Novel Applications of Shotgun Phage Display

Anna Rosander

Department of Microbiology
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2004

Acta Universitatis Agriculturae Sueciae

Agraria 450

Abstract

In a shotgun phage display library, theoretically, the entire proteome of a bacterium is represented. Phages displaying specific polypeptides can be isolated by affinity selection, while the corresponding gene remains physically linked to the gene product. The overall objective of the study in this thesis was to explore the shotgun phage display technique in new areas. Initially, it was used to study interactions between *Staphylococcus aureus* and an *in vivo* coated biomaterial. It was shown to be well suited for the identification of bacterial proteins that bind to *ex vivo* central venous catheters. Several known interactions were detected, but it was also found that β_2-glycoprotein I (β_2-GPI) is deposited on this type of biomaterial – a finding that is of interest both for the adherence of *S. aureus*, but perhaps also in view of the occurrence of autoantibodies in certain autoimmune diseases.

Further, it is of interest to identify the subset of extracellular proteins in a bacterium since they are involved in important functions like pathogenesis and symbiosis. A method that allows for the rapid and general isolation of extracellular proteins is desirable, and may prove particularly useful when applied to bacteria for which the genome sequences are not known. For this purpose, a specialised phage display method was developed to isolate extracellular proteins by virtue of the presence of signal peptides (SS phage display). It was successfully applied to *S. aureus* and, on a larger scale, to the symbiotic bacterium *Bradyrhizobium japonicum*.

In elaboration of the SS phage display method, an inducible antisense RNA system was incorporated to enable gene silencing of the isolated genes. A tetracycline-regulated promoter was inserted in such a way, that an antisense RNA covering the cloned gene could be expressed. The new element was shown to be compatible with the properties of SS phage display, and to promote gene expression upon induction on both the transcriptional and translational level. However, screening for clones affected by the induction of antisense RNA transcription was unsuccessful, and further developments of the system are required to improve the efficiency of this attractive application.

Keywords: Shotgun phage display, extracellular proteins, protein export, gene regulation, antisense RNA, biomaterial, *Staphylococcus aureus*, *Bradyrhizobium japonicum*, *Escherichia coli*

Author's address: Anna Rosander, Department of Microbiology, Swedish University of Agricultural Sciences, Box 7025, SE-750 07 UPPSALA, Sweden.
E-mail: anna.rosander@mikrob.slu.se
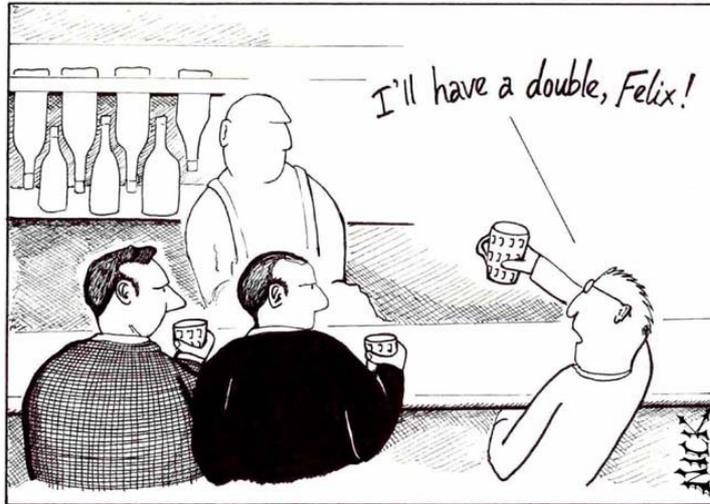
Cambridge, 1953. Shortly before discovering the structure of DNA, Watson and Crick, depressed by their lack of progress, visit the local pub.

Contents

Appendix

Papers I-V

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals:

I   Rosander, A., Bjerketorp, J., Frykberg, L. & Jacobsson, K. 2002. Phage display as a novel screening method to identify extracellular proteins. *Journal of Microbiological Methods* 51, 43-55.

II Rosander, A., Frykberg, L., Ausmees, N. & Müller, P. 2003. Identification of extracytoplasmic proteins in *Bradyrhizobium japonicum* using phage display. *Molecular Plant-Microbe Interactions* 16, 727-737.

III Rosander, A., Jacobsson, K. & Frykberg, L. Shotgun phage display in combination with an inducible antisense RNA system. Manuscript.

IV Bjerketorp, J., Rosander, A., Nilsson, M., Jacobsson, K. & Frykberg, L. Sorting of a *Staphylococcus aureus* phage display library against *ex vivo* biomaterial. Submitted.

V   Jacobsson, K., Rosander, A., Bjerketorp, J. & Frykberg, L. 2003. Shotgun phage display - selection for bacterial receptins or other exported proteins. *Biological Procedures Online* 5, 123-135.

Papers I and II are reprinted with permission from the publishers.

Background

We witness today a rapid development of genome sequencing. Less than ten years ago the first completed and assembled genome of an organism higher than a virus was published, *i.e.* that of *Haemophilus influenzae* (Fleischmann *et al.*, 1995). At present, complete genome sequences of more than 150 microbial organisms are known. The research frontier is moving from genomics to proteomics to elucidate the function of the encoded proteins. Powerful bioinformatic and proteomic tools provide valuable information, but still, analysis of genomic sequences have indicated that about one third of the open reading frames in a bacterial genome are of unknown function, *e.g.* 30% of the essential genes of the minimal bacterial genome of *Mycoplasma genitalium* (Hutchison *et al.*, 1999). In the post-genomic era, large scale experimental methods to complement and verify the *in silico* retrieved information are being developed, and have to continue to be so.

Proteins function in many different ways. In bacteria, one way is to communicate with or establish a connection to cells in the environment *e.g.* to eukaryotic host cells. The bacteria have extracellular proteins that are connected to the cell surface in different ways or secreted from the cell, and with which contact can be mediated. These proteins are of interest to study since the contact between microbial and eukaryotic organisms results in important consequences like symbiosis and pathogenesis. Of particular interest as a background to the present study are the extracellular proteins involved in the pathogenesis of *Staphylococcus aureus*. Several functions governed by these proteins provide important mechanisms for the bacteria to establish an infection. Colonisation of host tissues, or implanted materials, is promoted by adhesive proteins that bind to host proteins, generally those of the extracellular matrix such as collagen and fibronectin. The bacterial spread in tissues is promoted by *e.g.* host membrane-damaging toxins and various proteases. *S. aureus* expresses a number of factors that have the potential to interfere with host defence mechanisms. Plasma coagulating proteins may protect the bacteria by causing localised clotting, while serum component binding proteins can mask the bacteria from *e.g.* phagocytic immune defences. To provide nutrients for the bacteria, protein-, lipid-, and DNA-degrading enzymes are secreted into the environment. Several different types of toxins are released from *S. aureus*, which are likely to provoke symptoms of disease. Some have superantigen activity that disturbs the immune system of the host. In addition, communication between bacterial cells is mediated by so-called autoinducers, small proteins that monitor the presence of other bacteria in their surroundings.

This thesis deals with the isolation and characterisation of bacterial extracellular proteins by a method called phage display. In the following, a general introduction to the method is given as well as a background to extracellular proteins and protein export in bacteria, and to regulation of gene expression, focusing on antisense RNA regulation.

Introduction

Phage display

Phage display offers a powerful tool for the identification of protein-protein interactions and holds a key feature by supplying a physical link between a gene and a gene product of interest. Foreign polypeptides are displayed on the surface of a bacteriophage and are isolated by affinity selection in a procedure called panning. The first article on phage display was published in 1985 (Smith, 1985). A foreign gene fragment was inserted into the gene encoding one of the coat proteins of filamentous phage f1 in a non-disruptive way, creating a fusion protein. This fusion protein was shown to be incorporated into the phage coat and displayed on the phage surface. Such 'fusion phage' particles were found to retain the ability of infecting its host, at least partially, thus enabling propagation. Monoclonal antibodies directed against the foreign polypeptide encoded by the inserted gene fragment were used to isolate fusion phage in numbers 1000-fold over wild-type phage by affinity purification. The construction of random insert phage libraries was proposed, implying a cloning technique where a large number of clones can be investigated in a short period of time. Since this first achievement, phage display has evolved in different directions and is successfully used in a variety of applications.
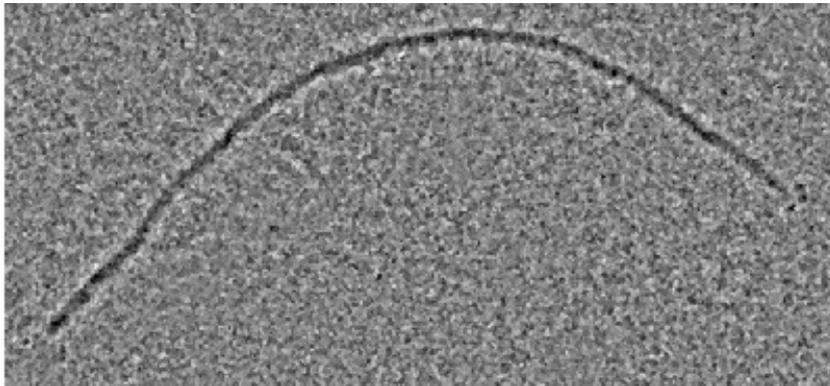
Filamentous phage

Structure

The filamentous phage constitute a large family of bacterial viruses, which infect a variety of Gram-negative bacteria, using pili as receptors. The Ff group, including the virtually identical strains M13, fd and f1 and their derivatives, is the most frequently used for phage display purposes. The phage capsid appears as a thin and flexible tubular structure of about 900 x 10 nm (figure 1). The single-stranded DNA genome is composed of approximately 6400 nucleotides from which 11 proteins are encoded. A non-coding intergenic region contains the signals for initiation of DNA synthesis, for termination of RNA synthesis, and the packaging signal. Although different designations have been assigned to the phage genes and proteins, the one used throughout this thesis applies Roman numerals *e.g.* gIII and pIII for gene III and protein III, respectively. Five coat proteins encompass the circular genome, three proteins are needed for DNA synthesis (pII, pV, and pX), and three others are involved in phage assembly and export (pI, pIV, and pXI). The phage coat consists mainly of pVIII, present in approximately 2800 copies. About five copies each of the minor coat proteins are located at the particle ends. Protein III and pVI are at the end involved in infection, and both are required for particle stability. Protein VII and pIX, necessary for efficient particle assembly, are at the other end (reviewed in *e.g.* Russel, 1995 and Russel *et al.*, 1997). The negatively charged amino-terminal (N-terminal) regions of mature pVIII molecules are exposed on the outer surface of the virion, whereas the carboxy-terminal (C-terminal) segments line the cylindrical hole generated by the protein sheath in which the DNA is contained. An apolar domain in the interior of the pVIII

8

molecule hydrophobically interacts with corresponding domains of other pVIII molecules, to hold the capsid subunits together (Glucksman *et al.*, 1992; Marvin *et al.*, 1994). Also, similar apolar domains are found in all four minor coat proteins suggesting that they are associated with the pVIII sheath in the same way as pVIII molecules are associated to each other. Protein VII has been proposed to interact with the pVIII array and to be shielded from the environment, whereas pIX interacts with pVII and is exposed to the environment. The pIII and pVI proteins form a stable entity in the virion (Gailus & Rasched, 1994). The C-terminal end of pVI in the virion is not accessible to anti-pVI serum (Endemann & Model, 1995), still, structural data suggest that it is exposed to the surrounding medium (Makowski, 1992).

A.

B.
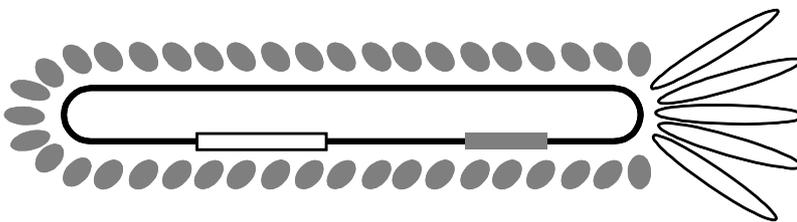
Figure 1. The filamentous phage. **A.** Electron micrograph of a negatively stained fd filamentous bacteriophage virion (http://www-ccmr-nmr.bioc.cam.ac.uk/~dam4/main.html; 16-Apr-2004. Printed with kind permission from Professor Richard Perham at the University of Cambridge, UK). The globular domains of pIII are visible at one end. **B.** Schematic picture of a wild-type filamentous phage. White areas represent gIII and pIII, grey areas represent gVIII and pVIII.

Infection

During filamentous phage infection, the host cell remains able to grow and divide since propagated phage leave the host without cell lysis, a feature not common in bacteriophage infection. Ff phages infect *Escherichia coli* cells *via* the F pili in a membrane-associated event. Infection is initiated by a pIII-pili adsorption process. Studies on pIII suggest a domain structure – domains D1 and D2 are exposed and

gVIII based systems

A.

B.

C. +

gIII based systems

D.

E.

F. +

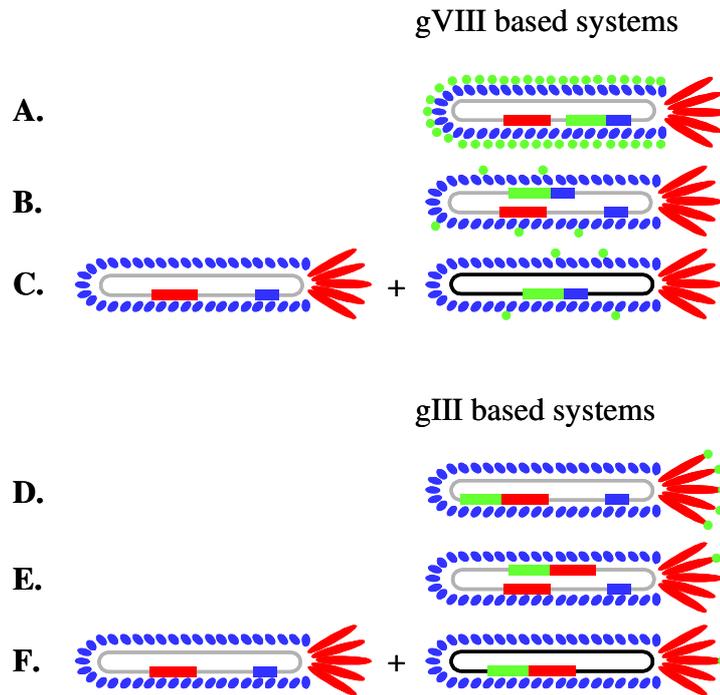*Figure 2*. Different phage display systems. Grey and black circular elements within phage particles represent a phage genome or a phagemid, respectively. Green boxes and spheres represent the foreign DNA insert and the corresponding gene product, respectively. Blue colour elements indicate gVIII or pVIII, red colour elements indicate gIII or pIII. Two wild-type helper phage particles associated with the use of phagemids are shown to the left.

susceptible to subtilisin digestion (Gray *et al.*, 1981), domain D3 is buried within the pVIII array (Kremser & Rasched, 1994). Two glycine-rich linker regions divide the three domains in pIII. The D2 domain of pIII, and possibly the D1-D2 interconnecting glycine-rich region (Nilsson *et al.*, 2000), interacts with the tip of the bacterial pilus. It is thought that the virion is drawn into contact with the cell by resorbtion of the pilus to the plasma membrane (Jacobson, 1972) where D1 attaches to the bacterial co-receptor outer membrane protein TolA (Riechmann & Holliger, 1997). Subsequent translocation of the phage DNA into the cytoplasm requires all products of the *tolQRA* genes, which are also needed for the uncoating of pVIII by membrane insertion (Click & Webster, 1998; Webster, 1991). In the cytoplasm, phage ssDNA is converted by host enzymes to a double-stranded replicative form, that serves as template for rolling circle replication and for phage gene expression. The dsDNA can be isolated from the cell and used for cloning of a foreign fragment.

