.
- Khil, J., Im, M., Heath, A., Ringdahl, U., Mundada, L., Cary Engleberg, N., and Fay, W.P. (2003) Plasminogen enhances virulence of group A streptococci by streptokinase-dependent and streptokinase-independent mechanisms. *J Infect Dis* **188**: 497-505.
- Kotarsky, H., Gustafsson, M., Svensson, H.G., Zipfel, P.F., Truedsson, L., and Sjöbring, U. (2001) Group A streptococcal phagocytosis resistance is independent of complement factor H and factor H-like protein 1 binding. *Mol Microbiol* **41**: 817-826.
- Kotb, M., Norrby-Teglund, A., McGeer, A., El-Sherbini, H., Dorak, M.T., Khurshid, A., Green, K., Peeples, J., Wade, J., Thomson, G., Schwartz, B., and Low, D.E. (2002) An immunogenetic and molecular basis for differences in outcomes of invasive group A streptococcal infections. *Nat Med* **8**: 1398-1404.
- Kreikemeyer, B., McIver, K.S., and Podbielski, A. (2003) Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. *Trends in Microbiology* **11**: 224-232.
- Kronvall, G. (1973) A surface component in group A, C, and G streptococci with non-immune reactivity for immunoglobulin G. *J Immunol* **111**: 1401-1406.
- Kronvall, G., Schönbeck, C., and Myhre, E. (1979a) Fibrinogen binding structures in beta-hemolytic streptococci group A, C, and G. Comparisons with receptors for IgG and aggregated beta 2-microglobulin. *Acta Pathol Microbiol Scand* **87**: 303-310.
- Kronvall, G., Simmons, A., Myhre, E.B., and Jonsson, S. (1979b) Specific absorption of human serum albumin, immunoglobulin A, and immunoglobulin G with selected strains of group A and G streptococci. *Infect Immun* **25**: 1-10.
- Kumar, V., Ali, S.R., Konrad, S., Zwirner, J., Verbeek, J.S., Schmidt, R.E., and Gessner, J.E. (2006) Cell-derived anaphylatoxins as key mediators of antibody-dependent type II autoimmunity in mice. *J Clin Invest* **116**: 512-520.
- Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**: 648-659.
- Lancefield, R.C. (1928) The antigenic complex of *Streptococcus haemolyticus*. I. Demonstration of a type-specific substance in extracts of *Streptococcus haemolyticus*. *J Exp Med* **47**: 91-103.

- Lancefield, R.C. (1933) A serological differentiation of human and other groups of hemolytic streptococci. *J Exp Med* **57**: 571-595.
- Lancefield, R.C., and Hare, R. (1935) The serological differentiation of pathogenic and non-pathogenic strains of hemolytic streptococci of parturient women. *J Exp Med* **61**: 335-349.
- Lancefield, R.C. (1962) Current knowledge of type-specific M antigens of group A streptococci. *J Immunol* **89**: 307-313.
- Lau, S.Y.M., Taneja, A.K., and Hodges, R.S. (1984) Synthesis of a model protein of defined secondary and quaternary structure - effect of chain-length on the stabilization and formation of 2-stranded alpha-helical coiled-coils. *J Biol Chem* **259**: 3253-3261.
- Lei, B., DeLeo, F.R., Hoe, N.P., Graham, M.R., Mackie, S.M., Cole, R.L., Liu, M., Hill, H.R., Low, D.E., Federle, M.J., Scott, J.R., and Musser, J.M. (2001) Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. *Nat Med* **7**: 1298-1305.
- Li, Z., Sakota, V., Jackson, D., Franklin, A.R., and Beall, B. (2003) Array of M protein gene subtypes in 1064 recent invasive group A streptococcus isolates recovered from the active bacterial core surveillance. *J Infect Dis* **188**: 1587-1592.
- Lindahl, G., and Åkerström, B. (1989) Receptor for IgA in group A streptococci: cloning of the gene and characterization of the protein expressed in *Escherichia coli*. *Mol Microbiol* **3**: 239-247.
- Lindahl, G., Sjöbring, U., and Johnsson, E. (2000) Human complement regulators: a major target for pathogenic microorganisms. *Curr Opin Immunol* **12**: 44-51.
- Liszewski, M.K., Farries, T.C., Lublin, D.M., Rooney, I.A., and Atkinson, J.P. (1996) Control of the complement system. *Adv Immunol* **61**: 201-283.