Assembly

All phage proteins, apart from the ones involved in DNA synthesis, are integral membrane proteins and move to the cell membrane after synthesis. Phage assembly and export take place simultaneously as the phage protrudes through the inner and outer membranes (Russel *et al.*, 1997). After phage DNA replication, the DNA is

10

covered by approximately 1500 pV molecules, which are replaced by pVIII molecules as the DNA is extruded through the assembly site. The DNA packaging signal is recognised by pVII and pIX, as this is the first end of the assembled phage to emerge from the cell (Russel & Model, 1989). When either pIII or pVI is absent, polyphage are produced. Hence, pIII and pVI are proposed to participate in the termination of assembly (Endemann & Model, 1995). Detailed analysis of deletion mutants revealed that D3 contains two functionally distinct subdomains: the C-terminal domain 1, which is involved in capping and stabilizing the assembled phage particle, and the C-terminal domain 2, which is required for incorporation into the phage coat and subsequent release of the assembled virion from the host membrane (Rakonjac *et al.*, 1999). Three phage-encoded proteins and at least one host-encoded protein (thioredoxin) are required for phage assembly, and are assumed to promote morphogenesis (Russel & Model, 1985; Russel, 1991). Assembly occurs at sites formed by multimers of additional phage-encoded proteins. Protein IV in the outer membrane interacts with pI and pXI in the inner membrane to form an assembly site through which phage particles are extruded (Russel *et al.*, 1997). The C-terminal half of pIV is located to the outer membrane while the N-terminal extends into the periplasm. It forms a stable 10-12 subunit oligomer and has been proposed to form a pore for virus assembly. Homologies to proteins involved in outer membrane protein export have been identified (Russel & Kazmierczak, 1993). The role of thioredoxin in phage assembly has never been completely elucidated, but apparently it acts as a DNA-handling protein, not a redox enzyme (Russel, 1995).

Phage display systems

The two following sections are reviewed in Smith & Petrenko, 1997.

Insertion in the filamentous phage genome

For phage display purposes, foreign DNA fragments have been inserted mainly into the genes of phage coat proteins pIII and pVIII, although all five coat proteins have been used for library construction. The small size of pVIII and the exposure of only the N-terminal part of the protein to the medium dictate the fusion site of foreign polypeptides only to the N-terminal end. For pIII with a larger fraction exposed, fusions have been made both within the molecule and to the N-terminal end. There are a variety of systems available for the fusion of peptides to these proteins. A foreign DNA insert in fusion with gVIII in the phage genome generates only one type of pVIII molecules (figure 2**A**). With this system, only about 6 amino acids can be fused to pVIII without the impairment of phage production (Iannolo *et al.*, 1995). If pIII is used (figure 2**D**), problems with impaired infectivity may occur if large polypeptide structures are fused to the molecule. For the display of larger polypeptides, supplementation of wild-type coat protein molecules can circumvent these problems. This can be achieved in two ways: i) by the introduction of a wild-type copy of gVIII or gIII to the phage genome in addition to the one used for insertion (figures 2**B** and 2**E**), or ii) the use of phagemid vectors in combination with helper phage (figures 2**C** and 2**F**). Both systems will produce mosaic phage particles of which the coats are composed of a mixture of recombinant and wild-type pIII or pVIII molecules, respectively.

Phagemids

To circumvent the impaired infectivity and enable display of larger polypeptides, the use of phagemid vectors in phage display was introduced in 1990 (Bass *et al.*, 1990). A phagemid is a combination of elements from the phage genome, such as the intergenic region and a gene encoding a phage coat protein, and a plasmid carrying a resistance marker. Thus, phagemids can be grown as plasmids or alternatively, packaged into a phage coat if all the structural proteins required for generating a complete phage is supplied by a wild-type helper phage. Usually, when using phagemids containing gIII, the D1 and D2 domains have been removed since cells expressing this domain are resistant to superinfection by helper phage. When using phagemid vectors based on gVIII, multivalent recombinant pVIII display is achieved. The amount of fusion protein varies, and is believed to depend on a number of different factors, such as the size, the folding, and the amino acid composition of the foreign protein.

Other display systems

When using the non-lytic filamentous phage, fusion coat proteins have to be compatible with the assembly transport process in *E coli*. Also, toxic polypeptides encoded by the foreign insert may affect *E. coli* viability and the continuous production of phage particles. These problems have been motivating factors for the development of display systems based on lytic phage. Phage λ, T4, and T7 have been used with promising results. Two capsid proteins of phage λ have been used for display purposes – protein D of the capsid head and pV of the tail. The most frequent use of phage λ display has been to reveal the minimal determinants of a particular function within a single gene (Hoess, 2002). Dispensable T4 capsid proteins Soc and Hoc have been used for the display of peptides and antibodies (Jiang *et al.*, 1997; Ren & Black, 1998), and the simultaneous display of different foreign polypeptides on these two proteins was proposed. The 10B protein of phage T7 has been employed mainly for peptide display purposes (Castillo *et al.*, 2001; Houshmand *et al.*, 1999). An interesting hybrid system of T7 based phage display and combinatorial synthetic chemistry was recently presented (Woiwode *et al.*, 2003). Cloning of cDNA fragments have been made, using both the D protein of phage λ as well as the 10B protein of phage T7 (Houshmand & Bergqvist, 2003; Santini *et al.*, 1998; Yamamoto *et al.*, 1999) – an attractive application, which is unsuitable for most filamentous phage display systems (see section 'Other applications'). Also, phage P4 and MS2 have been used for peptide display (Heal *et al.*, 1999; Lindqvist & Naderi, 1995), but so far, reports are scarce. In addition, protein display systems have been developed that use other display surfaces than phage, *e.g.* bacteria (reviewed in Georgiou *et al.*, 1997 and Samuelson *et al.*, 2002). Both Gram-negative and Gram-positive bacteria have been employed as carriers of foreign peptides, and diverse areas of usage have been proposed, such as the construction of (poly)peptide libraries for the isolation of high-affinity ligands to target molecules, generation of recombinant live vaccines, and bioremediation. One advantage of using cell surface display is that a powerful cell sorting technique can be applied for high throughput screening – the fluorescence activated cell-sorting (FACS) technique. Well-defined, cell-free *in vitro* methods have also been used for protein display. The physical link between

12

genotype and phenotype is managed by the use of *e.g.* ribosomes to which the selected peptide is associated as well as the template RNA (reviewed in Dower & Mattheakis, 2002).

Shotgun phage display

One application of phage display has been called Shotgun phage display, indicating its relation to shotgun cloning. In this approach, randomly fragmented bacterial chromosomal DNA is inserted into a phagemid vector. Theoretically, in a phage library constructed from the ligated material, polypeptides derived from all genes on the bacterial chromosome are represented. In phagemid vector pG8SAET (Zhang *et al.*, 1999; see figure in **V**), commonly used for construction of shotgun libraries, the main features surrounding the *Sna*BI cloning site are the staphylococcal protein A (Spa) derived promoter and signal peptide, the expression tag (E-tag), and the phage gene VIII. The signal peptide is required for the translocation of fusion protein to the cell membrane. The use of the Spa promoter and signal peptide was a result from the testing of several vectors (Jacobsson & Frykberg, 1998). The E-tag encodes a 15 amino acid epitope, which is recognised by monoclonal antibodies. It can be used to monitor the enrichment of correct clones after consecutive rounds of panning, and also to purify recombinant protein produced by the vector. The vector is based on gene VIII for multivalent display, which may result in the selection of low affinity interactions. Still, for the identification of different naturally occurring interactions, this could result in a more efficient selection even though the risk of selecting for irrelevant clones may also increase.

Shotgun phage display has successfully been used to identify staphylococcal and streptococcal receptins, *i.e.* 'microbial binding proteins that interact with mammalian target proteins' (Kronvall & Jonsson, 1999) as well as putative lectins (carbohydrate-binding proteins) in *Rhizobium leguminosarum* bv. *trifolii* (table 1). Gene fragment libraries have also been used to map the binding domains of bacterial receptins (Jacobsson *et al.*, 1997; Zhang *et al.*, 1999) as well as antigen epitopes recognised by monoclonal antibodies (*e.g.* in Petersen *et al.*, 1995).

Other applications

The most commonly used applications of phage display are antibody and peptide libraries. Among many excellent reviews on these applications I choose to refer to two recent ones, Hoogenboom, 2002 and Szardenings, 2003, respectively, in the following sections.

Antibody libraries

Phage display has been utilised for the production of recombinant antibodies as an alternative to conventional monoclonal or polyclonal antibodies. The production of recombinant antibodies provides a useful alternative approach for *e.g.* non-immunogenic or toxic antigens. Also, engineering of antibodies to improve affinity and specificity can be achieved. Antibodies are made up of two polypeptides – the

Table 1. *Bacterial receptins and putative lectins identified by shotgun phage display*

Source	Protein(s)	Ligand(s)	Reference(s)
Staphylococcus aureus	Sbi	IgG, β_2-glycoprotein I	Jacobsson & Frykberg, 1995; Zhang *et al.*, 1998; Zhang *et al.*, 1999
Staphylococcus aureus	vWbp	von Willebrand factor	Bjerketorp *et al.*, 2002
Staphylococcus aureus	Coa, Efb	Platelets	Heilmann *et al.*, 2002
Staphylococcus epidermidis	Fbe	Fibrinogen	Nilsson *et al.*, 1998
Staphylococcus epidermidis	Embp	Fibronectin	Williams *et al.*, 2002
Staphylococcus lugdunensis	Wbl	von Willebrand factor	Nilsson *et al.* 2004
Streptococcus dysgalactiae	DemA	Fibrinogen	Vasi *et al.*, 2000
Streptococcus equi subsp. *zooepidemicus*	FNZ	Fibronectin	Lindmark *et al.*, 1996
Streptococcus equi subsp. *equi*	SFS	Fibronectin	Lindmark & Guss, 1999
Streptococcus agalactiae	ScpB	Fibronectin	Beckmann *et al.*, 2002
Streptococcus agalactiae	Fgag	Fibrinogen	Jacobsson, 2003
Rhizobium leguminosarum bv. *trifolii*	Rap	Whole cells of *R. leguminosarum*	Ausmees *et al.*, 2001

heavy chain (H) and the light chain (L), both with constant (C) and variable (V) domains (C_{H1}, C_{H2}, C_{H3}, V_H and C_L, V_L, respectively). The variable domains of both chains combine to make the antigen-binding site. Several subfragments of an antibody molecule can be recognised, among these the Fv ($V_H + V_L$) and Fab (C_{H1}, $V_H + C_L$, V_L) fragments. They have been the basis for development of two strategies to display heterodimeric antibody molecules on the phage surface – single chain Fv (scFv) and Fab libraries. Unlike whole antibodies, Fab and scFv fragments can be functionally expressed in *E. coli* (Better *et al.*, 1988; Skerra & Pluckthun, 1988). To construct an scFv library, V_H and V_L cDNA is prepared by reverse transcription of mRNA obtained from B-lymphocytes. After PCR amplification, the two different cDNAs are assembled into a single gene using a DNA linker fragment. Fab-fragments are displayed by co-expression of both chains, one fused to pIII and the other soluble and targeted to the periplasmic space, where they associate non-covalently. This application of phage display can be used to generate libraries derived from lymphocyte pools of non-immunised or immunised donors. The former has the advantage that it can be used against any antigen, but to obtain good affinities, the library has to be rather large. The latter can only be used against immunogenic antigens but provides a high frequency of high affinity binders. Specialised libraries derived from B-lymphocytes in a patient with disease could enable the selection of specific molecules, such as tumor markers or autoimmune antibodies.

Peptide libraries

Random peptide libraries, or epitope libraries, are constructed by the fusion of chemically synthesised random oligonucleotides to the gene encoding a phage coat protein. They are either completely randomised or flanked by *e.g.* codons for

14

cysteine residues to provide a structurally constrained peptide after formation of a disulfide bond. Also, codons for defined amino acids can be comprised within the oligonucleotide to predefine a minimal affinity to a given target. Generally, the oligonucleotides contain 6-10 random triplets (Smith & Petrenko, 1997), where each triplet includes codons for all 20 natural amino acids. Nowadays, even with the inherent limitations of the phage display system, mainly at the stage of electroporation, libraries with 10^{10} unique clones can readily be attained. Still, most peptide libraries will not be able to cover all possible combinations of amino acids for random peptide libraries with more than 8 variable positions. Thus, attempts have been made to evolve peptides from a reduced amino acid code and to restrict the potential number of amino acids at any given position (Kamtekar et al., 1993).

Peptide libraries have been applied for several different purposes. The mapping of epitopes recognised by antibodies have been performed, e.g. to discover disease specific epitopes or identify a natural antigen, which can be of diagnostic and prognostic significance. Selection of enzyme inhibitors and substrates has been exploited by the use of peptide libraries. In particular, proteolytic enzymes have been implicated as the pathological agent in a number of specific states of a disease, why they are attractive therapeutic targets. Identification of receptor ligands has also been employed. A group of integrins has been studied thoroughly and peptides containing the RGD-motif as well as new binding motifs have been identified. More complicated biological systems, including the vasculature of whole organs and human patients, have also been used for in vivo applications, such as the homing to diseased areas. In the context of peptide libraries, it is near at hand to mention the use of scaffold proteins to which completely novel functions are introduced. Usually, surrogate antibody molecules are the objectives of this approach. An engineered variant of the native IgG-binding B domain of Spa, the Z domain (Nord et al., 1997), and the cellulose-binding domain (CBD) of cellobiohydrolase I from the filamentous fungus Trichoderma reesei (Smith et al., 1998), among others, have been used as protein frameworks on which some part, often loop structures, have been randomised to select for ligand-specific binders.

cDNA libraries

Another desirable application of phage display is the display of eukaryotic proteins from full-length cDNA fragments. However, during cDNA synthesis by poly(A)-priming of mRNA, translational stop codons are retained within the fragments, which prevents the production of fusion protein when fused to the N-terminus of pIII or pVIII. The integrity of the C-terminus of both proteins is essential for phage assembly, why fusions of foreign polypeptides are tolerated only at the N-terminus. Instead, the minor coat protein VI has been used for C-terminal display of cDNA encoded polypeptides. For instance, Fransen et al. (1999) reported the identification of novel peroxisomal enzymes from the screening of a cDNA phage library derived from rat liver, and Jespers et al. (1995) used a cDNA library from the parasite Ancylostoma caninum to identify novel serine protease inhibitors. Another solution to the problem made use of the Fos and Jun transcription factor proteins, which interact with each other and form a leucine zipper (Crameri & Suter, 1993). Further modifications by the introduction of cystein residues at the N-

15

and C-terminals of both proteins enabled covalent attachment between the two molecules. In the phagemid vector pJuFo, the *jun* gene is fused to the N-terminal of pIII, and the cDNA fragments are incorporated in fusion with the C-terminus of the *fos* gene. Both fusion proteins are translocated to the periplasmic space where they associate with each other. Following phage assembly, they are displayed as an entity on the phage surface. This system has been widely used to clone IgE antigens. Certain limitations with the use of cDNA fragments in phage display can be identified, such as the restricted codon usage of *E. coli* as a host and the lack of post-translational modifications. Furthermore, the folding capacity of *E. coli* has been shown to be limited, especially with regard to structures imbedded into membrane bilayers (Baneyx, 1999; Gottesman & Hendrickson, 2000).

Bacterial extracellular proteins and export

In this thesis, extracellular proteins are defined as cell membrane- or cell wall-associated proteins and proteins secreted into the environment. Extracellular proteins have been reported to constitute about 25 % of all proteins in bacteria (Pugsley, 1993; Tjalsma *et al.*, 2000) and are involved in a variety of important cell functions, such as organelle biogenesis, nutrient acquisition, and signal transduction. In particular, extracellular proteins play a key role in pathogenic or symbiotic bacteria by mediating interactions with their eukaryotic hosts. With regard to the abundance of these proteins, the diverse functions they possess, and the number of different microbial organisms that exist, an attempt to cover all extracellular proteins would be overwhelming. Therefore, I choose to discuss only a small group of extracellular proteins in the light of the pathogenicity of *S. aureus*, since this has special relevance to the study of this thesis.

S. aureus is part of our normal skin flora but is also an important human pathogen when circumstances permit. It causes a wide variety of diseases like scalded skin syndrome, arthritis, septicaemia, and food poisoning. Also, it is frequently isolated from implanted materials, to which it can adhere and thereby cause severe infections. As a result of the extensive use of antibiotics, the prevalence of resistance has gradually increased in *S. aureus* and multi-resistant strains occur, especially in hospital-acquired infections. Consequently, our means for treatment decreases, and the need for non-antibiotic related treatments emerges. Several efforts to develop vaccines against *S. aureus* have been, and are being, made (Lee, 2001). One approach used in vaccine development is to counteract what is believed to be the first and crucial step of staphylococcal infections – adherence to host proteins. Many proteins with affinity for mammalian ligands have been identified in *S. aureus* (table 2), and more have been predicted from the complete genome sequence of *S. aureus* (Kuroda *et al.*, 2001). These receptins are proposed to be important virulence factors and are as such potential candidates as vaccine components. They share a common feature in being extracellular proteins, most often equipped with an N-terminal Sec-dependent signal peptide. Many are targeted to the cell wall by a C-terminal LPXTG amino acid motif. Covalent attachment to the cell wall is achieved by enzymatic cleavage between the threonine and glycine residues followed by covalent linkage of threonine to the peptidoglycan of the cell wall (Navarre & Schneewind, 1999). In some cases, the

16

importance of receptins in staphylococcal infections has been investigated in animal models, and a few of them have been used for immunisation in animal infection models to explore their potential to provoke an immune response (Flock, 1999). Some success has been achieved, but the nature of the interactions between *S. aureus* and its host is complex and further studies are required to elucidate which components are critical for virulence.

Table 2. *Characterised receptins of* Staphylococcus aureus

Receptin	Ligand(s)	LPXTG motif	Reference(s)
Coa	Fibrinogen Prothrombin	-	Kaida *et al.*, 1987 Kawabata *et al.*, 1986
Efb	Fibrinogen	-	Boden & Flock, 1994
ClfA	Fibrinogen	LPDTG	McDevitt *et al.*, 1997
ClfB	Fibrinogen Cytokeratin 10	LPETG	Ni Eidhin *et al.*, 1998 O'Brien *et al.*, 2002
Spa	IgG, von Willebrand factor	LPETG	Uhlen *et al.*, 1984 Hartleib *et al.*, 2000
Sbi	IgG, β_2-glycoprotein I	-	Zhang *et al.*, 1998 Zhang *et al.*, 1999
FnbpA	Fibronectin, Fibrinogen	LPETG	Signas *et al.*, 1989 Wann *et al.*, 2000
FnbpB	Fibronectin	LPETG	Jonsson *et al.*, 1991
Bbp	Bone sialoprotein	LPETG	Tung *et al.*, 2000
Cna	Collagen	LPKTG	Patti *et al.*, 1992
EbpS	Elastin	-	Park *et al.*, 1999
vWbp	von Willebrand factor Prothrombin	-	Bjerketorp *et al.*, 2002 Friedrich *et al.*, 2003
Sak	Plasminogen	-	Sako & Tsuchida, 1983
StbA	Transferrin	LPKTG	Taylor & Heinrichs, 2002
IsdB	Hemoglobin	LPQTG	Mazmanian *et al.*, 2003
HarA	Haptoglobin	LPKTG	Dryla *et al.*, 2003
Map/Eap	Broad-spectrum recognition	-	Jonsson *et al.*, 1995
Emp	Broad-spectrum recognition	-	Hussain *et al.*, 2001
Ebh	Fibronectin	-	Clarke *et al.*, 2002

Protein export

Extracellular proteins are synthesised in the cytoplasm and have to be translocated to their site of action by specific targeting and transport mechanisms. Three major systems have been described for secretion of proteins across the bacterial cytoplasmic membrane: Sec (secretion pathway), SRP (signal-recognition particle), and Tat (twin-arginine translocation). Components of the Sec and SRP pathways are present in every organism (prokaryotic and eukaryotic) for which a genome has been completed. Homologues of the Tat family are not ubiquitous among bacteria, nor are they found in yeast or animals. They are, however, closely related to the pH-dependent protein import into the thylakoid lumen in plant chloroplasts (Pallen *et al.*, 2003).

Extracellular proteins are synthesised as precursors with signal peptides (cleavable) or signal anchors (non-cleavable). Most cleavable signal peptides have a tripartite structure – a net positive charged 1-5 residue n-domain, an α-helical 7-15 residue hydrophobic core (h-domain), and a polar 3-7 residue c-domain. Due to the positive charge of the n-domain, the N-terminus is retained in the cytosol since the electrical potential of the cytoplasmic membrane is negative on the inside. The h-domain is thought to insert into the lipid bilayer and may form a hairpin-like structure to facilitate insertion of the signal peptide into the translocation channel. The cleavage site is contained within the c-domain. The -1 and -3 residues relative to the cleavage site are recognised by the signal peptidase enzyme, which is responsible for signal peptide cleavage on the periplasmic side of the membrane. Despite the common general structure of signal peptides, no conservation in amino acid sequence is observed (von Heijne, 2002). Sec-dependent signal peptides of Gram-negative and Gram-positive bacteria differ mainly in length (24 and 29-31 amino acids, respectively), Gram-positive signal peptides being longer due to extensions in all of the three domains (von Heijne & Abrahmsen, 1989). Bacterial lipoproteins represent a special case of Sec-dependent signal peptides. The c-domain of lipoprotein signal peptides end with a 'lipobox' with the consensus sequence L(A,S,I)(G,A)-C ('-' indicates cleavage site). Specific lipoprotein signal peptidases cleave the lipobox in a reaction, where a lipid anchor is simultaneously added to the N-terminal cysteine in the mature chain. Signal peptides of the Tat pathway have an essentially invariant twin-arginine motif (providing the basis for the Tat designation) in the extended n-domain. The consecutive arginine residues are part of a consensus sequence contained within the signal peptide, S-R-R-x-F-L-K, where x is a polar amino acid or a glycine, and the other residues are found at frequencies exceeding 50%. Often, a positively charged residue is present in the c-region, suggested to act as a 'Sec-avoidance signal'. Also, Tat signal peptides have a lower average hydrophobicity of the h-region, due to the increased prevalence of glycine and threonine residues combined with decreased prevalence of leucine residues, relative to Sec signal peptides (Berks et al., 2003). In most cases, bacterial cytoplasmic membrane proteins do not have cleavable signal peptides but are instead recognised by the extended hydrophobic regions that become the transmembrane helices in the mature protein. Non-cleavable signal peptides can anchor proteins in the membrane, when they contain a sufficiently hydrophobic h-domain (Martoglio & Dobberstein, 1998). Based on studies of *E. coli*, cleavable signal peptides are rarely found in bacterial cytoplasmic membrane proteins, but it is very difficult to distinguish between cleavable and non-cleavable signal peptides (Broome-Smith et al., 1994).

Sec and SRP pathways

The Sec and SRP pathways merge at the Sec translocase (figure 3). Although exceptions have been identified, the Sec pathway mediates protein translocation in a post-translational manner, while the SRP pathway mainly mediates co-translational translocation. Chaperones, like SecB and SRP, keep the mature protein in an unfolded state. The post-translational pathway translocates mostly water-soluble proteins. Substrates with an N-terminal signal peptide are recognised by the SecA ATPase associated to the Sec core in the cytoplasmic membrane. The

core is made up of a SecYEG heterotrimer through which the substrate protein is transported by threading, requiring energy input in the form of ATP and the proton motive force (PMF). Additional proteins, SecD, SecF, and YajC, are membrane proteins involved in translocation, although their exact functions are not known. They appear to be associated as a heterotrimeric complex, and even though no catalytic activity has been assigned to the complex, cells depleted of SecD and Sec F are greatly deficient in protein export. In the post-translational pathway, insertion of integral membrane proteins into the cytoplasmic membrane normally also involves the SecYEG translocase *via* the SRP receptor FtsY associated with the translocase (Driessen & van der Does, 2002). Recent research shows that transmembrane helices move from the translocase *via* the translocase-associated protein YidC, into the membrane (Scotti *et al.*, 2000). Some inner membrane proteins are targeted directly to YidC, *e.g.* filamentous phage coat protein VIII (Samuelson *et al.*, 2000).
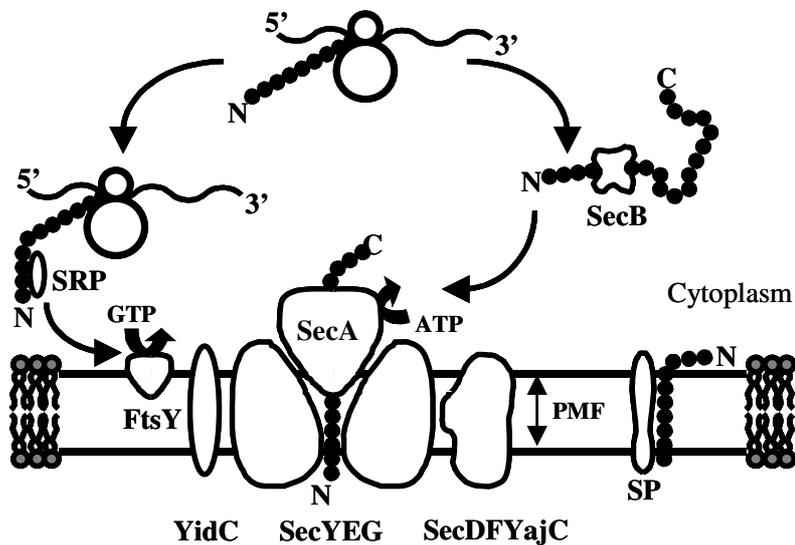
Figure 3. Schematic overview of the Sec and SRP pathways, see text for details. SP denotes signal peptidase. Adapted from Driessen & van der Does, 2002.

The Tat pathway

The majority of Tat substrates are fully folded redox proteins containing cofactor molecules. The physiological role of the bacterial Tat pathway is to allow the cell to complete the maturation of periplasmic proteins in the cytoplasm before they are exported, while maintaining the permeability barrier of the membrane to ions, especially protons. The integral membrane proteins TatA, TatB, and TatC are assumed to form the Tat translocase and transport is known to be energised by the PMF. The Tat proteins have been proposed to form two types of complex in *E. coli* – the Tat(A)BC (indicating a minor proportion of TatA protein) and TatA(B) (composed predominantly of TatA protein) complexes. Supposedly, the TatA(B) complex form the transport channel through which the substrate proteins are translocated, and Tat(A)BC form the substrate receptor complex, which

recruits the TatA(B) channel upon substrate binding. The Tat pathway is reviewed extensively in Berks *et al.*, 2003

Transport across the outer membrane

In Gram-negative bacteria, proteins located in the outer membrane (OM) or secreted into the surrounding environment have to cross the periplasm and the OM. Export systems mediating such translocation have to be able to handle structured polypeptide chains or completely folded proteins and, since energy sources are not known to be present at the OM, they have to be self-energized or have mechanisms to use the inner membrane (IM) energy scources. The following paragraphs on OM translocation refer to a review by Thanassi and Hultgren (2000).