- Lukomski, S., Nakashima, K., Abdi, I., Cipriano, V.J., Shelvin, B.J., Graviss, E.A., and Musser, J.M. (2001) Identification and characterization of a second extracellular collagen-like protein made by group A streptococcus: control of production at the level of translation. *Infect Immun* **69**: 1729-1738.
- Lupas, A.N., and Gruber, M. (2005) The structure of alpha-helical coiled coils. *Adv Protein Chem* **70**: 37-78.
- Lynch, N.J., Roscher, S., Hartung, T., Morath, S., Matsushita, M., Maennel, D.N., Kuraya, M., Fujita, T., and Schwaible, W.J. (2004) L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement. *J Immunol* **172**: 1198-1202.

- Malim, M.H., and Emerman, M. (2001) HIV-1 sequence variation: drift, shift, and attenuation. *Cell* **104**: 469-472.
- Manderson, A.P., Pickering, M.C., Botto, M., Walport, M.J., and Parish, C.R. (2001) Continual low-level activation of the classical complement pathway. *J Exp Med* **194**: 747-756.
- Manjula, B.N., and Fischetti, V.A. (1980) Tropomyosin-like seven residue periodicity in three immunologically distinct streptococcal M proteins and its implications for the antiphagocytic property of the molecule. *J Exp Med* **151**: 695-708.
- Mason, J.M., and Arndt, K.M. (2004) Coiled coil domains: stability, specificity, and biological implications. *ChemBiochem* **5**: 170-176.
- Mastellos, D., and Lambris, J.D. (2002) Complement: more than a 'guard' against invading pathogens? *Trends Immunol* **23**: 485-491.
- Mayer, D.C.G., Mu, J.-B., Feng, X., Su, X.-z., and Miller, L.H. (2002) Polymorphism in a *Plasmodium falciparum* erythrocyte-binding ligand changes its receptor specificity. *J Exp Med* **196**: 1523-1528.
- McGregor, K.F., Bilek, N., Bennett, A., Kalia, A., Beall, B., Carapetis, J.R., Currie, B.J., Sriprakash, K.S., Spratt, B.G., and Bessen, D.E. (2004a) Group A streptococci from a remote community have novel multilocus genotypes but share *emm* types and housekeeping alleles with isolates from worldwide sources. *J Infect Dis* **189**: 717-723.
- McGregor, K.F., Spratt, B.G., Kalia, A., Bennett, A., Bilek, N., Beall, B., and Bessen, D.E. (2004b) Multilocus sequence typing of *Streptococcus pyogenes* representing most known *emm* types and distinctions among subpopulation genetic structures. *J Bacteriol* **186**: 4285-4294.
- McLandsborough, L.A., and Cleary, P.P. (1995) Insertional inactivation of *virR* in *Streptococcus pyogenes* M49 demonstrates that VirR functions as a positive regulator of ScpA, FcRA, OF, and M protein. *FEMS Microbiol Lett* **128**: 45-52.
- Mora, M., Bensi, G., Capo, S., Falugi, F., Zingaretti, C., Manetti, A.G.O., Maggi, T., Taddei, A.R., Grandi, G., and Telford, J.L. (2005) Group A streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. *PNAS* **102**: 15641-15646.
- Morens, D.M., Folkers, G.K., and Fauci, A.S. (2004) The challenge of emerging and re-emerging infectious diseases. *Nature* **430**: 242-249.
- Morgan, B.P., and Harris, C.L. (1999) *Complement regulatory proteins*. San Diego: Academic Press.

- Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., Hayashi, H., Hattori, M., and Hamada, S. (2003) Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res* **13**: 1042-1055.
- Navarre, W.W., and Schneewind, O. (1999) Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* **63**: 174-229.
- Neth, O., Jack, D.L., Dodds, A.W., Holzel, H., Klein, N.J., and Turner, M.W. (2000) Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. *Infect Immun* **68**: 688-693.
- Ngampasutadol, J., Ram, S., Blom, A.M., Jarva, H., Jerse, A.E., Lien, E., Goguen, J., Gulati, S., and Rice, P.A. (2005) Human C4b-binding protein selectively interacts with *Neisseria gonorrhoeae* and results in species-specific infection. *Proc Natl Acad Sci U S A* **102**: 17142-17147.