The pathways evolved for conducting the task of OM translocation can be organised into six groups, four of which rely on a preceding Sec-dependent translocation through the IM. The autotransporter secretion pathway involves the translocation of proteins with a C-terminal β-domain, which presumably inserts into the OM as a β-barrel pore structure through which the α-domain, constituting the secreted mature protein and which can be cleaved off or not, can pass. Proteins exported by this pathway include proteins with various functions like proteases, toxins, adhesins, and invasins, and the prototypical member of the autotransporter family is the IgA1 protease of *Neisseria gonorrhoeae* (Pohlner *et al.*, 1987). The chaperone/usher pathway is named by the components necessary for this type of export – a periplasmic chaperone and an OM translocation channel termed an uscher. It is used for the assembly and secretion of a broad range of adhesive virulence structures on the bacterial surface, such as P and type 1 pili of *E. coli*. The periplasmic chaperone acts on release, folding, and capping of the pilus subunits to prevent premature interactions. Dissociation of the chaperones is triggered by interaction with the usher, through which the linear fiber of folded subunits can passage and adopt its helical conformation at the cell surface. The third Sec-dependent pathway through the OM is the type II secretory pathway, responsible for the export of a wide variety of enzymes and toxins. This system is more complex than those previously described, requiring between 12 and 16 accessory proteins, together referred to as the secreton. It is linked to the biogenesis of type 4 pili, as demonstrated by the sharing of many features, such as 10 homologous proteins, suggesting a common evolutionary origin. The homologous secreton component GspD belongs to the secretin superfamily, members of which are also required for, among other things, filamentous phage biogenesis (pIV). Many components of the secreton are associated with the IM and a model has been proposed, where ATP is used to effectuate conformational changes in the IM proteins, which are subsequently transmitted to the periplasmic domains and to the OM. The Sec-dependence of the type IV secretory pathway has not been established, although export of pertussis toxin (PT) of *Bordetella pertussis* by a type IV system most likely occurs by a two-step event, where the toxin subunits first traverse the IM *via* the Sec pathway. The type IV system has mainly been characterised by studies of the homologous T-DNA translocating VirB system of *Agrobacterium tumefaciens*, and the presence of homologous systems has so far only been identified in a few organisms (Burns, 1999).

20

The Sec-independent pathways include the type I, or ATP-binding cassette (ABC) protein exporters, and type III secretion. The type I system is employed by a wide range of Gram-negative bacteria for secretion of toxins, proteases, and lipases, and is represented by the prototypic *E. coli* α-hemolysin exporter (Thanabalu *et al.*, 1998). Proteins are transported directly from the cytoplasm to the outside of the cell. The secretion apparatus is made up of one IM ATP-binding ABC exporter, one IM-anchored periplasmic protein, and an OM protein. Instead of N-terminal signal peptides, the type I substrates possess a C-terminal secretion signal recognised by the secretion apparatus. During translocation, the IM-associated periplasmic protein interacts with both the ABC exporter and the channel-forming OM protein, allowing secretion without a periplasmic intermediate. The type III pathway, identified in a number of animal and plant pathogens, enables translocation of virulence proteins into the cytosol of target eukaryotic cells. Related secretion systems exist for flagellar biogenesis. About 20 components are required to form a large structure that spans both bacterial membranes and possibly the host cell membrane. The type III secretory system shares one homologous component with the type II system – a member of the secretin family of channel-forming proteins for export through the OM. The best-studied type III system is the translocation of approximately 12 proteins encoded by bacteria of the genus *Yersinia*, the *Yersinia* outer proteins (Yops), into target host cells (Anderson & Schneewind, 1999).

Protein secretion in Gram-positive bacteria

Protein secretion has mainly been studied in the Gram-negative bacterium *E. coli*. Unlike the case in Gram-negative bacteria, secreted proteins of Gram-positive bacteria only have to cross the cytoplasmic membrane to enter the extracellular environment. The general pathways of protein translocation (Sec, SRP, and to some extent Tat) are present in Gram-positive bacteria, although some features differ from the Gram-negative variants like the Sec-directing signal peptides and the absence of a SecB homologue in Gram-positive bacteria (van Wely *et al.*, 2001). *Bacillus subtilis* has arisen as a model organism for protein translocation in Gram-positive bacteria and possesses systems with some relation to the type I and type II secretion pathways of Gram-negative bacteria. For instance, ABC transporters are used for the export of ribosomally synthesised bacteriocins and pheromones. However, in contrast to type I substrates, these peptides are synthesised with defined N-terminal signal peptides. In addition, the *comG* and *comC* gene products of *B. subtilis* have N-terminal signal peptides homologous to those contained in the secretin superfamily (Chung & Dubnau, 1998; Pugsley, 1993; Tjalsma *et al.*, 2000). A few additional pathways have been proposed in Gram-positive organisms. An equivalent of the Gram-negative type III secretion system was identified in *Streptococcus pyogenes* – the cytolysin-mediated translocation (CMT) (Madden *et al.*, 2001). A cytotoxicity-triggering effector, *S. pyogenes* NAD-glycohydrolase (SPN), capable of producing a potent eukaryotic second messenger molecule, was shown to be translocated through a pore of cholesterol-dependent cytolysin streptolysin O (SLO) into the host cell cytosol. The high degree of conservation of cholesterol-dependent cytolysins among diverse species of Gram-positive pathogens suggested that CMT is a widespread

phenomena. Further, the early-secreted antigen target 6 kDa protein (ESAT-6) of *Mycobacterium tuberculosis* is secreted to the media despite the lack of a typical Sec-dependent signal peptide, and more than 20 homologues have been identified in the genome (Gey Van Pittius *et al.*, 2001; Pallen, 2002). Several of the genes lie in clusters that also encode large membrane-bound ATPases, which are proposed to represent a new secretion system in Gram-positive bacteria. A more recent report has shown that individual genes of the region of difference 1 (RD1) of *M. tuberculosis* are required for ESAT-6 export. The RD1 is one of several genetic regions that differ between an avirulent *Mycobacterium bovis* strain and virulent *M. tuberculosis* (Guinn *et al.*, 2004).

Regulation of gene expression

Both prokaryotic and eukaryotic organisms regulate the expression of their genes to adapt their phenotypes to particular extra- and intracellular environments. Three factors determine which genes that are expressed and the extent to which they are expressed – the concentration of mRNA transcript (determined by the rates of synthesis and degradation), the frequency with which the mRNA is translated, and the stability of the protein product. The most common form of gene regulation in prokaryotes is to control the initiation of transcription. This is accomplished by alterations, by different strategies, of RNA polymerase↔promoter recognition and interaction, *e.g.* by binding of regulatory proteins at or near the polymerase binding site for inhibited or enhanced transcription. Also, control of the termination of mRNA transcription, *e.g.* attenuation, provides an important way of regulating mRNA levels. Post-transcriptional regulation can be achieved by preventing ribosomes from initiating translation, *e.g.* by binding of a protein to a region of the mRNA overlapping with the ribosome binding site.

   Methods that provide a means for abolishing or controlling gene expression can be used to study essentiality and function of the gene products. Different gene knock-out methods are used in various bacterial systems to achieve functional inactivation of gene products, and have provided valuable information on the function of specific genes (*e.g.* Cheung *et al.*, 1995; de Lorenzo & Timmis, 1994). However, the screening of individual mutants to identify a set of genes, *e.g.* those of a pathogenic organism that are implicated in host survival or virulence, would be very time consuming. To overcome this limitation, new genome-wide techniques for the identification of *in vivo* expressed genes have been developed. Using *in vivo* expression technology (IVET), a positive selection of host-induced genes can be achieved. A gene fusion library is generated by insertion of random chromosomal fragments upstream of a gene that is required for survival in an animal model. Isolated clones represent genes expressed *in vivo*, but not all of these may be essential for infectivity (Angelichio & Camilli, 2002). Instead, the negative selection of infection related genes can be accomplished by using signature tagged mutagenesis (STM). The technique involves insertional mutagenesis and the labelling of each mutant with a unique signature tag, which can be used to identify the absent mutants by PCR amplification after a selective screen through an animal or cell culture model (Shea *et al.*, 2000). Moreover, genes with altered transcription rates caused by exogenous or endogenous stimuli

can be identified using differential display – a direct detection of mRNA profiles from an organism grown under various environmental conditions. Also, gene expression can be monitored on translational level rather than on transcriptional level using proteomic methods such as two-dimensional gel electrophoresis (2D-PAGE). This reveals effects of post-transcriptional, translational and post-translational regulating events on gene expression. The identification of *in vivo* expressed genes, using the methods mentioned above or others, is reviewed in Hautefort & Hinton, 2002.

Antisense RNA regulation

Naturally occurring antisense RNAs

Gene regulation by antisense RNAs is a naturally occurring event in all three kingdoms of life. Natural antisense RNA systems were first demonstrated in *E. coli* plasmids ColE1 (Tomizawa *et al.*, 1981) and R1 (Stougaard *et al.*, 1981) and thereafter, antisense RNAs have mainly been identified in prokaryotic organisms, often in their accessory elements – plasmids, phages, and transposons (Wagner & Simons, 1994). The name implies that these regulatory RNAs are complementary to the sense (coding) RNA and hence can bind to their target RNA and form a duplex. Base pairing with the target RNA affects gene expression at the post-transcriptional level, mostly by inhibition, and thereby the biological function. Antisense RNAs act through steric hindrance or degradation of target sequences. In prokaryotes, naturally occurring antisense RNAs are often small, diffusible, untranslated and highly structured regulatory RNAs. Typically, they are 35-150 nucleotides long and form 1-4 stem-loops (Brantl, 2002). Usually, the loops are the sites of critical interactions for the initiation of pairing to the target RNA. A YUNR (Y=pyrimidine, U=uracil, N=any nucleotide, R=purine) motif, similar to tRNA motifs that fold into so-called U-turn structures (Franch & Gerdes, 2000) is often found in these loops. Such U-turn structures provide a scaffold for the rapid interaction with complementary RNA. High binding rates are important for effective regulation by antisense RNA, but *in vitro* studies indicate that antisense RNAs do not pair rapidly to completion. Instead, complementary folded RNAs often arrest at the stage of stable binding intermediates due to topological constraints – great energy barriers have to be overcome to unwind the stable intramolecular stem structures (Wagner & Brantl, 1998).

Complementary structures are observed in most, but not in all, of the cognate target RNAs. Only a small number of antisense RNAs have been found to be of chromosomal origin, and they are mainly detected in bacteria. Most of these are *trans*-encoded, *i.e.* transcribed from distant loci, and reveal only partial complementarity to their targets. Regulation is most likely a result of partial or imperfect duplexes, and many appear to interact with the 5' untranslated region of their targets. Some are known to interact with multiple targets and can regulate a set of mRNAs, that are encoded by unlinked genes (Carpousis, 2003). An example of a bacterial *trans*-encoded regulatory RNA that acts on multiple targets is the antisense RNAIII of *S. aureus*. It also serves as an example of the rare case of activation by antisense RNA. The 512 nucleotide long RNAIII transcript is the effector molecule of the accessory gene regulator (*agr*) locus, known to control

many virulence determinants of *S. aureus*. Generally, staphylococcal genes encoding secreted proteins are activated by the presence of RNAIII, while genes encoding surface proteins are repressed. The translation of α-hemolysin (*hla*) is positively controlled by the competition of RNAIII with an inhibitory intramolecular RNA secondary structure that sequesters the *hla* ribosome-binding site. Upon RNAIII hybridisation to the *hla* mRNA, the ribosome binding site becomes accessible and translation of *hla* is enabled (Morfeldt *et al.*, 1995). Repression of the *spa* gene, encoding IgG-binding protein A, is also effectuated by RNAIII activity. It is believed that RNAIII acts, either directly or indirectly, at the transcriptional level, although it is possible that RNAIII affects the stability of the *spa* transcript. The mechanism by which RNAIII activates or inhibits expression of other virulence factors remains unknown (Johansson & Cossart, 2003).

Antisense RNAs can act as gene regulators by RNA-RNA interactions, but protein factors may also play a role by binding to antisense and target RNAs. In *E. coli*, Hfq, which is an RNA-binding protein related to the Sm proteins (small nuclear ribonucleoprotein-associated proteins) in eukaryotes, interact with many of the chromosome-encoded antisense RNAs to form ribonucleoproteins. A eukaryotic Sm-like sequence motif is present in many antisense RNAs, and Hfq recognises this motif both within antisense RNA and target mRNA molecules, and it can stabilise an interaction (Moller *et al.*, 2002; Zhang *et al.*, 2002).

Artificial antisense RNAs

The discovery of natural antisense RNA systems led to the idea of using artificial antisense RNAs as a means to control gene expression by specific inhibition of target RNA function. The first reports to demonstrate such an effect in prokaryotes and eukaryotes appeared in 1984 (Coleman *et al.*, 1984; Izant & Weintraub, 1984). Since then, two different approaches have evolved. First, the use of exogenous, short (15-25 bases), synthetic oligonucleotides, designed to hybridise to specific sequences of target mRNAs *via* Watson-Crick base pairs and predicted to act mainly by degradation of targeted mRNA by an RNase H mediated mechanism (Varga *et al.*, 1999). Second, the use of endogenous antisense RNAs, transcribed from DNA cloned in antisense orientation, which possibly fold into higher order structures and act on their target sequences as such (exemplified by one of the earliest artificial systems reported in Coleman *et al.*, 1984 and Green *et al.*, 1985). Both methods have met with successes and setbacks. Important factors to consider, when using oligonucleotides, are *e.g.* stability, toxicity and delivery, while antisense RNA approaches are more difficult to control because these molecules are often longer and may adopt complex structures. Also, the inhibitory effect may depend on whether intracellular host proteins will support or counteract the binding between antisense and target RNA molecules. Continuous developments and modifications of both lines are being made, mainly in the light of new insights in natural occurring regulatory RNAs (Delihas *et al.*, 1997; Engdahl *et al.*, 1997). Recently, genomic approaches to express large numbers of diverse antisense RNAs have been applied in prokaryotic (Forsyth *et al.*, 2002; Ji *et al.*, 2001; Wang & Kuramitsu, 2003) and eukaryotic (De Backer *et al.*, 2001) systems with the objective to identify genes critical for growth. Basically, in these systems, an inducible promoter drives the transcription of antisense RNAs from the inserted

24

random chromosomal fragments in a shotgun library. These antisense RNAs can act on their corresponding sense RNA transcript and down-regulate gene expression, which is detected by the loss of protein function.