- Nilson, B.H.K., Frick, I.M., Åkesson, P., Forsén, S., Björck, L., Åkerström, B., and Wikström, M. (1995) Structure and stability of protein H and the M1 protein from *Streptococcus pyogenes*. Implications for other surface proteins of gram-positive bacteria. *Biochemistry* **34**: 13688-13698.
- Nonaka, M.I., Wang, G., Mori, T., Okada, H., and Nonaka, M. (2001) Novel androgen-dependent promoters direct expression of the C4b-binding protein alpha-chain gene in epididymis. *J Immunol* **166**: 4570-4577.
- Nordström, T., Blom, A.M., Forsgren, A., and Riesbeck, K. (2004) The emerging pathogen *Moraxella catarrhalis* interacts with complement inhibitor C4b binding protein through ubiquitous surface proteins A1 and A2. *J Immunol* **173**: 4598-4606.
- Nyberg, P., Rasmussen, M., and Björck, L. (2004)  $\alpha$ 2-macroglobulin-proteinase complexes protect *Streptococcus pyogenes* from killing by the antimicrobial peptide LL-37. *J Biol Chem* **279**: 52820-52823.
- Ochsenbein, A.F., and Zinkernagel, R.M. (2000) Natural antibodies and complement link innate and acquired immunity. *Immunology Today* **21**: 624-630.
- Oehmcke, S., Podbielski, A., and Kreikemeyer, B. (2004) Function of the fibronectin-binding serum opacity factor of *Streptococcus pyogenes* in adherence to epithelial cells. *Infect Immun* **72**: 4302-4308.

- Okada, N., Liszewski, M.K., Atkinson, J.P., and Caparon, M. (1995) Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A streptococcus. *Proc Natl Acad Sci USA* **92**: 2489-2493.
- O'Toole, P., Stenberg, L., Rissler, M., and Lindahl, G. (1992) Two major classes in the M protein family in group A streptococci. *Proc Natl Acad Sci USA* **89**: 8661-8665.
- Pangburn, M.K., Schreiber, R.D., and Müller-Eberhard, H.J. (1981) Formation of the initial C3 convertase of the alternative complement pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3. *J Exp Med* **154**: 856-867.
- Pérez-Caballero, D., Albertí, S., Vivanco, F., Sánchez-Corral, P., and Rodríguez de Córdoba, S. (2000) Assessment of the interaction of human complement regulatory proteins with group A streptococcus. Identification of a high-affinity group A streptococcus binding site in FHL-1. *Eur J Immunol* **30**: 1243-1253.
- Pérez-Caballero, D., García-Laorden, I., Cortés, G., Wessels, M.R., de Córdoba, S.R., and Albertí, S. (2004) Interaction between complement regulators and *Streptococcus pyogenes*: binding of C4b-binding protein and factor H/factor H-like protein 1 to M18 strains involves two different cell surface molecules. *J Immunol* **173**: 6899-6904.
- Perez-Casal, J., Okada, N., Caparon, M.G., and Scott, J.R. (1995) Role of the conserved C-repeat region of the M protein of *Streptococcus pyogenes*. *Mol Microbiol* **15**: 907-916.
- Perkins, S.J., Chung, L.B., and Reid, K.B.M. (1986) Unusual ultrastructure of complement-component-C4b-binding protein of human complement by synchrotron X-ray scattering and hydrodynamic analysis. *Biochem J* **233**: 799-807.
- Phillips, R.E. (2002) Immunology taught by Darwin. *Nat Immunol* **3**: 987-989.
- Pleass, R.J., Areschoug, T., Lindahl, G., and Woof, J.M. (2001) Streptococcal IgA-binding proteins bind in the Calpha 2-Calpha 3 interdomain region and inhibit binding of IgA to human CD89. *J Biol Chem* **276**: 8197-8204.
- Podbielski, A., Schnitzler, N., Beyhs, P., and Boyle, M.D. (1996) M-related protein (Mrp) contributes to group A streptococcal resistance to phagocytosis by human granulocytes. *Mol Microbiol* **19**: 429-441.
- Prasadarao, N.V., Blom, A.M., Villoutreix, B.O., and Linsangan, L.C. (2002) A novel interaction of outer membrane protein A with C4b-binding protein mediates serum resistance of *Escherichia coli* K1. *J Immunol* **169**: 6352-6360.



- Price, J.D., Schaumburg, J., Sandin, C., Atkinson, J.P., Lindahl, G., and Kemper, C. (2005) Induction of a regulatory phenotype in human CD4<sup>+</sup> T cells by streptococcal M protein. *J Immunol* **175**: 677-684.
- Prodinger, W.M., Wurzner, R., Stoiber, H., and Dierich, M.P. (2003) Complement. In *Fundamental immunology*. Paul, W.E. (ed). New York: Lippincott Williams and Wilkins, pp. 1077-1103.
- Ram, S., Cullinane, M., Blom, A.M., Gulati, S., McQuillen, D.P., Monks, B.G., O'Connell, C., Boden, R., Elkins, C., Pangburn, M.K., Dahlbäck, B., and Rice, P.A. (2001) Binding of C4b-binding protein to porin: a molecular mechanism of serum resistance of *Neisseria gonorrhoeae*. *J Exp Med* **193**: 281-295.
- Rasmussen, M., Muller, H.P., and Björck, L. (1999) Protein GRAB of *Streptococcus pyogenes* regulates proteolysis at the bacterial surface by binding alpha2-macroglobulin. *J Biol Chem* **274**: 15336-15344.
- Rasmussen, M., and Björck, L. (2001) Unique regulation of SclB - a novel collagen-like surface protein of *Streptococcus pyogenes*. *Mol Microbiol* **40**: 1427-1438.
- Rattle, H. (1995) *An NMR primer for life scientists*. Fareham, UK: Partnership Press.
- Ren, B., Szalai, A.J., Hollingshead, S.K., and Briles, D.E. (2004) Effects of PspA and antibodies to PspA on activation and deposition of complement on the pneumococcal surface. *Infect Immun* **72**: 114-122.
- Retnoningrum, D.S., and Cleary, P.P. (1994) M12 protein from *Streptococcus pyogenes* is a receptor for immunoglobulin G3 and human albumin. *Infect Immun* **62**: 2387-2394.
- Rezcallah, M.S., Hodges, K., Gill, D.B., Atkinson, J.P., Wang, B., and Cleary, P.P. (2005) Engagement of CD46 and  $\alpha 5\beta 1$  integrin by group A streptococci is required for efficient invasion of epithelial cells. *Cell Microbiol* **7**: 645-653.
- Ringdahl, U., Svensson, M., Wistedt, A.C., Renné, T., Kellner, R., Müller-Esterl, W., and Sjöbring, U. (1998) Molecular co-operation between protein PAM and streptokinase for plasmin acquisition by *Streptococcus pyogenes*. *J Biol Chem* **273**: 6424-6430.
- Sandin, C., Linse, S., Areschoug, T., Woof, J.M., Reinholdt, J., and Lindahl, G. (2002) Isolation and detection of human IgA using a streptococcal IgA-binding peptide. *J Immunol* **169**: 1357-1364.
- Sandin, C. (2005) Interactions between streptococcal M proteins and human plasma proteins. *Doctoral thesis*. Lund University, pp. 63.

- Sandin, C., Carlsson, F., and Lindahl, G. (2006) Binding of human plasma proteins to *Streptococcus pyogenes* M protein determines the location of opsonic and non-opsonic epitopes. *Mol Microbiol* **59**: 20-30.
- Scharfstein, J., Ferreira, A., Gigli, I., and Nussenzweig, V. (1978) Human C4-binding protein. I. Isolation and characterization. *J Exp Med* **148**: 207-222.
- Schnitzler, N., Haase, G., Bussing, A., Kaufhold, A., Beyhs, P., and Podbielski, A. (1995) Measuring resistance to phagocytosis of group A and G streptococci: comparison of direct bactericidal assay and flow cytometry. *Med Microbiol Immunol (Berl)* **184**: 17-22.
- Schnitzler, N., Haase, G., Podbielski, A., Luttkien, R., and Schweizer, K.G. (1999) A co-stimulatory signal through ICAM- $\beta_2$  integrin-binding potentiates neutrophil phagocytosis. *Nat Med* **5**: 231-235.