Present investigation

Aims of the study

The aim of this study was to use and develop the shotgun phage display technique to identify and characterise extracellular proteins encoded by a bacterium, primarily those of *S. aureus*. The specific aims were to:

- evaluate shotgun phage display as a method to study the binding of *S. aureus* proteins to an *ex vivo* biomaterial

- design a specialised phage display method for the large-scale identification of genes encoding extracellular proteins in a bacterium

- explore the possibility to combine this method with an element for gene silencing

Results and discussion

General considerations of shotgun phage display libraries

The shotgun phage display technique is powerful in many ways. Perhaps its greatest advantage lies in the physical linkage between genotype and phenotype, why large numbers of different clones can easily be investigated. Furthermore, shotgun phage display libraries can be used in complex panning experiments, where traditional screening is not an optional method, *e.g.* against complex ligands like biomaterials or eukaryotic cells, or even in *in vivo* pannings. But even though, theoretically, all proteins encoded by a bacterium should be represented several times in a phage library, there are features of the system that may result in a bias of the clones isolated in a particular panning experiment. For instance, in panning an *S. aureus* library against a biomaterial, clones derived from two proteins were clearly dominating (**IV**, see below). These proteins were represented by several different clones, but certain clones appeared more often than others. When panning an *S. aureus* library against human serum it was observed that certain interactions were more frequently identified than others (Jacobsson & Frykberg, 1998). Seemingly, the high occurrence of IgG-binding clones masked the binding of proteins to less abundant serum components. Also, certain binding domains of a protein may dominate over other domains, even though they have affinity for the same ligand. For example, Sbi contains two IgG-binding domains (unpublished results) and vWbl have several von Willebrand factor-binding domains. Still, in both cases, only one ligand-binding domain is isolated in a standard panning experiment (Zhang *et al.*, 1998; Nilsson *et al.*, 2004). These biases can appear at different stages:

i)   It is conceivable that not all clones replicate equally well and/or that the expression of certain fusion proteins may be toxic to *E. coli*, resulting in a diminished representation of these clones in the library.

ii)   The display of fusion protein may depend on various factors, such as the expression of the protein, the compatibility with the *E. coli* protein translocation machinery and the periplasmic space, and the efficiency of incorporation into the phage coat.

iii)   Which clones that are isolated in a panning experiment is influenced by features like the affinity of the displayed polypeptide for the ligand, the concentration and accessibility of the ligand, and the conformation of the displayed polypeptide.

iv)   Different elution conditions may affect which clones that are present in the amplified phage stock for subsequent pannings. Gentle elution conditions may result in that a fraction of clones are not released from the ligand, while the use of harsh conditions could impair phage infectivity.

In **V**, the impact of different elution conditions on results from panning an *S. aureus* library against human serum is evaluated. The basis for the experiment was that there are known interactions between staphylococci and mammalian proteins for which the staphylococcal receptin has never been detected by the phage display technique, possibly because the binding to the ligand is not interrupted using standard elution conditions. Besides the usual low pH (pH 2) elution, 5M urea, 40% ethanol, 80% methanol, 5% Tween, 40% ethylene-glycol, and 2M NaCl were tested. Elution by low pH resulted in the highest frequency of correct clones, but one interaction never identified under low pH conditions was identified by elution with 2M NaCl. This suggests that the use of different elution conditions may result in the identification of new interactions. Besides different non-specific elution conditions, including others not tested in this study like alkaline buffers and proteases, specific elution *e.g.* by competitive elution may be used. In addition, elution can be completely omitted by proceeding to infection of *E. coli* directly in the microwell after washing. This approach was applied in pannings of a *Staphylococcus lugdunensis* phage library against von Willebrand factor. One domain of a repetitive von Willebrand factor-binding protein (vWbl) was identified when low pH elution conditions were used, but by using the 'elution by infection' method the same domain as well as an additional one were isolated (Nilsson *et al.*, 2004). Elution of bound phage is discussed in a review by Smith and Petrenko (1997).

As stated above, the conformation of the displayed polypeptide is of importance for the outcome of a panning experiment. Particularly, the correct folding and post-translational modifications of eukaryotic proteins can be difficult to achieve in phage display systems. When panning gene libraries made from cDNA of fibronectin and von Willebrand factor against staphylococcal receptins, no enrichment of clones was obtained. However, panning the same libraries against specific antibodies resulted in enrichment for certain domains of the eukaryotic proteins, which represent epitopes recognised by the polyclonal antibodies. This suggests that the bacterial receptins recognise domains of the eukaryotic proteins,

which are not correctly displayed on the phage surface (unpublished results discussed in **V**).

Accessibility of staphylococcal target proteins on biomaterial (IV)

S. aureus is a common coloniser of human skin and mucosa, but it is also an important pathogen in human diseases. In many countries, hospital-acquired infections by multi-resistant strains are common. In the United States, *S. aureus* is one of the three leading causes of hospital-based bloodstream infections within the acute care hospitals. The bacteria invade *via* broken skin or mucous membranes and can cause a wide variety of diseases. Patients at greatest risk are those who are immuno-compromised, those with chronic illnesses, and those whose treatment requires an invasive device. *S. aureus* and *Staphylococcus epidermidis* are the most frequently isolated bacteria from biomaterial-associated infections (von Eiff *et al.*, 2002). Different definitions for 'biomaterial' have been proposed; one example is 'a nonviable material used in a medical device, intended to interact with biological systems'. They can be of different origin, such as metals, ceramics, or polymers, and used during a limited time, *e.g.* intravenous catheters, or as permanent replacements, *e.g.* artificial hip joints. Implanted materials very quickly acquire a layer of plasma proteins, such as albumin, fibrinogen, fibronectin, and vitronectin, in a dynamic adherence process (Vroman *et al.*, 1980). The ability of staphylococci to adhere to the proteins deposited on biomaterials is an important factor for pathogenesis. Many staphylococcal receptins with affinity for plasma and extracellular matrix proteins have been identified and characterised, and they present a quite complex picture. There are several receptins that interact with the same ligand, and some receptins can interact with more than one ligand (table 2). The expression of these proteins under different conditions is elaborately regulated (Cheung *et al.*, 2004). The binding of *S. aureus* to various pure ligands coated on polymer surfaces has been demonstrated (table 3), but which receptins that are important for the *in vivo* adherence of *S. aureus* to host proteins is not completely elucidated.

Table 3. *Plasma and extracellular matrix proteins shown to promote* S. aureus *attachment to coated polymer surfaces*

Host protein	Reference(s)
Fibronectin	Vaudaux *et al.*, 1984
	Herrmann *et al.*, 1988
Fibrinogen	Herrmann *et al.*, 1988
Collagen	Switalski *et al.*, 1993
	Mohamed *et al.*, 1999
Vitronectin	Fuquay *et al.*, 1986
von Willebrand factor	Herrmann *et al.*, 1997
Laminin	Fuquay *et al.*, 1986
	Herrmann *et al.*, 1988
Thrombospondin	Herrmann *et al.*, 1991

With traditional methods, the interactions between *S. aureus* and biomaterials are studied by analysis of one specific protein at a time. For example, genes encoding

proteins with known binding properties are inactivated and the binding of the mutant to a biomaterial is investigated. Another way is to identify which proteins that are adsorbed to a biomaterial, after which the binding of *S. aureus* to pure fractions of these proteins can be tested. However, by using this approach, the identification of minor components on the biomaterial may fail. With a phage display library that displays all proteins encoded by an organism, potentially, all possible interactions to the biomaterial can be found in one single experiment. This approach also discriminates between surface accessible proteins and those hidden within the protein layers on the biomaterial. With the objective of finding novel interactions between staphylococcal receptins and components of biomaterials, an *S. aureus* shotgun phage display library was panned against an *ex vivo* biomaterial, a central venous catheter (CVC) removed from a patient after 2 days of implantation (**IV**). Although panning against a complex mixture of ligands, like human serum coated onto microwells, has been reported previously (Jacobsson & Frykberg, 1998; **V**), *in vivo* coated materials have never been investigated with this technique. From the panning, two main groups of clones were identified – clones containing the fibrinogen-binding domain of Coa and clones encoding either the β_2-glycoprotein I (β_2-GPI)-binding domain of Sbi alone or together with the IgG-binding domain. No novel receptins were identified, but the isolation of merely β_2-GPI-binding domains of Sbi suggested that β_2-GPI was present and accessible on the biomaterial. As far as we know, β_2-GPI has never been reported to adsorb to biomaterials. To confirm this finding, proteins were isolated from CVCs, which had been removed from different individuals and at different times, *i.e.* 6 days, 9 days, 13 days, and 18 days after insertion. In a Western blot, antibodies against human fibronectin, fibrinogen, IgG, and β_2-GPI showed that these proteins were recovered from the different biomaterial samples, with the exception that IgG could not be detected on the 9 day-material. Also, the accessibility of these proteins to staphylococcal receptins displayed on phage was established by the panning of specific phagestocks against the different CVC samples. Three phagestocks (Nbm69, Nbm29, and Nbm21) exposing the fibrinogen-binding domain of Coa, the IgG-binding and von Willebrand factor-binding domain of Spa, or the β_2-GPI-binding domain of Sbi, respectively, were shown to bind to all samples. The presence of β_2-GPI on biomaterial implicates a hitherto unknown opportunity for *S. aureus* to adhere to biomaterials. Protein Sbi has been shown to be a surface-associated protein of *S. aureus*, but lacks an LPXTG motif for covalent attachment to the cell wall (Zhang *et al.*, 1998). The presence of β_2-GPI on biomaterials is also of interest in the context of autoimmune disease. The antiphospholipid syndrome (APS) is characterised by the presence of pathogenic autoantibodies presumably directed against β_2-GPI (McNeil *et al.*, 1990), and for which the cause of appearance has not been established although an association with microbial pathogens has been demonstrated (McNeil *et al.*, 1991). Also, it has been suggested that a conformational change in β_2-GPI may be responsible for the exposure of autoepitopes (Borchman *et al.*, 1995; Chamley *et al.*, 1999; Wang *et al.*, 2000), and the binding of other plasma proteins to polymer surfaces has been shown to induce conformational changes (Barbucci & Magnani, 1994). Thus, one can speculate that the binding of β_2-GPI to biomaterial may be one route of induction of autoantibodies in APS.

28

*Signal sequence shotgun phage display (SS phage display) (**I** and **II**)*

Extracellular proteins are the means by which bacteria acquire nutrients, sense various stimuli in the environment, and adhere to eukaryotic host cells, among many other things. The identification of the whole subset of these important proteins in a bacterium is of interest, and several methods have been developed to accomplish this task (mentioned in **I**). Many involve the individual screening of clones, which is very time consuming. Phage display has the advantage that screening of many different clones can be performed in one experiment. The unusual non-lytic lifecycle of the filamentous phage, when phage particles are assembled in the *E. coli* cell membrane, allows for the development of the shotgun phage display technique to isolate 'all' extracellular proteins in a bacterium (**I**). In the following presentation and discussion, 'signal sequence' will refer to the DNA encoding a signal peptide. The pG8SAET phagemid vector was modified so that the *spa* signal sequence was removed and gVIII replaced by ΔgIII (C-terminal half, mainly D3), generating pG3DSS (schematic picture presented in **I**). The *spa* promoter and E-tag units were still intact. In order for a fusion protein expressed from pG3DSS to be incorporated into the phage coat, it has to be translocated to the *E. coli* cell membrane, where phage assembly occurs. In this system, the signal sequence has to be provided by the foreign insert, *i.e.* by fragments that contain genes encoding extracellular proteins. The general isolation of fusion proteins produced by such clones is achieved by affinity selection for clones displaying the E-tag (see 'Shotgun phage display' under 'Introduction'). This system was first applied to *S. aureus*, for which the genome sequence was available. Panning experiments with two cycles were performed, and after the first cycle about 50% of the clones were considered correct, *i.e.* with inserts encoding extracellular proteins, while the rest constituted background clones. In the second panning cycle, the amount of correct clones approached 100%, but a few clones were represented more frequently than others and tended to dominate. Both secreted proteins and membrane proteins were identified, including toxins, enzymes and transport proteins. A striking feature of the isolated clones was that they seemed to carry their own promoter and showed, in general, differential but low protein expression with regard to the E-tag. A limitation of the system is that the identification of the isolated clones requires individual handling, why only a fraction of the isolated clones can be analysed.

The first attempt to make an SS phage display library resulted in a rather small library, 8×10^5 unique clones, and was generated from DNA of a Gram-positive organism with a genome size of about 2,8 Mb. Later, the method has also been applied on the symbiotic Gram-negative bacterium *Bradyrhizobium japonicum* (**II**), as well as the probiotic bacterium *Lactobacillus reuteri* (Wall *et al.*, 2003), and *Staphylococcus epidermidis* (unpublished results). These libraries were considerably larger, 4×10^7 - 4×10^8 clones, while the genome sizes were comparable to *S. aureus* except for *B. japonicum* with a 9,1 Mb genome. A summary of the results obtained with the different libraries is presented in table 4. In general, the highest incidence of correct clones and, especially, unique clones was found with the *B. japonicum* and *S. epidermidis* libraries. For the *S. aureus* and *L. reuteri* libraries there was a higher incidence of background clones in the first panning cycles and a stronger dominance of certain clones in the second panning compared

to the other two libraries. *B. japonicum* can be expected to encode a higher number of extracellular proteins than the other organisms because of its substantially larger genome, and also, as a Gram-negative organism, it has additional cell compartments in which many extracellular proteins carry out important functions, *i.e.* the periplasm and the outer membrane A comparison between the results obtained with the different libraries should be made with the awareness of the different number of clones that have been analysed. Since there is a tendency of certain clones to dominate, to various extents, in all libraries, presumably it becomes increasingly difficult to find new unique proteins when additional clones are analysed, as observed for the *L. reuteri* library (Wall *et al.*, 2003). Putative lipoproteins were identified in all libraries except the one of *S. epidermidis*. Also, clones with inserts encoding signal peptides with putative Tat consensus sequences were identified with the *B. japonicum* and *L. reuteri* libraries, but the presence of components of the Tat translocase has not been established for these organisms and no information on the function of the proteins in question to support translocation *via* the Tat pathway, was available. An interesting finding is the isolation of clones encoding several transmembrane domains, since they can be expected to be more firmly attached to the *E. coli* cell membrane. A clone encoding 9 putative transmembrane domains was isolated from the *S. aureus* library and shown to promote display of the E-tag at the phage surface.

Table 4. *Summary of results from signal sequence shotgun phage display libraries*

Origin of foreign DNA fragments	A: Total number of clones investigated	B: Correct clones (% of A)	C: Unique clones (% of B)
S. aureus	105	84	42
B. japonicum	182	92	79
L. reuteri	263	62	33
S. epidermidis	64	89	61

With the many genome sequences and powerful bioinformatic tools available today, *in silico* scanning for extracellular proteins can be performed instead of using experimental methods. Still, the algorithms developed to execute these predictions are not likely to cover all true phenomenona as they are based on the limited information we possess. Thus, interesting information may be omitted by such handling, as illustrated by the following examples of clones that were isolated during the work with the *B. japonicum* and *S. epidermidis* libraries. A clone encoding a small protein with four characteristic cystein residues was identified from the *B. japonicum* library, apparently overlooked by the open reading frame-finding software. From scanning this protein against the microbial genome databases, another 16 homologous proteins were identified, 10 from *B. japonicum* and 6 from the closely related bacterium *Rhodopseudomonas palustris*. We denoted this group of proteins the 4C-protein family, which apart from the conserved cystein residues also share some other amino acid motifs. They were all predicted to have cleavable signal peptides by the SignalP algorithm (http://www.cbs.dtu.dk/services/SignalP/; 16-Apr-2004). At present, only one of the rhizobial 4C-proteins is annotated in the databases, while all proteins identified in *R. palustris* are annotated, but no functions have been proposed. Recently, yet

another member of the family was recognised in the protein databases – a protein identified in a genome screening of microbial populations in the Sargasso Sea (Venter *et al.*, 2004). This protein cannot be ascribed to a specific organism, but, at least, it is not identical to any of the previously mentioned proteins. It is also worth mentioning that 12 of the isolated clones from the *B. japonicum* library could not be identified in the published and publicly available genome of strain USDA110 (Kaneko *et al.*, 2002), but 5 of these were detected in the database for the uncompleted genome of strain USDA438, to which we had access. This suggests that, due to genetic variations between strains, an *in silico* identification of open reading frames in one strain may not be completely applicable to another un-sequenced strain, but by experimental methods such as the SS phage display these genomic variations can be detected. Another interesting clone, which encoded 19 amino acids of a putative 33 amino acid peptide, was obtained with the *S. epidermidis* library. The finding of a homologous peptide in the *S. aureus* genome strengthened our presumption that this was a true open reading frame. Today, these peptides are annotated from the completed genome sequences of *S. epidermidis* strain ATCC 12228 and *S. aureus* strain MW2 (figure 4), and can be detected in all available genome sequences for *S. aureus* and *S. epidermidis* strains. There is no function proposed for the annotated peptides and SignalP predictions result in questionable signal peptides, especially for the *S. aureus* peptide. The original phage display clone was obtained in the second panning cycle from which no background clones were identified. It cannot be considered a functional proof, but it strongly suggests that the peptide has a functional signal peptide and is exported to the staphylococcal cell membrane or to the outside of the cell. It is interesting to speculate that such a molecule could have an important function *e.g.* in quorum sensing or as an anti-microbial agent.

```
S. epidermidis ATCC12228:
MKKLAVILALAGAAFYGFKKYQNHVNQAPNIEY

S. aureus MW2:
MKKLAVILTLVGGLYFAFKKYQERVNQAPNIEY
```

Figure 4. Annotated peptides of *S. epidermidis* and *S. aureus*.

With the SS phage display method, extracellular proteins are isolated by virtue of signal peptide function and concurrently the corresponding partial gene is cloned. This allows for a further elaboration by combining this method with other techniques for *e.g.* gene silencing, as described in the following section.

*An inducible antisense RNA system combined with phage display (**III**)*

In the continuation of the SS phage display project, a way to down-regulate expression of the identified extracellular proteins could provide valuable information on their function. Down-regulation or inhibition of gene expression can be achieved by a variety of techniques, but the use of inducible antisense RNA systems on a genome-wide basis has recently been presented and shown to work quite well in various systems (De Backer *et al.*, 2001; Forsyth *et al.*, 2002; Ji *et al.*,

2001; Wang & Kuramitsu, 2003). It is an attractive approach to reduce gene expression instead of completely abolishing it, as with gene knock-out methods, especially for mutant phenotypes that are lethal. Paper **III** presents the concept of combining SS phage display with an inducible antisense RNA system for down-regulation of gene expression. A *Tn*10 based regulatory system was introduced in the antisense orientation of the E-tag and gIII so that the promoter could be regulated by the presence or non-presence of tetracycline or a derivative like anhydrotetracycline. The *Tn*10 transposable element encodes a repressor, which acts on two operator sequences to regulate the expression of itself as well as another *Tn*10-encoded protein. The repressor dissociates from the operator sequences by the addition of tetracycline. This system has been modified and applied to the phage λ left promoter for a tight regulation of transcription (Lutz & Bujard, 1997). After modification, this element was shown to be compatible with the SS phage display function, and to be induced by anhydrotetracycline on both the transcriptional and translational level. Several screenings to identify antisense RNA-affected clones were performed, such as screening for growth inhibition on standard growth medium and in the presence of high NaCl concentrations as well as screening of E-tag expression. Screenings were performed both on clones selected by the presence of signal peptides and on clones without any prior selection. In this study, no clones affected by antisense RNA transcription were identified.

There are at least four inducible antisense RNA genome libraries reported in the literature, as stated above. Two systems have been applied to *S. aureus* for the identification of essential genes. In (Ji *et al.*, 2001), a *Tn*10-based system combined with a *Bacillus*-derived promoter of the xylose utilisation operon (P_{xyl}) resulted in the identification of about 200 unique antisense RNA-affected genes following screening in *S. aureus*. The other system utilises the xylose-inducible pT5X promoter for the identification of about 650 unique growth affected genes (Forsyth *et al.*, 2002). A heterologous approach for the identification of essential genes common to both Gram-negative and Gram-positive bacteria was conducted by the screening of streptococcal genomic fragments in *E. coli* (Wang & Kuramitsu, 2003). The isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible P_{lac} promoter was used to regulate expression of the streptococcal genes, which resulted in the identification of 9 unique antisense RNA affected, hypothetically universal, essential genes. An antisense RNA and promoter interference combined approach has also been applied in a eukaryotic system for the identification of critical genes for growth of the pathogenic fungus *Candida albicans* (De Backer *et al.*, 2001). In this case, 86 genes were identified after the screening in *C. albicans* of both a genomic and a cDNA library upon activation of the GAL1 promoter in the presence of lithium acetate. Of the total number of clones screened, about 1%, 0,3%, 0,1%, and 4%, respectively, in the libraries presented above, were shown to represent unique genes affected by the induction of antisense RNA transcription. However, when comparing these figures, it must be kept in mind that the strategies used were quite different, *i.e.* it may be deceptive to compare heterologous systems with homologous systems and large-genome eukaryotic systems with prokaryotic systems. Still, with the number of clones screened in our study (for a correct comparison only the unselected clones should be considered) these figures suggest

that our system is less effective, since no antisense RNA affected clones were identified. Modifications of the antisense RNA phage display system to achieve an efficient silencing may include the use of a stronger antisense promoter, perhaps with an even more stringent transcriptional control, and insertion of smaller DNA fragments into the vector. This system may prove to be most effective for the screening of un-selected clones, but the concept of screening selected extracellular proteins offers the most attractive application.

Concluding remarks

Shotgun phage display offers a genome-wide tool for the identification of bacterial ligand binding proteins and characterisation of their binding activity. The results of the study of this thesis show, that this technique can be broadened in different ways. In summary,

- shotgun phage display can be used to investigate the presence and accessibility of bacterial target proteins on *in vivo* coated polymeric materials

- a specialised method for the identification of extracellular proteins was developed (SS phage display) and used for the general isolation of membrane-associated, cell wall-anchored, and secreted proteins, in both Gram-negative and Gram-positive bacteria

- an inducible promoter for antisense RNA transcription can be incorporated into the SS phage display system without impairing the SS phage display property, but the efficiency of the system is too low to allow for down regulation of protein expression

The future perspectives of this study are focused on the continued development of an improved gene silencing system combined with shotgun phage display. Such a system could be very useful for a number of applications but particularly for the investigation of virulence factors of pathogenic bacteria.

References

Anderson, D.M. & Schneewind, O. (1999) *Yersinia enterocolitica* type III secretion: an mRNA signal that couples translation and secretion of YopQ. *Mol Microbiol* **31**: 1139-1148.

Angelichio, M.J. & Camilli, A. (2002) *In vivo* expression technology. *Infect Immun* **70**: 6518-6523.

Ausmees, N., Jacobsson, K. & Lindberg, M. (2001) A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to a family of Rhizobium-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. *trifolii. Microbiology* **147**: 549-559.

Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli. Curr Opin Biotechnol* **10**: 411-421.

Barbucci, R. & Magnani, A. (1994) Conformation of human plasma proteins at polymer surfaces: the effectiveness of surface heparinization. *Biomaterials* **15**: 955-962.

Bass, S., Greene, R. & Wells, J.A. (1990) Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins* **8**: 309-314.

Beckmann, C., Waggoner, J.D., Harris, T.O., Tamura, G.S. & Rubens, C.E. (2002) Identification of novel adhesins from Group B streptococci by use of phage display reveals that C5a peptidase mediates fibronectin binding. *Infect Immun* **70**: 2869-2876.

Berks, B.C., Palmer, T. & Sargent, F. (2003) The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* **47**: 187-254.

Better, M., Chang, C.P., Robinson, R.R. & Horwitz, A.H. (1988) *Escherichia coli* secretion of an active chimeric antibody fragment. *Science* **240**: 1041-1043.

Bjerketorp, J., Nilsson, M., Ljungh, A., Flock, J.I., Jacobsson, K. & Frykberg, L. (2002) A novel von Willebrand factor binding protein expressed by *Staphylococcus aureus. Microbiology* **148**: 2037-2044.

Boden, M.K. & Flock, J.I. (1994) Cloning and characterization of a gene for a 19 kDa fibrinogen-binding protein from *Staphylococcus aureus. Mol Microbiol* **12**: 599-606.

Borchman, D., Harris, E.N., Pierangeli, S.S. & Lamba, O.P. (1995) Interactions and molecular structure of cardiolipin and beta 2-glycoprotein 1 (beta 2-GP1). *Clin Exp Immunol* **102**: 373-378.

Brantl, S. (2002) Antisense-RNA regulation and RNA interference. *Biochim Biophys Acta* **1575**: 15-25.

Broome-Smith, J.K., Gnaneshan, S., Hunt, L.A., Mehraein-Ghomi, F., Hashemzadeh-Bonehi, L., Tadayyon, M. & Hennessey, E.S. (1994) Cleavable signal peptides are rarely found in bacterial cytoplasmic membrane proteins (review). *Mol Membr Biol* **11**: 3-8.

Burns, D.L. (1999) Biochemistry of type IV secretion. *Curr Opin Microbiol* **2**: 25-29.

Carpousis, A.J. (2003) Degradation of targeted mRNAs in *Escherichia coli*: regulation by a small antisense RNA. *Genes Dev* **17**: 2351-2355.

Castillo, J., Goodson, B. & Winter, J. (2001) T7 displayed peptides as targets for selecting peptide specific scFvs from M13 scFv display libraries. *J Immunol Methods* **257**: 117-122.

Chamley, L.W., Duncalf, A.M., Konarkowska, B., Mitchell, M.D. & Johnson, P.M. (1999) Conformationally altered beta 2-glycoprotein I is the antigen for anti-cardiolipin autoantibodies. *Clin Exp Immunol* **115**: 571-576.

Cheung, A.L., Wolz, C., Yeaman, M.R. & Bayer, A.S. (1995) Insertional inactivation of a chromosomal locus that modulates expression of potential virulence determinants in *Staphylococcus aureus*. *J Bacteriol* **177**: 3220-3226.

Cheung, A.L., Bayer, A.S., Zhang, G., Gresham, H. & Xiong, Y.Q. (2004) Regulation of virulence determinants *in vitro* and *in vivo* in *Staphylococcus aureus*. *FEMS Immunol Med Microbiol* **40**: 1-9.

Chung, Y.S. & Dubnau, D. (1998) All seven comG open reading frames are required for DNA binding during transformation of competent *Bacillus subtilis*. *J Bacteriol* **180**: 41-45.

Clarke, S.R., Harris, L.G., Richards, R.G. & Foster, S.J. (2002) Analysis of Ebh, a 1.1-megadalton cell wall-associated fibronectin-binding protein of *Staphylococcus aureus*. *Infect Immun* **70**: 6680-6687.

Click, E.M. & Webster, R.E. (1998) The TolQRA proteins are required for membrane insertion of the major capsid protein of the filamentous phage f1 during infection. *J Bacteriol* **180**: 1723-1728.

Coleman, J., Green, P.J. & Inouye, M. (1984) The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. *Cell* **37**: 429-436.

Crameri, R. & Suter, M. (1993) Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene* **137**: 69-75.

De Backer, M.D., Nelissen, B., Logghe, M., Viaene, J., Loonen, I., Vandoninck, S., de Hoogt, R., Dewaele, S., Simons, F.A., Verhasselt, P., Vanhoof, G., Contreras, R. & Luyten, W.H. (2001) An antisense-based functional genomics approach for identification of genes critical for growth of *Candida albicans*. *Nat Biotechnol* **19**: 235-241.

de Lorenzo, V. & Timmis, K.N. (1994) Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol* **235**: 386-405.

Delihas, N., Rokita, S.E. & Zheng, P. (1997) Natural antisense RNA/target RNA interactions: possible models for antisense oligonucleotide drug design. *Nat Biotechnol* **15**: 751-753.

Dower, W.J. & Mattheakis, L.C. (2002) In vitro selection as a powerful tool for the applied evolution of proteins and peptides. *Curr Opin Chem Biol* **6**: 390-398.

Driessen, A.J.M. & van der Does, C. (2002) Protein export in bacteria. In *Protein targeting, transport & translocation*. (ed. Dalbey, R.E. & von Heijne, G.), 47-73. Academic Press, London, UK.

Dryla, A., Gelbmann, D., von Gabain, A. & Nagy, E. (2003) Identification of a novel iron regulated staphylococcal surface protein with haptoglobin-haemoglobin binding activity. *Mol Microbiol* **49**: 37-53.

Endemann, H. & Model, P. (1995) Location of filamentous phage minor coat proteins in phage and in infected cells. *J Mol Biol* **250**: 496-506.

Engdahl, H.M., Hjalt, T.A. & Wagner, E.G. (1997) A two unit antisense RNA cassette test system for silencing of target genes. *Nucleic Acids Res* **25**: 3218-3227.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenny, K., Sutton, W., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrman, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. & Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496-512.

Flock, J.I. (1999) Extracellular-matrix-binding proteins as targets for the prevention of *Staphylococcus aureus* infections. *Mol Med Today* **5**: 532-537.

Forsyth, R.A., Haselbeck, R.J., Ohlsen, K.L., Yamamoto, R.T., Xu, H., Trawick, J.D., Wall, D., Wang, L., Brown-Driver, V., Froelich, J.M., C, K.G., King, P., McCarthy, M., Malone, C., Misiner, B., Robbins, D., Tan, Z., Zhu Zy, Z.Y., Carr, G., Mosca, D.A., Zamudio, C., Foulkes, J.G. & Zyskind, J.W. (2002) A genome-wide strategy for the identification of essential genes in *Staphylococcus aureus*. *Mol Microbiol* **43**: 1387-1400.