- Schwalbe, R.A., Dahlbäck, B., and Nelsestuen, G.L. (1990) Independent association of serum amyloid P component, protein S, and complement C4b with complement C4b-binding protein and subsequent association of the complex with membranes. *J Biol Chem* **265**: 21749-21757.
- Sammelweis, I. (1861) *Die Aetiologie, der Begriff und die Prophylaxis des Kindbettfiebers (Pest: CA Hartleben)*. Translation to English by Carter, KC: *The etiology, concept and prophylaxis of chilbedfever*, University of Wisconsin Press, 1983.
- Sen, J.W., and Heegaard, N.H. (2002) Serum amyloid P component does not circulate in complex with C4-binding protein, fibronectin or any other major protein ligand. *Scand J Immunol* **56**: 85-93.
- Shakhnovich, E., Abkevich, V., and Ptitsyn, O. (1996) Conserved residues and the mechanism of protein folding. *Nature* **379**: 96-98.
- Skehel, J.J., and Wiley, D.C. (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* **69**: 531-569.
- Smoot, J.C., Barbian, K.D., Van Gompel, J.J., Smoot, L.M., Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M., Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M., Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G., and Musser, J.M. (2002) Genome sequence and comparative microarray analysis of serotype M18 group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. *PNAS* **99**: 4668-4673.
- Steinhauer, D.A., and Skehel, J.J. (2002) Genetics of influenza viruses. *Annu Rev Genet* **36**: 305-332.

- Stenberg, L., O'Toole, P., and Lindahl, G. (1992) Many group A streptococcal strains express two different immunoglobulin-binding proteins, encoded by closely linked genes: characterization of the proteins expressed by four strains of different M-type. *Mol Microbiol* **6**: 1185-1194.
- Stenberg, L., O'Toole, P.W., Mestecky, J., and Lindahl, G. (1994) Molecular characterization of protein Sir, a streptococcal cell surface protein that binds both immunoglobulin A and immunoglobulin G. *J Biol Chem* **269**: 13458-13464.
- Stevens, D.L. (1999) The flesh-eating bacterium: what's next? *J Infect Dis* **179** Suppl 2: S366-374.
- Stevens, D.L., and Kaplan, E.L. (2000) *Streptococcal Infections. Clinical aspects, microbiology and molecular pathogenesis*. New York: Oxford University Press.
- Stevens, J., Corper, A.L., Basler, C.F., Taubenberger, J.K., Palese, P., and Wilson, I.A. (2004) Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* **303**: 1866-1870.
- Strömberg, N., Marklund, B.I., Lund, B., Ilver, D., Hamers, A., Gaastra, W., Karlsson, K.A., and Normark, S. (1990) Host-specificity of uropathogenic *Escherichia coli* depends on differences in binding specificity to Gal $\alpha$  1-4Gal-containing isoreceptors. *EMBO J* **9**: 2001-2010.
- Sumby, P., Porcella, S.F., Madrigal, A.G., Barbian, K.D., Virtaneva, K., Ricklefs, S.M., Sturdevant, D.E., Graham, M.R., Vuopio-Varkila, J., Hoe, N.P., and Musser, J.M. (2005) Evolutionary origin and emergence of a highly successful clone of serotype M1 group A streptococcus involved multiple horizontal gene transfer events. *J Infect Dis* **192**: 771-782.
- Sun, H., Ringdahl, U., Homeister, J.W., Fay, W.P., Engleberg, N.C., Yang, A.Y., Rozek, L.S., Wang, X., Sjöbring, U., and Ginsburg, D. (2004) Plasminogen is a critical host pathogenicity factor for group A streptococcal infection. *Science* **305**: 1283-1286.
- Swanson, J., Hsu, K.C., and Gotschlich, E.C. (1969) Electron microscopic studies on streptococci. I. M antigen. *J Exp Med* **130**: 1063-1091.
- Terao, Y., Kawabata, S., Kunitomo, E., Murakami, J., Nakagawa, I., and Hamada, S. (2001) Fba, a novel fibronectin-binding protein from *Streptococcus pyogenes*, promotes bacterial entry into epithelial cells, and the *fba* gene is positively transcribed under the Mga regulator. *Mol Microbiol* **42**: 75-86.

- Thern, A., Stenberg, L., Dahlbäck, B., and Lindahl, G. (1995) Ig-binding surface proteins of *Streptococcus pyogenes* also bind human C4b-binding protein (C4BP), a regulatory component of the complement system. *J Immunol* **154**: 375-386.