Franch, T. & Gerdes, K. (2000) U-turns and regulatory RNAs. *Curr Opin Microbiol* **3**: 159-164.

Fransen, M., Van Veldhoven, P.P. & Subramani, S. (1999) Identification of peroxisomal proteins by using M13 phage protein VI phage display: molecular evidence that mammalian peroxisomes contain a 2,4-dienoyl-CoA reductase. *Biochem J* **340**: 561-568.

Friedrich, R., Panizzi, P., Fuentes-Prior, P., Richter, K., Verhamme, I., Anderson, P.J., Kawabata, S., Huber, R., Bode, W. & Bock, P.E. (2003) Staphylocoagulase is a prototype for the mechanism of cofactor-induced zymogen activation. *Nature* **425**: 535-539.

Fuquay, J.I., Loo, D.T. & Barnes, D.W. (1986) Binding of *Staphylococcus aureus* by human serum spreading factor in an *in vitro* assay. *Infect Immun* **52**: 714-717.

Gailus, V. & Rasched, I. (1994) The adsorption protein of bacteriophage fd and its neighbour minor coat protein build a structural entity. *Eur J Biochem* **222**: 927-931.

Georgiou, G., Stathopoulos, C., Daugherty, P.S., Nayak, A.R., Iverson, B.L. & Curtiss, R., 3rd (1997) Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* **15**: 29-34.

Gey Van Pittius, N.C., Gamieldien, J., Hide, W., Brown, G.D., Siezen, R.J. & Beyers, A.D. (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biol* **2**: RESEARCH0044.

Glucksman, M.J., Bhattacharjee, S. & Makowski, L. (1992) Three-dimensional structure of a cloning vector. X-ray diffraction studies of filamentous bacteriophage M13 at 7 A resolution. *J Mol Biol* **226**: 455-470.

Gottesman, M.E. & Hendrickson, W.A. (2000) Protein folding and unfolding by *Escherichia coli* chaperones and chaperonins. *Curr Opin Microbiol* **3**: 197-202.

Gray, C.W., Brown, R.S. & Marvin, D.A. (1981) Adsorption complex of filamentous fd virus. *J Mol Biol* **146**: 621-627.

Green, P.J., Mitchell, C. & Inouye, M. (1985) Effect of *mic* gene structure on repressor activity in the OmpA system. *Biochimie* **67**: 763-767.

Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S. & Sherman, D.R. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **51**: 359-370.

Hartleib, J., Kohler, N., Dickinson, R.B., Chhatwal, G.S., Sixma, J.J., Hartford, O.M., Foster, T.J., Peters, G., Kehrel, B.E. & Herrmann, M. (2000) Protein A is the von Willebrand factor binding protein on *Staphylococcus aureus*. *Blood* **96**: 2149-2156.

Hautefort, I. & Hinton, J.C.D. (2002) Molecular methods for monitoring bacterial gene expression during infection. *Molecular Cellular Microbiology* **31**: 55-91.

Heal, K.G., Hill, H.R., Stockley, P.G., Hollingdale, M.R. & Taylor-Robinson, A.W. (1999) Expression and immunogenicity of a liver stage malaria epitope presented as a foreign peptide on the surface of RNA-free MS2 bacteriophage capsids. *Vaccine* **18**: 251-258.

Heilmann, C., Herrmann, M., Kehrel, B.E. & Peters, G. (2002) Platelet-binding domains in 2 fibrinogen-binding proteins of *Staphylococcus aureus* identified by phage display. *J Infect Dis* **186**: 32-39.

Herrmann, M., Vaudaux, P.E., Pittet, D., Auckenthaler, R., Lew, P.D., Schumacher-Perdreau, F., Peters, G. & Waldvogel, F.A. (1988) Fibronectin, fibrinogen, and laminin act as mediators of adherence of clinical staphylococcal isolates to foreign material. *J Infect Dis* **158**: 693-701.

Herrmann, M., Suchard, S.J., Boxer, L.A., Waldvogel, F.A. & Lew, P.D. (1991) Thrombospondin binds to *Staphylococcus aureus* and promotes staphylococcal adherence to surfaces. *Infect Immun* **59**: 279-288.

Herrmann, M., Hartleib, J., Kehrel, B., Montgomery, R.R., Sixma, J.J. & Peters, G. (1997) Interaction of von Willebrand factor with *Staphylococcus aureus*. *J Infect Dis* **176**: 984-991.

Hoess, R.H. (2002) Bacteriophage lambda as a vehicle for peptide and protein display. *Curr Pharm Biotechnol* **3**: 23-28.

Hoogenboom, H.R. (2002) Overview of antibody phage-display technology and its applications. *Methods Mol Biol* **178**: 1-37.

Houshmand, H., Froman, G. & Magnusson, G. (1999) Use of bacteriophage T7 displayed peptides for determination of monoclonal antibody specificity and biosensor analysis of the binding reaction. *Anal Biochem* **268**: 363-370.

Houshmand, H. & Bergqvist, A. (2003) Interaction of hepatitis C virus NS5A with La protein revealed by T7 phage display. *Biochem Biophys Res Commun* **309**: 695-701.

Hussain, M., Becker, K., von Eiff, C., Schrenzel, J., Peters, G. & Herrmann, M. (2001) Identification and characterization of a novel 38.5-kilodalton cell surface protein of *Staphylococcus aureus* with extended-spectrum binding activity for extracellular matrix and plasma proteins. *J Bacteriol* **183**: 6778-6786.

Hutchison, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O. & Venter, J.C. (1999) Global transposon mutagenesis and a minimal Mycoplasma genome. *Science* **286**: 2165-2169.

Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. (1995) Modifying filamentous phage capsid: limits in the size of the major capsid protein. *J Mol Biol* **248**: 835-844.

Izant, J.G. & Weintraub, H. (1984) Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* **36**: 1007-1015.

Jacobson, A. (1972) Role of F pili in the penetration of bacteriophage fl. *J Virol* **10**: 835-843.

Jacobsson, K. & Frykberg, L. (1995) Cloning of ligand-binding domains of bacterial receptors by phage display. *Biotechniques* **18**: 878-885.

Jacobsson, K., Jonsson, H., Lindmark, H., Guss, B., Lindberg, M. & Frykberg, L. (1997) Shot-gun phage display mapping of two streptococcal cell-surface proteins. *Microbiol Res* **152**: 121-128.

Jacobsson, K. & Frykberg, L. (1998) Gene VIII-based, phage-display vectors for selection against complex mixtures of ligands. *Biotechniques* **24**: 294-301.

Jacobsson, K. (2003) A novel family of fibrinogen-binding proteins in *Streptococcus agalactiae*. *Vet Microbiol* **96**: 103-113.

Jespers, L.S., Messens, J.H., De Keyser, A., Eeckhout, D., Van den Brande, I., Gansemans, Y.G., Lauwereys, M.J., Vlasuk, G.P. & Stanssens, P.E. (1995) Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Biotechnology (N Y)* **13**: 378-382.

Ji, Y., Zhang, B., Van, S.F., Horn, Warren, P., Woodnutt, G., Burnham, M.K. & Rosenberg, M. (2001) Identification of critical staphylococcal genes using conditional phenotypes generated by antisense RNA. *Science* **293**: 2266-2269.

Jiang, J., Abu-Shilbayeh, L. & Rao, V.B. (1997) Display of a PorA peptide from *Neisseria meningitidis* on the bacteriophage T4 capsid surface. *Infect Immun* **65**: 4770-4777.

Johansson, J. & Cossart, P. (2003) RNA-mediated control of virulence gene expression in bacterial pathogens. *Trends Microbiol* **11**: 280-285.

Jonsson, K., Signas, C., Muller, H.P. & Lindberg, M. (1991) Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. The complete nucleotide sequence and characterization of the second gene. *Eur J Biochem* **202**: 1041-1048.

Jonsson, K., McDevitt, D., McGavin, M.H., Patti, J.M. & Hook, M. (1995) *Staphylococcus aureus* expresses a major histocompatibility complex class II analog. *J Biol Chem* **270**: 21457-21460.

Kaida, S., Miyata, T., Yoshizawa, Y., Kawabata, S., Morita, T., Igarashi, H. & Iwanaga, S. (1987) Nucleotide sequence of the staphylocoagulase gene: its unique COOH-terminal 8 tandem repeats. *J Biochem (Tokyo)* **102**: 1177-1186.

Kamtekar, S., Schiffer, J.M., Xiong, H., Babik, J.M. & Hecht, M.H. (1993) Protein design by binary patterning of polar and nonpolar amino acids. *Science* **262**: 1680-1685.

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M. & Tabata, S. (2002) Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res* **9**: 189-197.

Kawabata, S., Miyata, T., Morita, T., Iwanaga, S. & Igarashi, H. (1986) The amino acid sequence of the procoagulant- and prothrombin-binding domain isolated from staphylocoagulase. *J Biol Chem* **261**: 527-531.

Kremser, A. & Rasched, I. (1994) The adsorption protein of filamentous phage fd: assignment of its disulfide bridges and identification of the domain incorporated in the coat. *Biochemistry* **33**: 13954-13958.

Kronvall, G. & Jonsson, K. (1999) Receptins: a novel term for an expanding spectrum of natural and engineered microbial proteins with binding properties for mammalian proteins. *J Mol Recognit* **12**: 38-44.

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lian, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Murakami, H., Hosoyama, A., Mizutani-Ui, Y., Takahashi, N.K., Sawano, T., Inoue, R., Kaito, C., Sekimizu, K., Hirakawa, H., Kuhara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. & Hiramatsu, K. (2001) Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* **357**: 1225-1240.

Lee, J.C. (2001) Development of Antistaphylococcal Vaccines. *Curr Infect Dis Rep* **3**: 517-524.

Lindmark, H., Jacobsson, K., Frykberg, L. & Guss, B. (1996) Fibronectin-binding protein of *Streptococcus equi* subsp. *zooepidemicus*. *Infect Immun* **64**: 3993-3999.

Lindmark, H. & Guss, B. (1999) SFS, a novel fibronectin-binding protein from *Streptococcus equi*, inhibits the binding between fibronectin and collagen. *Infect Immun* **67**: 2383-2388.

Lindqvist, B.H. & Naderi, S. (1995) Peptide presentation by bacteriophage P4. *FEMS Microbiol Rev* **17**: 33-39.

Lutz, R. & Bujard, H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**: 1203-1210.

Madden, J.C., Ruiz, N. & Caparon, M. (2001) Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria. *Cell* **104**: 143-152.

Makowski, L. (1992) Terminating a macromolecular helix. Structural model for the minor proteins of bacteriophage M13. *J Mol Biol* **228**: 885-892.

Martoglio, B. & Dobberstein, B. (1998) Signal sequences: more than just greasy peptides. *Trends Cell Biol* **8**: 410-415.

Marvin, D.A., Hale, R.D., Nave, C. & Helmer-Citterich, M. (1994) Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe. *J Mol Biol* **235**: 260-286.

Mazmanian, S.K., Skaar, E.P., Gaspar, A.H., Humayun, M., Gornicki, P., Jelenska, J., Joachmiak, A., Missiakas, D.M. & Schneewind, O. (2003) Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* **299**: 906-909.

McDevitt, D., Nanavaty, T., House-Pompeo, K., Bell, E., Turner, N., McIntire, L., Foster, T. & Hook, M. (1997) Characterization of the interaction between the *Staphylococcus aureus* clumping factor (ClfA) and fibrinogen. *Eur J Biochem* **247**: 416-424.

McNeil, H.P., Simpson, R.J., Chesterman, C.N. & Krilis, S.A. (1990) Anti-phospholipid antibodies are directed against a complex antigen that includes a lipid-binding inhibitor of coagulation: beta 2-glycoprotein I (apolipoprotein H). *Proc Natl Acad Sci U S A* **87**: 4120-4124.

McNeil, H.P., Chesterman, C.N. & Krilis, S.A. (1991) Immunology and clinical importance of antiphospholipid antibodies. *Adv Immunol* **49**: 193-280.

Mohamed, N., Teeters, M.A., Patti, J.M., Hook, M. & Ross, J.M. (1999) Inhibition of *Staphylococcus aureus* adherence to collagen under dynamic conditions. *Infect Immun* **67**: 589-594.

Moller, T., Franch, T., Hojrup, P., Keene, D.R., Bachinger, H.P., Brennan, R.G. & Valentin-Hansen, P. (2002) Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9**: 23-30.

Morfeldt, E., Taylor, D., von Gabain, A. & Arvidson, S. (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAIII. *Embo J* **14**: 4569-4577.

Navarre, W.W. & 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.

Ni Eidhin, D., Perkins, S., Francois, P., Vaudaux, P., Hook, M. & Foster, T.J. (1998) Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of *Staphylococcus aureus*. *Mol Microbiol* **30**: 245-257.

Nilsson, M., Frykberg, L., Flock, J.I., Pei, L., Lindberg, M. & Guss, B. (1998) A fibrinogen-binding protein of *Staphylococcus epidermidis*. *Infect Immun* **66**: 2666-2673.

Nilsson, N., Malmborg, A.C. & Borrebaeck, C.A. (2000) The phage infection process: a functional role for the distal linker region of bacteriophage protein 3. *J Virol* **74**: 4229-4235.

Nilsson, M., Bjerketorp, J., Wiebensjö, A., Ljung, A., Frykberg, L. & Guss, B. (2004) A von Willebrand factor-binding protein from *Staphylococcus lugdunensis*. *FEMS Microbiol Lett*. In press.

Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M. & Nygren, P.A. (1997) Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat Biotechnol* **15**: 772-777.

O'Brien, L.M., Walsh, E.J., Massey, R.C., Peacock, S.J. & Foster, T.J. (2002) *Staphylococcus aureus* clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. *Cell Microbiol* **4**: 759-770.

Pallen, M.J. (2002) The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? *Trends Microbiol* **10**: 209-212.

Pallen, M.J., Chaudhuri, R.R. & Henderson, I.R. (2003) Genomic analysis of secretion systems. *Curr Opin Microbiol* **6**: 519-527.

Park, P.W., Broekelmann, T.J., Mecham, B.R. & Mecham, R.P. (1999) Characterization of the elastin binding domain in the cell-surface 25-kDa elastin-binding protein of *Staphylococcus aureus* (EbpS). *J Biol Chem* **274**: 2845-2850.

Patti, J.M., Jonsson, H., Guss, B., Switalski, L.M., Wiberg, K., Lindberg, M. & Hook, M. (1992) Molecular characterization and expression of a gene encoding a *Staphylococcus aureus* collagen adhesin. *J Biol Chem* **267**: 4766-4772.

Petersen, G., Song, D., Hugle-Dorr, B., Oldenburg, I. & Bautz, E.K. (1995) Mapping of linear epitopes recognized by monoclonal antibodies with gene-fragment phage display libraries. *Mol Gen Genet* **249**: 425-431.

Pohlner, J., Halter, R., Beyreuther, K. & Meyer, T.F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325**: 458-462.

Pugsley, A.P. (1993) The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**: 50-108.

Rakonjac, J., Feng, J. & Model, P. (1999) Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *J Mol Biol* **289**: 1253-1265.

Ren, Z. & Black, L.W. (1998) Phage T4 SOC and HOC display of biologically active, full-length proteins on the viral capsid. *Gene* **215**: 439-444.

Riechmann, L. & Holliger, P. (1997) The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell* **90**: 351-360.

Russel, M. & Model, P. (1985) Thioredoxin is required for filamentous phage assembly. *Proc Natl Acad Sci U S A* **82**: 29-33.

Russel, M. & Model, P. (1989) Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. *J Virol* **63**: 3284-3295.

Russel, M. (1991) Filamentous phage assembly. *Mol Microbiol* **5**: 1607-1613.

Russel, M. & Kazmierczak, B. (1993) Analysis of the structure and subcellular location of filamentous phage pIV. *J Bacteriol* **175**: 3998-4007.

Russel, M. (1995) Moving through the membrane with filamentous phages. *Trends Microbiol* **3**: 223-228.

Russel, M., Linderoth, N.A. & Sali, A. (1997) Filamentous phage assembly: variation on a protein export theme. *Gene* **192**: 23-32.

Sako, T. & Tsuchida, N. (1983) Nucleotide sequence of the staphylokinase gene from *Staphylococcus aureus*. *Nucleic Acids Res* **11**: 7679-7693.

Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J. & Dalbey, R.E. (2000) YidC mediates membrane protein insertion in bacteria. *Nature* **406**: 637-641.

Samuelson, P., Gunneriusson, E., Nygren, P.A. & Stahl, S. (2002) Display of proteins on bacteria. *J Biotechnol* **96**: 129-154.

Santini, C., Brennan, D., Mennuni, C., Hoess, R.H., Nicosia, A., Cortese, R. & Luzzago, A. (1998) Efficient display of an HCV cDNA expression library as C-terminal fusion to the capsid protein D of bacteriophage lambda. *J Mol Biol* **282**: 125-135.

Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W., von Heijne, G., van der Does, C., Driessen, A.J., Oudega, B. & Luirink, J. (2000) YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *Embo J* **19**: 542-549.

Shea, J.E., Santangelo, J.D. & Feldman, R.G. (2000) Signature-tagged mutagenesis in the identification of virulence genes in pathogens. *Curr Opin Microbiol* **3**: 451-458.

Signas, C., Raucci, G., Jonsson, K., Lindgren, P.E., Anantharamaiah, G.M., Hook, M. & Lindberg, M. (1989) Nucleotide sequence of the gene for a fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. *Proc Natl Acad Sci U S A* **86**: 699-703.

Skerra, A. & Pluckthun, A. (1988) Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science* **240**: 1038-1041.

Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**: 1315-1317.

Smith, G.P. & Petrenko, V.A. (1997) Phage Display. *Chem Rev* **97**: 391-410.

Smith, G.P., Patel, S.U., Windass, J.D., Thornton, J.M., Winter, G. & Griffiths, A.D. (1998) Small binding proteins selected from a combinatorial repertoire of knottins displayed on phage. *J Mol Biol* **277**: 317-332.

Stougaard, P., Molin, S. & Nordstrom, K. (1981) RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc Natl Acad Sci U S A* **78**: 6008-6012.

Switalski, L.M., Patti, J.M., Butcher, W., Gristina, A.G., Speziale, P. & Hook, M. (1993) A collagen receptor on *Staphylococcus aureus* strains isolated from patients with septic arthritis mediates adhesion to cartilage. *Mol Microbiol* **7**: 99-107.

Szardenings, M. (2003) Phage display of random peptide libraries: applications, limits, and potential. *J Recept Signal Transduct Res* **23**: 307-349.

Taylor, J.M. & Heinrichs, D.E. (2002) Transferrin binding in *Staphylococcus aureus*: involvement of a cell wall-anchored protein. *Mol Microbiol* **43**: 1603-1614.

Thanabalu, T., Koronakis, E., Hughes, C. & Koronakis, V. (1998) Substrate-induced assembly of a contiguous channel for protein export from *E.coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *Embo J* **17**: 6487-6496.

Thanassi, D.G. & Hultgren, S.J. (2000) Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr Opin Cell Biol* **12**: 420-430.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D., Bron, S. & van Dijl, J.M. (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* **64**: 515-547.

Tomizawa, J., Itoh, T., Selzer, G. & Som, T. (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc Natl Acad Sci U S A* **78**: 1421-1425.

Tung, H., Guss, B., Hellman, U., Persson, L., Rubin, K. & Ryden, C. (2000) A bone sialoprotein-binding protein from *Staphylococcus aureus*: a member of the staphylococcal Sdr family. *Biochem J* **345 Pt 3**: 611-619.

Uhlen, M., Guss, B., Nilsson, B., Gotz, F. & Lindberg, M. (1984) Expression of the gene encoding protein A in *Staphylococcus aureus* and coagulase-negative staphylococci. *J Bacteriol* **159**: 713-719.

Wagner, E.G. & Simons, R.W. (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* **48**: 713-742.

Wagner, E.G. & Brantl, S. (1998) Kissing and RNA stability in antisense control of plasmid replication. *Trends Biochem Sci* **23**: 451-454.

Wall, T., Roos, S., Jacobsson, K., Rosander, A. & Jonsson, H. (2003) Phage display reveals 52 novel extracellular and transmembrane proteins from *Lactobacillus reuteri* DSM 20016(T). *Microbiology* **149**: 3493-3505.

van Wely, K.H., Swaving, J., Freudl, R. & Driessen, A.J. (2001) Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol Rev* **25**: 437-454.

Wang, B. & Kuramitsu, H.K. (2003) Assessment of the utilization of the antisense RNA strategy to identify essential genes in heterologous bacteria. *FEMS Microbiol Lett* **220**: 171-176.

Wang, S.X., Sun, Y.T. & Sui, S.F. (2000) Membrane-induced conformational change in human apolipoprotein H. *Biochem J* **348 Pt 1**: 103-106.

Wann, E.R., Gurusiddappa, S. & Hook, M. (2000) The fibronectin-binding MSCRAMM FnbpA of *Staphylococcus aureus* is a bifunctional protein that also binds to fibrinogen. *J Biol Chem* **275**: 13863-13871.

Varga, L.V., Toth, S., Novak, I. & Falus, A. (1999) Antisense strategies: functions and applications in immunology. *Immunol Lett* **69**: 217-224.

Vasi, J., Frykberg, L., Carlsson, L.E., Lindberg, M. & Guss, B. (2000) M-like proteins of *Streptococcus dysgalactiae*. *Infect Immun* **68**: 294-302.

Vaudaux, P., Suzuki, R., Waldvogel, F.A., Morgenthaler, J.J. & Nydegger, U.E. (1984) Foreign body infection: role of fibronectin as a ligand for the adherence of *Staphylococcus aureus*. *J Infect Dis* **150**: 546-553.

Webster, R.E. (1991) The tol gene products and the import of macromolecules into *Escherichia coli*. *Mol Microbiol* **5**: 1005-1011.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H. & Smith, H.O. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.

Williams, R.J., Henderson, B., Sharp, L.J. & Nair, S.P. (2002) Identification of a fibronectin-binding protein from *Staphylococcus epidermidis*. *Infect Immun* **70**: 6805-6810.

Woiwode, T.F., Haggerty, J.E., Katz, R., Gallop, M.A., Barrett, R.W., Dower, W.J. & Cwirla, S.E. (2003) Synthetic compound libraries displayed on the surface of encoded bacteriophage. *Chem Biol* **10**: 847-858.

von Eiff, C., Peters, G. & Heilmann, C. (2002) Pathogenesis of infections due to coagulase-negative staphylococci. *Lancet Infect Dis* **2**: 677-685.

von Heijne, G. & Abrahmsen, L. (1989) Species-specific variation in signal peptide design. Implications for protein secretion in foreign hosts. *FEBS Lett* **244**: 439-446.

von Heijne, G. (2002) Targeting sequences. In *Protein targeting, transport & translocation*. (ed. Dalbey, R.E. & von Heijne, G.), 35-46. Academic Press, London, UK.

Vroman, L., Adams, A.L., Fischer, G.C. & Munoz, P.C. (1980) Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. *Blood* **55**: 156-159.

Yamamoto, M., Kominato, Y. & Yamamoto, F. (1999) Phage display cDNA cloning of protein with carbohydrate affinity. *Biochem Biophys Res Commun* **255**: 194-199.

Zhang, A., Wassarman, K.M., Ortega, J., Steven, A.C. & Storz, G. (2002) The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9**: 11-22.

Zhang, L., Jacobsson, K., Vasi, J., Lindberg, M. & Frykberg, L. (1998) A second IgG-binding protein in *Staphylococcus aureus*. *Microbiology* **144**: 985-991.

Zhang, L., Jacobsson, K., Strom, K., Lindberg, M. & Frykberg, L. (1999) *Staphylococcus aureus* expresses a cell surface protein that binds both IgG and beta2-glycoprotein I. *Microbiology* **145**: 177-183.

Acknowledgements

The department of microbiology provides an excellent soil in which to grow as a microbiologist, researcher and human being, thanks to the staff and students who all contribute to the friendly and inspiring atmosphere. I extend my sincere gratitude to all those who have supported and helped me during my postgraduate studies, and especially to:

Associate Professor **Lars Frykberg**, my main supervisor, for your continuous guidance and encouragement during these years. It has been a pleasure to be a part of the progressive work that originates from you.

My assisting supervisors Dr **Karin Jacobsson**, for contributing to the progress of my work with your great knowledge and clear-sighted view on things, and Professor **Bengt Guss** for your interest in my work, for your dedicated leadership of our group, and for spreading joy and enthusiasm.

Professor Emeritus **Martin Lindberg**, who initiated my undergraduate student project, without which I have never had this good opportunity to embark upon postgraduate studies. Your constructive criticism of my written work is gratefully acknowledged.

Associate Professor **Christer Signäs** for your prompt assistance on computer matters and your excellent teaching skills during my introductory studies in microbiology.

Joakim Bjerketorp, it has been fun! Thanks for all the nice talks in our room and in the lab, and for being a great person to work with and a good friend. I hope our paths will continue to cross, as they have in the past, in sometimes surprising ways.

My younger fellow PhD students, **Åsa Karlström** and **Jonas Lannergård**, who provided me with insight into the latest research on streptococci, for your support and your contribution to the nice spirit in our group.

The former members of the adhesion group: My room mates **Nora Ausmees**, 'the success story', for initiation of and collaboration on the *B. japonicum* project, and for many pleasant moments together, and **Azar Pooyan**, who turned me into a 'pysslare', for being very supportive during the writing of this thesis and for bringing insight into the Iranian/Persian culture. **Martin Nilsson**, whose humour and interesting real life stories I have missed lately, for many good advice and good times in the lab. **Lihong Zhang,** with whom I worked together on protein Sbi, for a nice and stimulating collaboration. **Josef Vasi** and **Hans Lindmark**, for being very kind to me during my first years as a PhD student.

Dr **Peter Müller** at the Philipps University Marburg, Germany, for our fruitful collaboration on the identification of extracytoplasmic proteins in *B. japonicum*. To widen one's view beyond the familiar grounds is always a rewarding experience.

Torun Wall, **Stefan Roos** and **Hans Jonsson** in the food and feed microbiology group, for showing an interest in my work and involving me as a co-author on the *Lactobacillus reuteri* SS phage display project.

My good friends from the undergraduate studies in agricultural sciences, who have supplied me with information on interesting topics from their own postgraduate studies at various departments. **Sigrid Agenäs**, for all your support during the writing phase of this thesis, particularly for the statuette 'Ångest' which has kept me as sharp-minded as I ever could be (!), but even more for your out-standing friendship during good times and bad. **Annica Pontén**, who put up with me as a lab-partner during all those undergraduate years, for showing me what accuracy means, and that motorcycling is not just for men. **Gabriella Lindgren**, who is a source of inspiration for anyone seeking academic career, and for always inviting me to parties. **Sofia Wretblad**, for nice chats in the corridors of the Genetic Centre on scientific, and more often non-scientific, matters.

Mina vänner utanför universitets trygga murar, ingen nämnd och ingen glömd, för att ni får mig att känna att det jag gör är svårt och viktigt, och för att ni berikar mitt liv på så många olika sätt.

Sist men inte minst, min familj, min trygghet: Mina föräldrar, **Stina** och **Hans Lindén**, för allt ert stöd och för att ni har gett mig ett rikt liv. Mina bröder och förebilder, **Johan**, **Magnus** och **Staffan**: om ni går universitetsvägen så gör jag det också! **Annika** och **Åsa**, för att ni för med er nya, friska vindar till vår familj. Mina svärföräldrar, **Inger** och **Ragnar Rosander**, för att ni har öppnat ert hem för mig och för att ni alltid hjälper oss med stort och smått, inte minst nu under avhandlingsskrivandet. **Pia** och **Patrik**, för alla era vettiga tankar om viktiga saker i livet och för många trevliga och avkopplande stunder hemma hos er. Min älskade **Pär**, som fick mig att framstå som något av en Excel-expert, för att jag alltid känner ditt stöd och din kärlek, och vår son **David** som skänker oss så mycket glädje och lycka.

It was on a short-cut through the hospital kitchens that Albert was first approached by a member of the Antibiotic Resistance.

My contributions to the papers of this thesis have been as follows:

I   Participated in planning of the project and in analysis of the results. Performed the construction of the phage display vector and plasmids, as well as the *S. aureus* library. Performed a major part of the pannings, screening, and sequencing. Main writer of the paper.

II Performed a major part of the construction of the library, the pannings, sequencing, and analysis of the obtained sequences. Participated in writing the paper, mainly the parts concerning the phage display method.

III Participated in planning of the project. Performed the construction of pG3ASUP and pG3ASUP-Spa as well as the phage library. Performed a major part of the investigation of the library and pG3ASUP-Spa. Main writer of the paper.

IV Involved in sequencing of the clones obtained from the initial panning. Performed a major part of the pannings with specific phage stocks against different CVCs and pure ligands. Minor part in writing the paper.

V   This paper contains a general overview of shotgun phage display and I have contributed with the part describing SS phage display.