- Thern, A., Wästfelt, M., and Lindahl, G. (1998) Expression of two different antiphagocytic M proteins by *Streptococcus pyogenes* of the OF<sup>+</sup> lineage. *J Immunol* **160**: 860-869.
- Van Holde, K.E., Johnson, C., and Ho, P.S. (1988) NMR Spectroscopy. In *Principles of Physical Biochemistry*: Prentice Hall, pp. 484-524.
- van Luenen, H.G.A.M., Kieft, R., Mussmann, R., Engstler, M., ter Riet, B., and Borst, P. (2005) Trypanosomes change their transferrin receptor expression to allow effective uptake of host transferrin. *Mol Microbiol* **58**: 151-165.
- Verschoor, A., and Carroll, M.C. (2004) Complement and its receptors in infection. In *The innate immune response to infection*. Kaufmann, S.H.E., Medzhitov, R. and Gordon, S. (eds). Washington D.C.: ASM press, pp. 219-240.
- Villiers, M.B., Reboul, A., Thielens, N.M., and Colomb, M.G. (1981) Purification and characterization of C4-binding protein from human serum. *FEBS Lett* **132**: 49-54.
- von Pawel-Rammingen, U., Johansson, B.P., and Björck, L. (2002) IdeS, a novel streptococcal cysteine proteinase with unique specificity for immunoglobulin G. *EMBO J* **21**: 1607-1615.
- Walport, M.J. (2001a) Complement. First of two parts. *N Engl J Med* **344**: 1058-1066.
- Walport, M.J. (2001b) Complement. Second of two parts. *N Engl J Med* **344**: 1140-1144.
- Webb, J.H., Blom, A.M., and Dahlbäck, B. (2002) Vitamin K-dependent protein S localizing complement regulator C4b-binding protein to the surface of apoptotic cells. *J Immunol* **169**: 2580-2586.
- Webb, J.H., Blom, A.M., and Dahlbäck, B. (2003) The binding of protein S and the protein S-C4BP complex to neutrophils is apoptosis dependent. *Blood Coagul Fibrinolysis* **14**: 355-359.
- Weissman, S.J., Moseley, S.L., Dykhuizen, D.E., and Sokurenko, E.V. (2003) Enterobacterial adhesins and the case for studying SNPs in bacteria. *Trends Microbiol* **11**: 115-117.
- 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. *Proc Natl Acad Sci USA* **88**: 8317-8321.
- Wessels, M.R. (2005) Streptolysin S. *J Infect Dis* **192**: 13-15.
- Wexler, D.E., Chenoweth, D.E., and Cleary, P.P. (1985) Mechanism of action of the group A streptococcal C5a inactivator. *Proc Natl Acad Sci USA* **82**: 8144-8148.

- Whatmore, A.M., and Kehoe, M.A. (1994) Horizontal gene transfer in the evolution of group A streptococcal *emm*-like genes: gene mosaics and variation in Vir regulons. *Mol Microbiol* **11**: 363-374.
- Whitnack, E., and Beachey, E.H. (1982) Antipsonic activity of fibrinogen bound to M protein on the surface of group A streptococci. *J Clin Invest* **69**: 1042-1045.
- Widdowson, J.P., Maxted, W.R., and Grant, D.L. (1970) The production of opacity in serum by group A streptococci and its relationship with the presence of M antigen. *J Gen Microbiol* **61**: 343-353.
- Wilson, I.A., and Cox, N.J. (1990) Structural basis of immune recognition of influenza virus hemagglutinin. *Annu Rev Immunol* **8**: 737-771.
- Wistedt, A.C., Ringdahl, U., Müller-Esterl, W., and Sjöbring, U. (1995) Identification of a plasminogen-binding motif in PAM, a bacterial surface protein. *Mol Microbiol* **18**: 569-578.
- Zeng, X., Zhu, H., Lashuel, H.A., and Hu, J.C. (1997) Oligomerization properties of GCN4 leucine zipper e and g position mutants. *Protein Sci* **6**: 2218-2226.
- Ziccardi, R.J., Dahlbäck, B., and Müller-Eberhard, H.J. (1984) Characterization of the interaction of human C4b-binding protein with physiological ligands. *J Biol Chem* **259**: 13674-13679.